جامعة سيدي محمد بن عبد الله بفاس
ⵜⴰⵙⴷⴰⵡⵉⵜ ⵙⵉⴷⵉ ⵎⵓⵃⵎⵎⴰⴷ ⴱⵏ ⵄⴱⴷⵍⵍⴰⵀ ⵏ ⴼⴰⵙ
UNIVERSITÉ SIDI MOHAMED BEN ABDELLAH DE FES

كلية العلوم والتقنيات فاس
ⵜⴰⵖⵉⵡⴰⵏⵜ ⵏ ⵜⵎⴰⵙⵙⴰⵏⵉⵏ ⴷ ⵜⵙⵏⵉⴳⵉⵜⵉⵏ
FACULTÉ DES SCIENCES ET TECHNIQUES DE FÈS

N° d'ordre : ……..

Centre des Etudes Doctorales : Sciences et Techniques et Sciences Médicales
Formation Doctorale : SISPMI (Sciences de l'Ingénieur, Sciences Physiques, Mathématiques et Informatique)
Discipline : Sciences et Techniques d'Information et de Communication
Spécialité : Informatique
Laboratoire : Intelligence Artificielle, Sciences de données et Systèmes Émergents (LIASSE)

Thèse pour l'obtention de DOCTORAT

Application des Algorithmes Machine Learning dans les Stratégies E-Marketing

Préparée par : Manal Loukili

Soutenue le 22/12/2023 devant le jury :

Nom et Prénom	Grade	Etablissement	Qualité
Moulhime EL BEKKALI	PES	Ecole Nationale des Sciences Appliquées de Fès	Président
Nour El Houda CHAOUI	PES	Ecole Supérieure de Technologie de Kenitra	Rapporteur
Said AGOUJIL	PES	Ecole Nationale de Commerce et de Gestion de Meknès	Rapporteur
Bouchra FRIKH	PES	Ecole Supérieure de Technologie de Fès	Rapporteur
Mohamed TALIBI ALAOUI	PES	Faculté des Sciences et Techniques de Fès	Examinateur
Abdellatif TAGHZOUTI	PES	Ecole Nationale de Commerce et de Gestion de Fès	Examinateur
Mohammed EL GHAZI	PES	Ecole Supérieure de Technologie de Fès	Co-Directeur de thèse
Fayçal MESSAOUDI	PH	Ecole Nationale de Commerce et de Gestion de Fès	Directeur de thèse

ÉPIGRAPHE



Le machine learning est

le moteur de l'avenir du marketing.

Marc Benioff, PDG de Salesforce

DÉCLARATION

Je déclare par la présente que cette thèse est mon œuvre originale et qu'elle a été entièrement rédigée par mes soins. J'ai dûment mentionné toutes les sources d'information qui ont servi de références bibliographiques dans la thèse.

De plus, cette thèse n'a jamais été soumise à aucun diplôme dans aucune université auparavant.

REMERCIEMENTS

Tout d'abord, je souhaite exprimer ma sincère gratitude envers mes directeurs de thèse, le professeur Fayçal Messaoudi et le professeur Mohammed El Ghazi, pour leur soutien constant, leur patience, leur motivation et leur immense savoir. Cette thèse n'aurait pas été réalisable sans leur encouragement. Ils ont fait preuve d'un professionnalisme exemplaire, d'une disponibilité continue, ont été une source précieuse de conseils avisés et ont également montré un grand sens de l'humour. Ils m'ont guidé dans la voie de la recherche scientifique à long terme et m'ont enseigné les méthodologies nécessaires pour atteindre un niveau élevé de rigueur dans la recherche.

Je tiens à exprimer ma gratitude envers les examinateurs internes et externes de ma thèse qui ont bien accepté d'évaluer et de juger ce travail. Je tiens également à remercier chaleureusement tous les professeurs du laboratoire IASSE. Mes remerciements vont également à mes chers amis de l'École Nationale des Sciences Appliquées et de l'Université Sidi Mohammed Ben Abdellah, qui m'ont soutenu sans hésitation chaque fois que j'en avais besoin.

Je tiens à exprimer ma profonde reconnaissance envers les professeurs des formations doctorales pour leur soutien et leurs conseils pertinents qui m'ont été d'une extrême utilité tout au long de mon parcours de doctorante.

Je souhaite exprimer ma gratitude envers mes parents, Monsieur Charif Omar Loukili et Madame Intissar Bousseta, pour leur amour constant et précieux, ainsi que leur soutien et leur patience inébranlables. Merci de m'avoir toujours poussé à donner le meilleur de moi-même et d'avoir éclairé mon avenir. Merci de m'avoir aidé à devenir la personne que je suis aujourd'hui. Je tiens également à exprimer ma profonde reconnaissance envers mes frères, Riad et Achraf. Les sacrifices qu'ils ont consentis et l'amour qu'ils m'ont prodigué ont été essentiels et déterminants pour tout ce que j'ai accompli au cours de mes études. Je vous aimerai toujours.

Enfin, à ma grand-mère Hania, que ton âme repose en paix, mon amour pour toi est éternel.

Résumé

RÉSUMÉ

Cette thèse a pour objectif d'explorer l'utilisation des algorithmes d'apprentissage automatique (Machine Learning) dans les stratégies de marketing digital. Le marketing digital a connu une évolution rapide au fil des années, passant des méthodes traditionnelles aux techniques numériques avancées. Dans ce contexte, le machine learning se présente comme un outil puissant permettant l'analyse et la génération d'informations précieuses à partir des données massives engendrées par le marketing digital. Cette thèse propose une vue d'ensemble conceptuelle et théorique du marketing digital et du machine learning.

Elle examine les diverses techniques du marketing digital, notamment le marketing des moteurs de recherche, le marketing des médias sociaux, le marketing automatisé, le marketing de contenu, et bien d'autres. De même, elle présente les stratégies du e-marketing, en précisant leurs caractéristiques, leurs composants, ainsi que leurs étapes de développement. En ce qui concerne le machine learning, elle explore en détail les bases théoriques, les types d'apprentissage, ainsi que les méthodes d'évaluation des modèles.

De plus, elle propose une revue de la littérature approfondie concernant l'application du machine learning dans le marketing digital. Elle examine les diverses utilisations du machine learning dans le domaine du e-marketing. Elle présente également les tendances et les évolutions futures dans ce domaine. De plus, elle met en avant les contributions originales de cette thèse à la littérature. Ces contributions sont divisées en plusieurs parties, chacune se focalisant sur une application spécifique du machine learning dans le e-marketing.

Ces parties incluent la prédiction du churn des clients, l'amélioration de la fidélisation des clients grâce au deep learning et aux techniques de gestion des données déséquilibrées, les systèmes de recommandation dans le e-commerce, la prédiction de la fraude liée aux paiements électroniques, l'analyse des sentiments des commentaires sur les produits en vue de recommandations dans le e-commerce, la prédiction des ventes futures grâce à l'utilisation des algorithmes de machine learning, et enfin, l'application du machine learning pour la tarification dynamique dans le e-marketing.

Cette thèse met en évidence le potentiel du machine learning dans l'amélioration des stratégies de e-marketing. Les résultats obtenus démontrent que l'application des algorithmes

de machine learning peut apporter des avantages significatifs dans plusieurs domaines clés du marketing digital.

Mots clés :

Machine learning, marketing digital, analyse de sentiments, gestion de marque, systèmes de recommandation, tarification dynamique, optimisation de la navigation, gestion de la publicité, chatbots, prédiction de la demande, ciblage des clients, engagement des utilisateurs, segmentation du marché, détection de fraudes, expérience client, churn des clients, fidélisation des clients, deep learning, données déséquilibrées, e-commerce, fraude au paiement électronique, prédiction des ventes.

ABSTRACT

The aim of this thesis is to explore the use of machine learning algorithms in digital marketing strategies. Digital marketing has evolved rapidly over the years, moving from traditional methods to advanced digital techniques. In this context, machine learning presents itself as a powerful tool for analyzing and generating valuable insights from the massive data generated by digital marketing. This thesis provides a conceptual and theoretical overview of digital marketing and machine learning.

It examines various digital marketing techniques, including search engine marketing, social media marketing, automated marketing, content marketing, and many others. It also presents e-marketing strategies, outlining their characteristics, components, and development stages. With regard to machine learning, it explores in detail the theoretical foundations, types of learning, and methods for evaluating models.

It also provides an in-depth literature review on the application of machine learning in digital marketing. It examines the various uses of machine learning in e-marketing. It also presents trends and future developments in this field. In addition, it highlights the original contributions of this thesis to the literature. These contributions are divided into several parts, each focusing on a specific application of machine learning in e-marketing.

These parts include the prediction of customer churn, the improvement of customer loyalty through deep learning and unbalanced data management techniques, recommendation systems in e-marketing, the prediction of fraud related to electronic payments, sentiment analysis of product reviews for recommendations in e-marketing, the prediction of future sales through the use of machine learning algorithms, and finally, the application of machine learning for dynamic pricing in e-marketing.

This thesis highlights the potential of machine learning to improve e-marketing strategies. The results obtained demonstrate that the application of machine learning algorithms can bring significant benefits in several key areas of digital marketing.

Keywords:

Machine learning, digital marketing, sentiment analysis, brand management, recommendation systems, dynamic pricing, navigation optimization, advertising management, chatbots, demand prediction, customer targeting, user engagement, market segmentation, fraud detection, customer experience, churn prediction, customer retention, deep learning, imbalanced data, e-commerce, electronic payment fraud, sales prediction.

Production Scientifique

PRODUCTION SCIENTIFIQUE

Articles de Revues :

- **Loukili, M., Messaoudi, F., & El Ghazi, M. (2022).** *Supervised Learning Algorithms for Predicting Customer Churn with Hyperparameter Optimization*. International Journal of Advances in Soft Computing & Its Applications, 14(3), 49-63. doi: 10.15849/IJASCA.221128.04
- **Loukili, M., Messaoudi, F., & El Ghazi, M. (2023).** *Machine Learning Based Recommender System for E-Commerce*. IAES International Journal of Artificial Intelligence (IJ-AI), 12(4), 1803-1811. doi: 10.11591/ijai.v12.i4
- **Loukili, M., Messaoudi, F., & El Ghazi, M. (2023).** *Sentiment Analysis of Product Reviews for E-Commerce Recommendation based on Machine Learning*. International Journal of Advances in Soft Computing & Its Applications, 15(1), 1-13. doi: 10.15849/IJASCA.230320.01
- **Loukili, M., Messaoudi, F., & El Youbi, R. (2023).** *Implementation of Machine Learning Algorithms for Customer Churn Prediction*. Journal of Information Systems and Telecommunication, 11(43), 196-208. doi: 10.52547/jist.34208.11.43.196
- **Loukili, M., Messaoudi, F., & El Ghazi, M. (2024).** *Defending Against Digital Thievery: A Machine Learning Approach to Predict E-Payment Fraud*. International Journal of Management Practice.
- **El Youbi, R., Messaoudi, F., & Loukili, M. (2024).** *Improving E-commerce Personalization using the Apriori Recommender System: A Machine Learning Approach*. International Journal of Advances in Soft Computing & Its Applications, 16(1).
- **Loukili, M., Messaoudi, F., & El Ghazi, M. (2024).** *Enhancing Customer Retention through Deep Learning and Imbalanced Data Techniques*. Iraqi Journal of Science (IJS), 65(5).
- **Messaoudi, F., & Loukili, M. (2024).** *E-Commerce Personalized Recommendations: A Deep Neural Collaborative Filtering Approach*. Operations Research Forum.

Chapitres de Livres :

- **Loukili, M., Messaoudi, F.** *Machine Learning, Deep Neural Network and Natural Language Processing Based Recommendation System*. In: *Kacprzyk, J., Ezziyyani, M., Balas, V.E. (eds) International Conference on Advanced Intelligent Systems for Sustainable Development.* AI2SD 2022. Lecture Notes in Networks and Systems, vol 637. **Springer**, Cham. **2023**. doi: 10.1007/978-3-031-26384-2_7.
- **Loukili, M., Messaoudi, F., El Ghazi, M. & Azirar, H.** *Predicting Future Sales: A Machine Learning Algorithm Showdown*. *Springer Lecture Notes in Networks & Systems, proceeding of the 5th International Conference on Artificial Intelligence and Smart Environments*. ICAISE 2023. **Springer**. **2023**.
- **Loukili, M., Messaoudi, F. & El Ghazi, M.** *Personalizing Product Recommendations using Collaborative Filtering in Online Retail: A Machine Learning Approach*. In: *2023 International Conference on Information Technology (ICIT)*, Amman, Jordan, pp. 19-24, **IEEE Xplore Digital Library**. **2023**. doi: 10.1109/ICIT58056.2023.10226042.
- **Loukili, M., Messaoudi, F. & Azirar, H.** *E-Payment Fraud Prediction using Supervised Learning Algorithms. Advances in Emerging Financial Technology and Digital Money, proceeding of the International Fintech Congress*. IFC 2022. **Taylor & Francis**, (ISBN: 9781032667447), **2023**.
- **Messaoudi, F., Loukili, M. & El Ghazi, M.** *Demand Prediction Using Sequential Deep Learning Model*. In: *2023 International Conference on Information Technology (ICIT),* Amman, Jordan, pp. 577-582, **IEEE Xplore Digital Library**. **2023**. doi: 10.1109/ICIT58056.2023.10225930.
- **Loukili, M., Messaoudi, F., El Ghazi, M. & Azirar, H.** *Leveraging Machine Learning for Dynamic Pricing in Digital Marketing*. In: *New Generation Digital Systems*. NGDS'23. **Springer**. **2023**.
- **El Youbi, R., Messaoudi, F. & Loukili, M.** *Machine Learning-driven Dynamic Pricing Strategies in E-Commerce*. In: *2023 14th International Conference on Information and Communication Systems (ICICS)*, Irbid, Jordan, pp. 1-5, **IEEE Xplore Digital Library**. **2023**. doi: 10.1109/ICICS60529.2023.10330541.

Communications de conférences :

Conférences nationales :

- **CAIMCASE 2021** ; Conférence sur l'intelligence artificielle et les mathématiques pour les défis en sciences appliquées et en ingénierie : « *L'impact du Machine Learning sur l'Expérience Client – Cas de Netflix* ». Agadir, Maroc.
- **CAIMCASE 2021** ; Conférence sur l'intelligence artificielle et les mathématiques pour les défis en sciences appliquées et en ingénierie : « *Artificial Intelligence in Digital Marketing* ». Agadir, Maroc.

Conférences internationales :

- **ICRAMCS 2022** ; Fourth Edition of the International Conference on Research in Applied Mathematics And Computer Science : « *Les Algorithmes Machine Learning pour la Détection de l'Attrition des Clients* ». Casablanca, Maroc.
- **AI2SD 2022** ; The International Conference on Advanced Intelligent Systems for Sustainable Development : « *Machine Learning, Deep Neural Network and Natural Language Processing Based Recommendation System* ». Rabat, Maroc.
- **DEDS 2022** ; The Second Edition of the International Conference on Differential Equations and Dynamical Systems : « *Sales Forecasting Using Recurrent Neural Networks* ». Bizerte, Tunisie.
- **DEDS 2022** ; The Second Edition of the International Conference on Differential Equations and Dynamical Systems : « *Customer Segmentation Using the k-Means Clustering Algorithm* ». Bizerte, Tunisie.
- **ICAISE 2022** ; The 4th International Conference on Artificial Intelligence and Smart Environments : « *Machine Learning Algorithms for Sales Forecasting: Comparative Analysis* ». Errachidia, Maroc.
- **IFC 2022** ; International Fintech Congress : « *E-Payment Fraud Prediction using Supervised Learning Algorithms* ». Marrakech, Maroc.
- **IFC 2022** ; International Fintech Congress : « *Fintech and Financial Inclusion: Mobile Payment Case* ». Marrakech, Maroc.
- **ICIT 2023** ; The 11th International Conference on Information Technology : « *Personalizing*

Product Recommendations using Collaborative Filtering in Online Retail: A Machine Learning Approach ». Amman, Jordanie.

- **ICAISE 2023** ; The 5th International Conference on Artificial Intelligence and Smart Environments : « *Predicting Future Sales: A Machine Learning Algorithm Showdown* ». Errachidia, Maroc.
- **ICIT 2023** ; The 11th International Conference on Information Technology : « *Demand Prediction Using Sequential Deep Learning Model* ». Amman, Jordanie.
- **NGDS 2023** ; 7th International Conference NGNS on New Generation Digital Systems : « *Leveraging Machine Learning for Dynamic Pricing in Digital Marketing* ». Casablanca, Maroc.
- **ICICS 2023** ; The 14th International Conference on Information and Communication Systems : « *Machine Learning-driven Dynamic Pricing Strategies in E-Commerce* ». Irbid, Jordanie.

Journées doctorales :

- **JDPD SIT 2020** ; La quatrième édition de la Journée Doctorale et Postdoctorale des Systèmes d'Information et Télécommunications : « *Comparative Study of the different Machine Learning Algorithms used in Digital Marketing* ». Fès, Maroc.
- **JDPD SIT 2020** ; La quatrième édition de la Journée Doctorale et Postdoctorale des Systèmes d'Information et Télécommunications : « *Applications of Machine Learning in E-commerce* ». Fès, Maroc.
- **JDPD SIT 2022** ; La cinquième édition de la Journée Doctorale et Postdoctorale des Systèmes d'Information et Télécommunications : « *Analyse Comparative des Algorithmes d'Apprentissage Automatique pour la Prévision des Ventes* ». Fès, Maroc.
- **JDPD SIT 2022** ; La cinquième édition de la Journée Doctorale et Postdoctorale des Systèmes d'Information et Télécommunications : « *Le Machine Learning pour la Prédiction du Churn des Consommateurs dans le Secteur des Télécommunications* ». Fès, Maroc.
- **JCUSMBA 2023** ; La première édition des Jeunes Chercheurs de l'USMBA : « *L'Application du Machine Learning dans les Stratégies E-Marketing* ». Fès, Maroc.

TABLE DES MATIÈRES

ÉPIGRAPHE I
DÉCLARATION II
REMERCIEMENTS III
RÉSUMÉ V
ABSTRACT VII
PRODUCTION SCIENTIFIQUE IX
TABLE DES MATIÈRES XIII
LISTE DES FIGURES XX
LISTE DES TABLEAUX XXII
LISTE DES ABRÉVIATIONS XXIV
INTRODUCTION GÉNÉRALE 1
Chapitre 1. Aperçu conceptuel et théorique de l'e-marketing et du machine learning 6
Introduction 6
1. Marketing digital : définitions, histoire, techniques et stratégies 6
1.1. Définitions 6
1.1.1. Marketing 6
1.1.2. Marketing digital 6
1.2. Du marketing au marketing digital 7
1.3. Techniques du marketing digital 8
1.3.1. Marketing des moteurs de recherche (SEM) 8
1.3.2. Marketing des médias sociaux (SMM) 9
1.3.3. Optimisation des médias sociaux (SMO) 10
1.3.4. Marketing automatisé des médias sociaux (SMA) 10
1.3.5. Marketing de contenu 11
1.3.6. Email Marketing 12
1.3.7. Marketing d'influence 12
1.3.8. Publicité native 13
1.3.9. Remarketing 13
1.3.10. Marketing viral 14
1.3.11. Marketing d'affiliation 15
1.3.12. Marketing mobile 15
1.3.13. Publicité display 16
1.3.14. Marketing vidéo 16

1.4. Stratégies e-marketing : définition, caractéristiques et composants, et étapes de développement 16
1.4.1. Définition d'une stratégie e-marketing 16
1.4.2. Caractéristiques et composants d'une stratégie e-marketing 17
1.4.3. Étapes de développement des stratégies e-marketing 18
1.5. Types des stratégies e-marketing 19
1.5.1. Stratégies de marché cible 19
1.5.2. Stratégie de segmentation du marché 19
1.5.3. Stratégie de marketing ciblé 20
1.5.4. Stratégie d'implantation sur le marché 20
1.5.5. Stratégie d'excellence 20
1.5.6. Stratégie d'ajout d'une toute nouvelle industrie 21
1.5.7. Stratégie d'ajout de lignes de production à des produits 21
1.5.8. Stratégies d'expansion des lignes de produits existantes 21
1.5.9. Stratégie de modification des caractéristiques implicites des produits 21
1.5.10. Stratégie de repositionnement du produit sur le marché 21
1.5.11. Stratégie de tarification 21
1.5.11.1. Stratégie de tarification dynamique (Dynamic Pricing) 22
1.5.11.2. Stratégie de différenciation des prix 22
1.5.11.3. Stratégie de prix promotionnels 23
1.5.12. Stratégie de modification des canaux de distribution 25
1.5.13. Stratégie de gestion des relations clients 26
2. Machine Learning : fondements théoriques 26
2.1. Apprentissage supervisé 28
2.1.1. Algorithmes de classification 28
2.1.2. Algorithmes de régression 29
2.1.3. Algorithmes d'étiquetage de séquence 29
2.2. Apprentissage non-supervisé 30
2.2.1. Clustering 30
2.2.2. Réduction de dimensionnalité 31
2.2.3. Détection d'anomalies 32
2.2.4. Modèles génératifs 33
2.2.5. Self-organizing maps 34
2.3. Apprentissage semi-supervisé 35
2.3.1. Méthodes basées sur les graphes 35
2.3.2. Self-learning 36

2.3.3. Co-training ... 37
2.3.4. Multi-view learning ... 37
2.4. Apprentissage par renforcement ... 38
2.4.1. Apprentissage par renforcement basé sur un modèle ... 39
2.4.2. Apprentissage par renforcement sans modèle ... 39
2.4.3. Apprentissage par renforcement hors politique ... 40
2.4.4. Apprentissage par renforcement sur politique ... 40
2.4.5. Apprentissage par renforcement hiérarchique ... 41
2.4.6. Apprentissage par renforcement profond ... 41
2.5. Deep Learning ... 42
2.5.1. Réseaux de neurones convolutionnels ... 43
2.5.2. Réseaux de neurones récurrents ... 43
2.5.3. Autoencodeurs ... 44
2.5.4. Réseaux génératifs antagonistes ... 45
2.6. Transfer learning ... 47
2.6.1. Fine-tuning ... 47
2.6.2. Extraction de caractéristiques ... 48
2.6.3. Multi-task learning ... 48
2.7. Traitement Automatique du Langage Naturel ... 49
3. Méthodes d'évaluation des modèles machine learning ... 50
3.1. Evaluation des algorithmes d'apprentissage supervisé ... 50
3.1.1. Classification ... 50
3.1.2. Régression ... 52
3.2. Evaluation des algorithmes d'apprentissage non supervisé ... 53
3.2.1. Clustering ... 53
3.2.2. Association ... 53
3.3. Evaluation des algorithmes d'apprentissage semi supervisé ... 54
3.4. Evaluation des algorithmes d'apprentissage par renforcement ... 54
Conclusion ... 55
Chapitre 2. Application du machine learning dans le e-marketing, tendances et développements futurs : état de l'art ... 57
Introduction ... 57
1. Revue de littérature portant sur l'application du machine learning dans le e-marketing ... 57
1.1. Produit ... 58
1.1.1. Analyse de sentiments ... 58
1.1.2. Gestion de marque ... 61

1.1.3. Systèmes de recommandation ... 62
1.2. Prix ... 65
1.2.1. Tarification dynamique ... 65
1.2.2. Segmentation des prix ... 66
1.3. Place ... 67
1.3.1. Optimisation de la navigation sur les sites web ... 67
1.3.2. Optimisation du référencement ... 68
1.4. Promotion ... 69
1.4.1. Gestion de publicité ... 69
1.4.2. Chatbots ... 70
1.4.3. Prédiction de la demande ... 72
1.5. Personnes ... 75
1.5.1. Prédiction du désabonnement des clients ... 75
1.5.2. Ciblage des clients ... 78
1.5.3. Engagement ... 79
1.6. Processus ... 80
1.6.1. Segmentation du marché ... 80
1.6.2. Détection de fraudes ... 83
1.6.3. Amélioration de l'expérience client ... 84
1.7. Preuve physique ... 85
1.7.1. Qualité du contenu ... 85
1.7.2. Réputation en ligne ... 86
2. Tendances et Développements Futurs du Machine Learning dans le Marketing Digital ... 87
2.1. Traitement du langage naturel (NLP) ... 87
2.2. Réseaux adversoriels génératifs ... 88
2.3. Hyperpersonnalisation ... 89
2.4. Analyse prédictive ... 89
2.5. Optimisation de la recherche vocale ... 90
2.6. Enchères automatisées ... 91
2.7. Vision par ordinateur ... 92
2.8. Réalité augmentée ... 93
2.9. Edge Computing ... 94
Conclusion ... 95
Chapitre 3. Contributions à la littérature ... 98
Introduction ... 98

Partie I. Le machine learning pour la prédiction du churn des clients avec optimisation des hyperparamètres 98

1. Contexte 98

2. Matériaux 99

2.1. Description du problème 99

2.2. Description de l'ensemble de données 100

3. Modèles de classification d'apprentissage automatique utilisés 101

4. Méthodologie 103

4.1. Prétraitement des données 103

4.2. Ensembles d'entraînement et de test 107

4.3. Mesures de performance du modèle 108

4.4. Optimisation des hyperparamètres 109

5. Evaluation et la comparaison des modèles 109

6. Synthèse 112

Partie II. Amélioration de la fidélisation des clients grâce au deep learning et aux techniques de données déséquilibrées 113

1. Contexte 113

2. Méthodologie 114

2.1. Prétraitement des données 114

2.2. Modèle prédictif 117

2.3. Entraînement 118

2.4. Évaluation 119

2.5. Traitement des données déséquilibrées 120

2.6. Comparaison des modèles 122

3. Résultats et discussion 123

4. Synthèse 125

Partie III. Le machine learning pour les systèmes de recommandation dans le e-commerce 126

1. Contexte 126

2. Aperçu des systèmes de recommandation 127

2.1. Filtrage collaboratif 128

2.2. Réseau de confiance 128

2.3. Modèle latent 129

2.4. Règles d'association 130

3. Méthodologie 132

3.1. Visualisation des données 133

3.2. Prétraitement des données 133

3.3. Modélisation des règles d'association 134
4. Évaluation du modèle et résultats 135
5. Synthèse 136
Partie IV. Se défendre contre le vol numérique : une approche machine learning pour prédire la fraude au paiement électronique 137
1. Contexte 137
2. Fondements de l'apprentissage automatique pour la détection de la fraude 140
2.1. Systèmes de détection de fraude basés sur l'apprentissage automatique 140
2.2. Fonctionnement des systèmes de détection de fraude basés sur l'apprentissage automatique 141
2.3. Algorithmes d'apprentissage automatique pour la prédiction de paiements frauduleux 141
3. Méthodologie 143
3.1. Base de données 144
3.2. Prétraitement des données 145
3.3. Ingénierie des fonctionnalités 145
3.4. Encodage à chaud (one-hot) 146
3.5. Fractionnement des données 146
3.6. Suréchantillonnage 146
3.7. Évaluation du modèle 146
4. Résultats et discussion 147
4.1. Résultats 147
4.2. Discussion 148
5. Synthèse 149
Partie V. Analyse des sentiments à travers les commentaires sur les produits pour la recommandation dans le e-commerce : une approche machine learning 151
1. Contexte 151
2. Méthodologie 152
2.1. Description de l'ensemble de données 152
2.2. Prétraitement et exploration des données 153
2.3. Analyse de la polarité 154
2.4. Prétraitement du texte 155
2.5. Mots vides, stemming et lemmatisation 156
2.6. Analyse du nuage de mots 156
2.7. Modèles prédictifs 158
3. Résultats expérimentaux et discussion 158
3.1. Matrice de confusion 158

3.2. Indicateurs de performance 159
3.3. Courbe AUC-ROC 160
3.4. Discussion 161
4. Synthèse 162
Partie VI. Prédire les ventes futures en utilisant les algorithmes machine learning 163
1. Contexte 163
2. Méthodologie 164
2.1. Description et visualisation des données 165
2.2. Analyse exploratoire des données 166
2.3. Détermination de la stationnarité 167
2.4. Observation des retards 168
2.5. Modélisation régressive de l'ensemble de données 169
2.6. Résultats de la prévision des ventes 169
3. Comparaison des modèles et résultats 171
4. Synthèse 173
Partie VII. Le machine learning au service de la tarification dynamique dans le e-marketing 174
1. Contexte 174
2. Cadre théorique 175
3. Méthodologie 176
3.1. Dataset 176
3.2. Prétraitement des données 178
3.3. Ingénierie des caractéristiques 178
3.4. Réglage des hyperparamètres 179
3.5. Architecture du modèle d'apprentissage par renforcement 179
3.6. Entraînement du modèle 181
4. Evaluation du modèle et résultats 181
5. Synthèse 183
Conclusion 184
Conclusion Générale 186
Bibliographie 190

LISTE DES FIGURES

Figure 1. « Les 7P de l'e-marketing ». 57

Figure 2. « Le pipeline de l'étude de prédiction du churn des clients en utilisant les algorithmes machine learning ». 103

Figure 3. « Répartition des clients résiliants et non résiliants ». 105

Figure 4. « Répartition des variables "TotalCharges", "MonthlyCharges" et "tenure" par rapport à la variable "Churn" ». 105

Figure 5. « La répartition des variables "Partner", "Dependents", "SeniorCitizen" et "Contract" par rapport à la variable "Churn" ». 106

Figure 6. « Le nombre de services par rapport à la variable "Churn" ». 107

Figure 7. « La corrélation entre la variable cible et les autres variables ». 107

Figure 8. « Les étapes de la méthodologie adoptée pour la prédiction du désabonnement des clients en utilisant les réseaux neuronaux profonds et les techniques de données déséquilibrées ». 114

Figure 9. « Observation de l'ensemble des variables discrètes par rapport à la variable cible ''Exited''. . 116

Figure 10. « La distribution des variables continues par rapport à la variable cible ». 116

Figure 11. « Architecture du modèle Deep Neural Network ». 118

Figure 12. « Classes déséquilibrées de la variable cible ''Exited'' ». 120

Figure 13. « Processus du système de recommandation proposé ». 133

Figure 14. « Ventes au détail du commerce électronique dans le monde ». 137

Figure 15. « Architecture du système de paiement électronique ». 138

Figure 16. « Flowchart de l'étude de prédiction de paiements frauduleux ». 144

Figure 17. « Normalisation des montants des transactions ». 145

Figure 18. « Le pipeline de l'étude de prédiction des recommandations des clients ». 152

Figure 19. « Proportions des produits recommandés et non recommandés ». 153

Figure 20. « Variable cible "Recommend_IND" en fonction de la longueur du texte ». 153

Figure 21. « Distribution des évaluations des produits ». 154

Figure 22. « La variable cible "Recommend_IND" par rapport à l'évaluation du produit ». 154

Figure 23. « La polarité du texte des commentaires des clients ». 155

Figure 24. « Les sentiments par rapport à la variable cible "Recommend_IND" en pourcentage ». 156

Figure 25. « Nuage de mots positifs relatif au commentaires des clients ». 157

Figure 26. « Nuage de mots négatifs relatif au commentaires des clients ». 157

Figure 27. « Implémentation des modèles de classification (script) ». 158

Figure 28. « Courbe AUC-ROC des modèles de prédiction des sentiments des commentaires des clients ». 161
Figure 29. « Architecture de la méthodologie adoptée pour la prévision des ventes ». 165
Figure 30. « Distribution des ventes par jour et des ventes totales par magasin ». 167
Figure 31. « Graphique des ventes mensuelles ». 167
Figure 32. « Graphique des ventes mensuelles après la transformation de différenciation ». 168
Figure 33. « Graphique d'autocorrélation ». 168
Figure 34. « Graphique d'autocorrélation partielle ». 169
Figure 35. « Graphique de prévision des ventes pour le modèle RF ». 170
Figure 36. « Graphique de prévision des ventes pour le modèle XGBoost ». 170
Figure 37. « Graphique de prévision des ventes pour le modèle Linear Regression ». 171
Figure 38. « Graphique de prévision des ventes pour le modèle LSTM ». 171
Figure 39. « Étapes de l'étude de la tarification dynamique ». 176
Figure 40. « Architecture du modèle d'apprentissage par renforcement ». 180

LISTE DES TABLEAUX

Tableau 1. « L'ensemble de données du churn des clients ».....100
Tableau 2. « Liste des caractéristiques du dataset utilisé ». 101
Tableau 3. « Résultats de vérification des données manquantes ». 103
Tableau 4. « Matrice de confusion ». 108
Tableau 5. « Matrices de confusion des modèles proposés ». 109
Tableau 6. « Mesures de performance du modèle SVM ». 110
Tableau 7. « Mesures de performance du modèle k-NN ».....110
Tableau 8. « Mesures de performance du modèle Log. R ». 110
Tableau 9. « Mesures de performance du modèle RF ».....111
Tableau 10. « Mesures de performance des modèles proposés avant le réglage des hyperparamètres ». 111
Tableau 11. « Mesures de performance des modèles proposés après le réglage des hyperparamètres ». 111
Tableau 12. « L'ensemble de données utilisé pour la prédiction du désabonnement des clients ».....115
Tableau 13. « La matrice de corrélation ».....117
Tableau 14. « Évaluation du modèle DNN sans techniques d'échantillonnage des données ».....123
Tableau 15. « Évaluation du modèle DNN avec différentes techniques d'échantillonnage des données ». 123
Tableau 16. « Le dataset transactionnel utilisé pour la recommandation ».....133
Tableau 17. « L'ensemble des données relatives à tous les produits achetés par chaque client ». 134
Tableau 18. « Ensemble de données sur les règles d'association générées par le modèle ».....134
Tableau 19. « Ensemble de données de règles d'association appliquées aux paniers des clients ».....135
Tableau 20. « Évaluation du système de recommandation ».....135
Tableau 21. « Description du dataset relatif à la détection des payements frauduleux ».....144
Tableau 22. « Matrice de confusion des modèles utilisés pour la détection de fraude ».....147
Tableau 23. « Évaluation du modèle XGBoost ».....147
Tableau 24. « Évaluation du modèle CatBoost ». 148
Tableau 25. « Évaluation du modèle AdaBoost ». 148
Tableau 26. « Dataset sur les commentaires et les évaluations des clients ». 152
Tableau 27. « Dataset sur la polarité et le sentiment dans les commentaires des clients ».....155
Tableau 28. « Dataset contenant le texte des commentaires et la variable cible ». 157

Tableau 29. « Matrices de confusion des modèles utilisés pour la prédiction des sentiments des commentaires des clients ». 159
Tableau 30. « Valeurs d'exactitude, de précision, de rappel et de score f1 des modèles utilisés ». 160
Tableau 31. « Les fonctions utilisées pour la prédiction des ventes ». 165
Tableau 32. (a) « Data frame des ventes quotidiennes par magasin », (b) « Data frame des ventes mensuelles totales ».....166
Tableau 33. « Data frame des ventes mensuelles totales ». 169
Tableau 34. « Valeurs de RMSE, MAE et R^2 pour chaque modèle ».....172
Tableau 35. « Dataset utilisé pour la tarification dynamique ».....177
Tableau 36. « Résultats du modèle DQN par rapport à la stratégie de référence ».....182

LISTE DES ABRÉVIATIONS

A/B : Test A/B
A3C : Asynchronous Advantage Actor-critic
ACP : Analyse en Composantes Principales
ADASYN : Adaptive Synthetic
AIDA : Attention, Intérêt, Désir, Action
AIS : Artificial Immune Systems
AR : Réalité augmentée
AUC : Area Under the Curve
BPTT : Backpropagation through Time
CCA : Canonical Correlation Analysis (Analyse de corrélation canonique)
CNN : Convolutional Neural Networks (Réseaux de neurones convolutionnels)
CNN-LSTM : Convolutional Neural Network Long Short-Term Memory
CRF : Conditional Random Fields
CTR : Click-Through Rate (Taux de clics)
DBSCAN : Density-Based Spatial Clustering of Applications with Noise
DDQN : Double DQN
DDPG : Deep Deterministic Policy Gradients
DQN : Deep Q-Network
E-marketing : marketing digital ou marketing numérique
EM : Expectation-Maximization
eWOM : Electronic Word of Mouth (Bouche-à-oreille électronique)
FN : Faux Négatif
FP : Faux Positif
FP-Growth : Algorithme Frequent Pattern Growth
GAN : Generative Adversarial Networks (Réseaux génératifs antagonistes)
GRU : Gated Recurrent Unit (Un type de réseau de neurones récurrents)
GSOM : Growing Self-Organizing Map
HMM : Hidden Markov Models
IA : Intelligence Artificielle
IWGAN : Improved Wasserstein GAN (Une amélioration des WGAN)
k-NN : K-Nearest Neighbor (k-Le plus proche voisin)
LDA : Linear Discriminant Analysis
LOF : Local Anomaly Factor
Log R : Logistic Regression (Régression logistique)
LSGAN : Least squares GAN (Une variante des GAN)
LSTM : Long Short-Term Memory (Mémoire à long terme - Mémoire à court terme)
MAE : Mean Absolute Error (Erreur absolue moyenne)
MDP : Markov Decision Process (Processus de décision markovien)
ML : Machine Learning
MLP : Multilayer Perceptron
MSE : Mean Squared Error (Erreur quadratique moyenne)
POS : Part-of-Speech (Partie du discours)

PPO : Proximal Policy Optimization
Q-learning : Algorithme d'apprentissage par renforcement
R^2 : Coefficient de Détermination
RF : Random Forest
RFM : Récence, Fréquence, Monétaire
RMSE : Root Mean Square Error (La racine de l'erreur quadratique moyenne)
RNN : Recurrent Neural Networks
SARSA : State-Action-Reward-State-Action
SGD : Stochastic Gradient Descent (Une technique d'optimisation)
SMA : Social Media Automation (Marketing automatisé des médias sociaux)
SMM : Social Media Marketing (Marketing des médias sociaux)
SMO : Social Media Optimization (Optimisation des médias sociaux)
SMOTE : Synthetic Minority Oversampling Technique
SMS : Short Message Service
SOM : Self-Organizing Maps
SVD : Singular Value Decomposition
SVM : Support Vector Machine (Machine à vecteurs de support)
SVM-SMOTE : Support Vector Machine SMOTE
t-SNE : t-distributed Stochastic Neighbor Embedding
TBPTT : Truncated backpropagation through time (Une variante de BPTT)
TF-IDF : Term Frequency-Inverse Document Frequency
URL : Uniform Resource Locator
VAE : Variational AutoEncoder
TN : Vrai Négatif
TP : Vrai Positif
VR : Réalité virtuelle

Introduction Générale

INTRODUCTION GÉNÉRALE

Dans un monde en constante effervescence, où les entreprises se livrent une compétition acharnée pour se démarquer au sein de l'immensité de l'univers en ligne, émerge un nouvel art, une symphonie digitale d'une créativité sans limites : le marketing digital. Tel un chef d'orchestre, il guide les entreprises vers de nouveaux sommets, leur offrant des opportunités insoupçonnées pour comprendre leurs clients, améliorer leur expérience en ligne et élaborer des stratégies commerciales innovantes.

Au cœur de cette révolution, l'intelligence artificielle (IA) et l'apprentissage automatique (Machine Learning, ML) jouent le rôle de virtuoses accompagnateurs, offrant des harmonies technologiques inédites, tandis que le marketing numérique transcende les frontières traditionnelles du commerce, redéfinissant notre vision du monde et la gestion des affaires.

Ces dernières années ont été marquées par une croissance rapide du marketing digital, également désigné sous les termes marketing numérique ou e-marketing, ainsi que de ses aspects au sein de la nouvelle économie mondiale (López García et al., 2019). Cette évolution a profondément transformé notre perception du monde et la gestion des entreprises. Cependant, avant d'expliquer l'importance et les objectifs du marketing digital dans le présent document, il est essentiel de définir les notions de "marketing" et de "digital".

Le terme "digital" se réfère à l'utilisation ou au stockage de données ou d'informations sous forme de signaux numériques, impliquant l'utilisation des technologies informatiques (Khan et al., 2022). Quant au terme "marketing", il englobe les actions de promotion et de commercialisation de produits ou de services, y compris les analyses de marché et la publicité (Infante et Mardikaningsih, 2022).

Cependant, dans cette thèse, la définition centrale se concentre sur le marketing digital, qui englobe le marketing exploitant Internet et les technologies digitales connectées, y compris les smartphones, les ordinateurs, ainsi que tout autre média et plate-forme digitale, dans le but de promouvoir des produits et des services (Fahmi et al., 2022).

L'intelligence artificielle (IA) est couramment définie comme un ensemble de théories et de techniques visant à développer des programmes informatiques complexes capables de simuler certaines caractéristiques de l'intelligence humaine, telles que le raisonnement et l'apprentissage (Zhang et al., 2023). Dans le domaine du marketing digital, l'IA joue un rôle

crucial, en mettant en avant deux notions clés : le Big Data et le Machine Learning (George et al., 2023).

Le Machine Learning, ou apprentissage automatique, représente l'une des technologies d'IA les plus répandues pour améliorer les performances de l'e-marketing (Sadeq et al., 2023). Il repose sur l'idée de donner aux machines un accès aux données et de les laisser apprendre par elles-mêmes. En d'autres termes, le Machine Learning peut être défini comme un programme permettant à une machine d'apprendre de manière automatique, lui permettant ainsi d'accomplir plusieurs tâches, particulièrement complexes, sans qu'elle soit explicitement programmée pour le faire (Janiesch et al., 2021).

Le Big Data joue un rôle essentiel dans le domaine du Machine Learning. Il permet d'accélérer la courbe d'apprentissage d'un système de Machine Learning : plus il reçoit de données, plus il apprend, ce qui le rend plus précis (Al-Jarrah et al., 2015). Le Big Data est devenu un atout marketing incontournable (Verhoef et al., 2016).

Ces dernières années, l'IA a suscité un vif intérêt dans de nombreuses phases du e-marketing (Kumanova-Larde et Hristov, 2019). Le concept d'intelligence artificielle a été introduit pour la première fois dans le marketing en 1998 à l'Université de Columbia dans le contexte des bibliothèques numériques (O'Halloran et Nowaczyk, 2019). À cette époque, de tels algorithmes étaient utilisés pour fournir des recommandations simples aux clients en fonction de ce que l'algorithme "supposait" que le public apprécierait. En 2013, les marketeurs ont commencé à reconnaître l'IA comme un outil d'analyse de données dans le e-marketing pour prendre des décisions basées sur les données (Saura, 2021).

Actuellement, le domaine du e-marketing requiert une manipulation plus précise et rapide des vastes volumes de données, et il convient de noter que les algorithmes d'apprentissage automatique ont connu des avancées significatives (Xie et He, 2022). En conséquence, l'objectif de cette thèse est d'explorer les possibilités d'intégration des techniques d'apprentissage automatique dans le domaine du e-marketing. Cette démarche revêt une importance particulière en raison de la demande croissante de traitement de quantités massives de données dans le domaine du e-marketing. De plus, l'impact potentiel des algorithmes d'apprentissage automatique sur les processus du e-marketing, tels que la planification, la mise en œuvre, la mesure et l'optimisation, mérite une réflexion approfondie, car il offre des perspectives d'efficacité accrue grâce à l'automatisation.

Cette thèse se propose d'explorer en profondeur l'application du machine learning dans les stratégies de e-marketing, en examinant des questions de recherche cruciales, telles que celles posées ci-dessous :

• *Comment le machine learning peut-il être mobilisé pour anticiper et prévenir le churn des clients dans le contexte du marketing digital, et quelles sont les méthodologies les plus pertinentes à cet égard ?*

• *Dans quelle mesure les techniques de deep learning et de gestion des données déséquilibrées peuvent-elles être exploitées pour renforcer la fidélisation des clients dans le domaine du marketing digital ?*

• *Quelles approches basées sur le machine learning pour les systèmes de recommandation peuvent être mises en œuvre pour optimiser les stratégies de vente dans le secteur du e-commerce ?*

• *De quelle manière l'analyse des sentiments des clients, reposant sur des algorithmes machine learning, peut-elle contribuer à une meilleure appréhension des préférences et des besoins des clients dans le domaine du marketing digital ?*

• *En quoi l'usage du machine learning dans le domaine du marketing digital peut-il concourir à la prédiction des ventes et de la demande ?*

• *Quels défis majeurs se posent et quelles méthodologies optimales peuvent être employées pour détecter la fraude dans les transactions électroniques en ayant recours aux techniques du machine learning dans le cadre du marketing digital ?*

• *Comment le machine learning peut-il être mis en application pour la tarification dynamique dans le domaine du marketing digital, en vue d'optimiser les stratégies tarifaires et de maximiser les revenus ?*

Cette thèse adopte l'organisation suivante :

Le premier chapitre s'attèle à fournir un aperçu conceptuel et théorique du e-marketing et du machine learning. Il s'articule autour de l'exposition des concepts fondamentaux du marketing digital et du machine learning, en mettant en exergue leur pertinence au sein du contexte du marketing en ligne. Des définitions précises du e-marketing et des concepts connexes sont présentées, accompagnées d'une explication détaillée des techniques et approches du

machine learning. Ce chapitre établit les bases nécessaires à la compréhension des implications du machine learning dans les stratégies du marketing digital.

Dans le deuxième chapitre, une revue de littérature exhaustive est entreprise afin d'explorer les recherches antérieures consacrées à l'application du machine learning dans le domaine du e-marketing. Les travaux pertinents sont scrutés en vue de mettre en lumière les différentes méthodes et techniques du machine learning employées dans les stratégies de marketing digital. De surcroît, ce chapitre scrute les tendances actuelles et les développements futurs dans ce domaine, en offrant un aperçu des avancées récentes et des perspectives prometteuses.

Le troisième chapitre expose les contributions originales de la recherche. Il traite des questions de recherche précédemment identifiées et offre des analyses détaillées accompagnées de solutions pratiques. Les diverses applications du machine learning dans les stratégies de marketing digital sont passées en revue, mettant en évidence les défis spécifiques et les meilleures pratiques associées à leur mise en œuvre. Les résultats des études menées sont exposés, et des recommandations pertinentes sont formulées dans le but d'aider les professionnels du marketing à optimiser leurs performances grâce à l'intégration du machine learning.

Enfin, une conclusion générale synthétise les principaux apports et contributions de cette thèse et leurs implications pratiques dans la recherche. Elle met en relief l'importance grandissante du machine learning au sein des stratégies du marketing digital, tout en soulignant les avantages potentiels pour les entreprises. Des perspectives de recherche future sont également évoquées pour encourager davantage le développement et l'application du machine learning dans le domaine du marketing en ligne.

Chapitre 1

Aperçu conceptuel et théorique de l'e-marketing et du machine learning

Sommaire

INTRODUCTION 6
MARKETING DIGITAL : DEFINITIONS, HISTOIRE, TECHNIQUES ET STRATEGIES 6
MACHINE LEARNING : FONDEMENTS THEORIQUES 26
METHODES D'EVALUATION DES MODELES MACHINE LEARNING 50
CONCLUSION 55

Chapitre 1. Aperçu conceptuel et théorique de l'e-marketing et du machine learning

Ce chapitre propose un aperçu conceptuel et théorique du e-marketing et du machine learning. Il constitue une introduction essentielle pour comprendre les fondements et les principes de ces deux domaines qui jouent un rôle central dans le paysage numérique d'aujourd'hui. De plus, il établit les bases théoriques et conceptuelles nécessaires pour comprendre en profondeur le e-marketing et le machine learning. Il offre une vision d'ensemble des principes fondamentaux et des concepts clés qui seront explorés plus en détail dans les chapitres suivants de ce manuscrit.

1. Marketing digital : définitions, histoire, techniques et stratégies

1.1. Définitions

1.1.1. Marketing

Le marketing est un terme anglo-saxon introduit en France dans les années 1950 et se définit comme l'ensemble de ce qui contribue à attirer le plus grand nombre de nouveaux clients et à les maintenir en leur fidélisant (Le Meunier-FitzHugh, 2021). Il représente généralement les différents mécanismes que l'entreprise utilise effectivement pour vendre ses produits (Zouaoui et Allouche, 2023).

Le marketing désigne toutes les activités qu'une entreprise utilise pour promouvoir ses produits et services et améliorer sa part de marché (Buttle, 1996). Pour réussir, le marketing requiert une combinaison de savoir-faire en matière de publicité, de vente et de capacité à fournir des biens aux utilisateurs finaux (Srivastava et al., 1999). Cette activité est normalement menée par des professionnels spécifiques ou des spécialistes du marketing qui peuvent travailler en interne ou en externe avec d'autres sociétés de marketing.

1.1.2. Marketing digital

Le terme marketing digital, également connu sous le nom de marketing en ligne ou marketing numérique, désigne l'utilisation des canaux numériques pour promouvoir et commercialiser des produits et des services dans le but d'atteindre les consommateurs (Peter et Dalla Vecchia, 2021). Ces canaux numériques comprennent les sites web, les appareils mobiles, les médias sociaux, les moteurs de recherche et d'autres plateformes en ligne similaires (Viglia et

Viglia, 2014). L'émergence de l'internet dans les années 1990 a joué un rôle clé dans le développement et la popularisation du marketing digital (Ngai, 2003).

Le marketing digital repose sur de nombreux principes similaires à ceux du marketing traditionnel et est souvent considéré comme une nouvelle approche permettant aux entreprises de toucher les consommateurs et de comprendre leur comportement (Chaffey et Smith, 2013). De nombreuses entreprises intègrent désormais des techniques de marketing traditionnel et numérique dans leurs stratégies globales (Wind et Mahajan, 2002). Cela leur permet de tirer parti des avantages spécifiques offerts par chaque canal et de maximiser leur portée et leur impact auprès du public cible.

Dans un environnement où les consommateurs sont de plus en plus connectés et où les technologies numériques continuent de se développer, le marketing digital joue un rôle essentiel pour les entreprises qui souhaitent rester compétitives et atteindre leur public de manière efficace et pertinente.

1.2. Du marketing au marketing digital

L'évolution du marketing a été façonnée par les avancées technologiques et les besoins changeants des entreprises et des consommateurs (Jamshido'g'li et al., 2020). Tout a commencé dans les civilisations anciennes, lorsque les marchands faisaient la promotion de leurs produits par le bouche-à-oreille ou en les exposant dans des lieux publics. Au fur et à mesure que la société et la technologie ont progressé, le domaine du marketing a également évolué (Brooks, 1980).

Un développement significatif a été l'introduction de publicités imprimées dans les journaux et les magazines. Cela a permis aux entreprises de toucher un public plus large et de cibler des segments démographiques spécifiques grâce à l'utilisation de publications spécialisées (Turow, 2005).

À mesure que le XXe siècle avançait, l'introduction de la radio et de la télévision a révolutionné le domaine de la publicité. Les entreprises pouvaient désormais toucher un public de masse grâce à la publicité radiodiffusée, et les spots télévisés sont devenus un élément essentiel du marketing moderne (Turow, 1997).

À la fin du XXe siècle, le marketing par publipostage est apparu comme un moyen populaire pour les entreprises d'envoyer des matériaux promotionnels directement aux foyers des

consommateurs (Turow, 2008). Cela a permis aux entreprises de cibler des individus ou des foyers spécifiques avec des efforts de marketing personnalisés.

Avec l'adoption généralisée d'Internet à la fin du XXe et au début du XXIe siècles, les entreprises ont commencé à utiliser le courrier électronique et les sites web comme nouveaux canaux pour toucher et engager les consommateurs (Applegate et al., 2006). L'émergence des plateformes de médias sociaux au début du XXIe siècle a offert encore plus d'opportunités aux entreprises de se connecter avec leurs clients en ligne (Hanna et al., 2011).

Ces dernières années, l'adoption d'algorithmes machine learning a encore transformé le domaine du marketing (Fallucchi et al., 2020). Ces algorithmes permettent aux entreprises d'analyser les données des clients et d'optimiser leurs efforts de marketing de diverses manières, telles que la personnalisation des recommandations et l'optimisation du ciblage publicitaire (Lee et Shin, 2020).

L'évolution du marketing était un processus continu, stimulé par les avancées technologiques et les besoins et préférences changeants des entreprises et des consommateurs. De ses modestes débuts dans les civilisations anciennes, le marketing est devenu un domaine sophistiqué et axé sur les données, utilisant les dernières technologies pour toucher et engager les clients de manière de plus en plus efficace et efficiente (Shah et Murthi, 2021).

1.3. Techniques du marketing digital

Les techniques du marketing digital, également appelées techniques de marketing numérique ou e-marketing, désignent l'ensemble des tactiques et outils utilisés pour promouvoir des produits ou des services sur Internet. Ces techniques visent à améliorer la visibilité d'une entreprise en ligne, à attirer du trafic qualifié, à générer des leads et à stimuler les ventes.

1.3.1. Marketing des moteurs de recherche (SEM)

Le marketing des moteurs de recherche (SEM) regroupe l'ensemble des techniques marketing visant à promouvoir un site web en utilisant les moteurs de recherche, principalement Google. Cette stratégie marketing permet de maximiser la visibilité d'un site web sur les résultats de recherche en combinant deux approches complémentaires : l'optimisation pour les moteurs de recherche (SEO) et la publicité sur les moteurs de recherche (SEA).

Le SEO consiste en l'optimisation du contenu et de la structure du site web pour améliorer sa position sur les pages de résultats des moteurs de recherche. L'objectif est de rendre le site web plus facilement identifiable et compréhensible par les robots d'indexation des moteurs de recherche. Le SEO comprend des techniques telles que l'optimisation des mots clés, des balises meta, de la structure des URL et du contenu en général.

Le SEA, quant à lui, est une technique de publicité en ligne qui permet de placer des annonces en haut des pages de résultats des moteurs de recherche en fonction de mots clés spécifiques. Cette technique est payante et permet de bénéficier d'une visibilité immédiate. Les annonces peuvent être sous forme de texte, d'image ou de vidéo.

En combinant les deux techniques, les entreprises peuvent maximiser leur visibilité sur les moteurs de recherche et attirer des visiteurs qualifiés sur leur site web. Il est important de mettre en place une stratégie de SEM adaptée aux objectifs de l'entreprise pour obtenir des résultats efficaces et durables.

1.3.2. Marketing des médias sociaux (SMM)

Le marketing des médias sociaux (SMM) est une stratégie de marketing numérique qui vise à promouvoir une marque, un produit ou un service sur les réseaux sociaux tels que Facebook, Instagram, Twitter, LinkedIn, YouTube, TikTok, etc. Cette stratégie permet aux entreprises de toucher une large audience de manière personnalisée, d'interagir avec leur communauté et d'accroître leur notoriété. Le SMM comprend plusieurs techniques, à savoir :

- La création de contenu : produire des contenus de qualité et pertinents pour attirer et engager l'audience.
- La gestion de communauté : interagir avec les abonnés et répondre à leurs commentaires et questions.
- La publicité sur les réseaux sociaux : diffuser des annonces sponsorisées pour atteindre une audience plus large et ciblée.
- Le marketing d'influence : travailler avec des influenceurs pour promouvoir une marque ou un produit.
- Les concours et les jeux : organiser des concours ou des jeux pour augmenter l'engagement de l'audience et accroître la notoriété de la marque.

Le SMM permet aux entreprises de rester en contact avec leur public et de créer une communauté fidèle qui peut contribuer à augmenter leur visibilité et leur rentabilité.

1.3.3. Optimisation des médias sociaux (SMO)

L'optimisation des médias sociaux (SMO) est une stratégie visant à améliorer la visibilité d'un site web sur les réseaux sociaux. Contrairement au SMM qui se concentre sur la promotion de la marque sur les réseaux sociaux, le SMO se concentre sur l'optimisation du site web pour une meilleure visibilité sur les réseaux sociaux. Le SMO comprend plusieurs techniques, notamment :

- L'optimisation des profils : optimiser les profils sur les réseaux sociaux pour améliorer la visibilité et l'engagement.
- Le partage social : permettre aux utilisateurs de partager facilement le contenu du site web sur les réseaux sociaux.
- Les boutons de partage social : ajouter des boutons de partage sur le site web pour encourager le partage de contenu.
- L'intégration des réseaux sociaux : intégrer les icônes des réseaux sociaux sur le site web pour permettre aux visiteurs de suivre la marque sur les réseaux sociaux.
- L'optimisation du contenu pour les réseaux sociaux : adapter le contenu du site web pour qu'il soit facilement partageable sur les réseaux sociaux.

Le SMO est essentiel pour renforcer la présence en ligne d'une entreprise et améliorer sa crédibilité auprès de son public.

1.3.4. Marketing automatisé des médias sociaux (SMA)

Le marketing automatisé des médias sociaux (SMA) est une stratégie de marketing numérique qui utilise des outils et des logiciels pour automatiser les tâches de gestion des réseaux sociaux. Cette stratégie permet aux entreprises de gagner du temps et d'optimiser leur présence sur les réseaux sociaux. Le SMA comprend plusieurs techniques, telles que :

- La planification de contenu : planifier le contenu à publier sur les réseaux sociaux à l'avance.
- L'automatisation des publications : publier automatiquement le contenu prévu aux heures définies.

- L'automatisation des messages : automatiser les réponses aux messages et commentaires des utilisateurs.
- Le suivi des performances : suivre automatiquement les performances des publications sur les réseaux sociaux.
- La gestion de la publicité : optimiser automatiquement les campagnes publicitaires sur les réseaux sociaux.

Grâce au SMA, les entreprises peuvent maximiser leur efficacité sur les réseaux sociaux et améliorer leur retour sur investissement. Le SMA permet également aux entreprises de mieux comprendre leur public cible et d'adapter leur stratégie de marketing en conséquence.

1.3.5. Marketing de contenu

Le marketing de contenu est une stratégie de marketing numérique qui consiste à créer et à diffuser du contenu informatif, utile et attractif pour attirer et fidéliser un public cible. Cette stratégie permet aux entreprises de développer leur présence en ligne, d'augmenter leur visibilité et de renforcer leur image de marque.

Le marketing de contenu comprend plusieurs sous-techniques, notamment :

- La création de blogs : créer un blog sur le site web de l'entreprise pour publier régulièrement des articles sur des sujets pertinents pour son public cible.
- La création de vidéos : créer des vidéos informatives et attractives pour promouvoir les produits ou services de l'entreprise.
- La création d'infographies : créer des infographies pour présenter des données de manière visuelle et attractive.
- La création d'e-books : créer des livres électroniques informatifs sur des sujets pertinents pour le public cible.
- La création de newsletters : créer des newsletters régulières pour partager les dernières nouvelles de l'entreprise, des offres spéciales, etc.
- La participation aux forums et réseaux sociaux : participer aux discussions en ligne sur des forums et des réseaux sociaux pour partager des informations pertinentes et établir des relations avec le public cible.

En utilisant ces stratégies, les entreprises peuvent attirer un public cible qualifié, renforcer leur expertise et leur autorité dans leur domaine d'activité, et augmenter leur taux de conversion.

1.3.6. Email Marketing

L'email marketing est une stratégie de marketing numérique qui consiste à utiliser l'email pour promouvoir des produits ou services, communiquer avec les clients et prospects, et renforcer l'image de marque de l'entreprise. Cette stratégie permet aux entreprises de toucher un public cible directement dans leur boîte de réception et de personnaliser leurs communications en fonction des intérêts et des préférences du public cible.

L'email marketing comprend plusieurs sous-techniques, notamment :

- La création de listes de diffusion : collecter les adresses email des clients et prospects intéressés par les produits ou services de l'entreprise pour créer une liste de diffusion.
- La segmentation de la liste de diffusion : segmenter la liste de diffusion en fonction des intérêts, des préférences et des comportements d'achat des abonnés pour personnaliser les communications.
- La création de newsletters : créer des newsletters régulières pour partager les dernières nouvelles de l'entreprise, des offres spéciales, des promotions, etc.
- L'envoi de courriels de promotion : envoyer des courriels promotionnels pour promouvoir des produits ou services spécifiques auprès des abonnés intéressés.
- L'envoi de courriels de suivi : envoyer des courriels de suivi personnalisés pour remercier les clients pour leur achat, les inviter à revenir, etc.

L'email marketing permet aux entreprises de toucher un public cible directement dans leur boîte de réception, personnaliser leurs communications et augmenter leur taux de conversion.

1.3.7. Marketing d'influence

Le marketing d'influence est une stratégie de marketing numérique qui consiste à collaborer avec des personnalités influentes sur les réseaux sociaux pour promouvoir des produits ou services auprès de leur audience. Cette stratégie permet aux entreprises de toucher un public cible de manière authentique et crédible, en utilisant l'influence de la personnalité pour renforcer leur image de marque et augmenter leur visibilité. Il existe plusieurs techniques marketing d'influence telles que :

- L'identification des influenceurs : identifier les personnalités influentes sur les réseaux sociaux qui sont pertinentes pour le public cible de l'entreprise.
- La collaboration avec les influenceurs : collaborer avec les influenceurs pour promouvoir les produits ou services de l'entreprise auprès de leur audience via des publications sponsorisées, des revues de produits, des événements en direct, etc.
- La mesure des résultats : mesurer les résultats de la campagne de marketing d'influence pour évaluer son impact et ajuster la stratégie si nécessaire.

Le marketing d'influence permet d'atteindre un public cible de manière authentique et crédible en utilisant l'influence des personnalités sur les réseaux sociaux, renforcer l'image de marque et augmenter leur visibilité.

1.3.8. Publicité native

La publicité native est une stratégie de marketing numérique qui consiste à intégrer des publicités dans le contenu d'un site web ou d'une application mobile, de manière à ce qu'elles ressemblent à des éléments naturels du contenu. Cette stratégie permet aux entreprises de toucher leur public cible de manière non intrusive, en offrant une expérience publicitaire plus agréable et pertinente pour l'utilisateur. La publicité native comprend :

- La création de publicités natives : créer des publicités qui s'intègrent naturellement dans le contenu du site web ou de l'application mobile, en utilisant des éléments tels que des titres attractifs, des images percutantes et des descriptions précises.
- Le choix des emplacements de publicités : choisir les emplacements où les publicités natives seront affichées pour maximiser leur visibilité et leur pertinence pour le public cible.
- La mesure des résultats : mesurer les résultats de la campagne de publicité native pour évaluer son impact et ajuster la stratégie si nécessaire.

En adoptant ces techniques, les entreprises peuvent offrir une expérience publicitaire plus agréable et pertinente pour l'utilisateur, tout en touchant leur public cible de manière non intrusive.

1.3.9. Remarketing

Le remarketing est une stratégie de marketing numérique qui consiste à cibler des personnes qui ont déjà visité le site web ou l'application mobile de l'entreprise, mais qui n'ont

pas effectué d'achat ou d'action spécifique. Cette stratégie permet aux entreprises de retenter de convaincre ces visiteurs de faire un achat ou de s'engager avec l'entreprise. Le remarketing comprend les techniques suivantes :

- Le suivi des visiteurs : installer des outils de suivi pour identifier les visiteurs du site web ou de l'application mobile qui ne sont pas encore convertis en clients.
- La création de publicités personnalisées : créer des publicités personnalisées pour cibler ces visiteurs, en utilisant des éléments tels que des offres spéciales, des recommandations de produits ou des messages de rappel.
- Le ciblage de publicités sur d'autres plateformes : diffuser des publicités personnalisées sur d'autres plateformes telles que les réseaux sociaux ou les sites web tiers pour toucher les visiteurs qui ont quitté le site web ou l'application mobile.
- L'analyse des résultats : mesurer les résultats de la campagne de remarketing pour évaluer son efficacité et ajuster la stratégie si nécessaire.

Le remarketing permet de convaincre les visiteurs qui ont quitté leur site web ou leur application mobile de faire un achat ou de s'engager avec l'entreprise, en utilisant des publicités personnalisées pour répondre à leurs besoins et à leurs intérêts spécifiques.

1.3.10. Marketing viral

Le marketing viral est une stratégie de marketing en ligne qui vise à promouvoir un produit ou un service en encourageant les individus à partager une information ou un contenu à leur entourage. Les vidéos virales sont l'un des moyens les plus populaires pour atteindre cet objectif. En créant des vidéos divertissantes ou informatives qui sont susceptibles de plaire aux utilisateurs, les entreprises peuvent les inciter à les partager avec leurs amis et leur famille, augmentant ainsi leur portée. Les concours et les cadeaux sont également des techniques courantes de marketing viral. Les entreprises peuvent offrir des prix en échange de contenu généré par les utilisateurs, comme des photos ou des vidéos partagées sur les réseaux sociaux. Cela peut aider à inciter les utilisateurs à partager le contenu de la marque et à attirer l'attention sur ses produits ou services. Le marketing viral peut aider une entreprise à étendre sa portée et à attirer de nouveaux clients en utilisant la puissance de l'engagement et de la recommandation sociale.

1.3.11. Marketing d'affiliation

Le marketing d'affiliation est une technique de marketing en ligne dans laquelle une entreprise récompense un affilié pour chaque vente générée par l'affilié. Les sites de coupons et de deals sont des plateformes populaires pour cette technique, car elles permettent aux entreprises de partager des offres exclusives avec des partenaires affiliés. Les affiliés peuvent ensuite promouvoir ces offres auprès de leur public et gagner une commission sur chaque vente effectuée grâce à leur lien de référence. Les sites de contenu et de critiques sont également des partenaires courants dans le marketing d'affiliation. Les entreprises peuvent collaborer avec des blogueurs et des critiques pour promouvoir leurs produits auprès de leur public. Les affiliés peuvent alors utiliser leurs liens de référence pour diriger les lecteurs vers les pages de produits, où ils peuvent effectuer des achats. Les sites de fidélité sont une autre technique de marketing d'affiliation populaire. Les entreprises peuvent inciter leurs clients à promouvoir leurs produits auprès de leur public en offrant des récompenses ou des incitations pour chaque vente générée par leur lien de référence. Dans l'ensemble, le marketing d'affiliation est une technique efficace pour augmenter les ventes et la portée en utilisant des partenaires affiliés pour promouvoir des produits ou des offres exclusives.

1.3.12. Marketing mobile

Le marketing mobile est une technique e-marketing qui utilise les appareils mobiles pour atteindre les clients. Le SMS marketing est une stratégie courante dans laquelle les entreprises envoient des messages promotionnels aux clients via SMS. Les messages peuvent inclure des offres spéciales, des codes de réduction ou des rappels pour des événements à venir. Les in-app ads sont une autre technique de marketing mobile dans laquelle les entreprises placent des publicités au sein des applications mobiles. Les publicités peuvent prendre la forme de bannières, de vidéos ou d'annonces intégrées au contenu de l'application. Les push notifications sont des notifications envoyées directement au dispositif mobile du client. Ces notifications peuvent inclure des messages de marketing ou des rappels pour les événements à venir. Les entreprises peuvent utiliser des push notifications pour rappeler aux clients des offres spéciales ou pour les informer de mises à jour importantes. Le marketing mobile est une technique efficace pour atteindre les clients en déplacement et pour les inciter à acheter ou à interagir avec une entreprise.

1.3.13. Publicité display

La publicité display est une technique e-marketing qui consiste à diffuser des annonces publicitaires graphiques sur des sites Web et des plateformes de médias sociaux. Les bannières publicitaires sont l'un des types d'annonces les plus courants en publicité display. Ces publicités visuellement attractives sont placées sur des sites web ou des plateformes de médias sociaux pour inciter les clients à cliquer dessus. Les publicités Rich Media sont un autre type de publicité display qui inclut des éléments interactifs tels que l'audio, la vidéo ou l'animation. Ces publicités sont conçues pour attirer l'attention des clients et encourager l'interaction avec la marque. Les pop-up ads sont un autre type de publicité display qui s'ouvre dans une nouvelle fenêtre ou un nouvel onglet lorsqu'un client visite un site Web. Ces publicités peuvent être intrusives, mais elles peuvent être efficaces pour attirer l'attention des clients et les inciter à interagir avec la marque. Les entreprises utilisent la publicité display pour accroître la visibilité de leur marque, attirer de nouveaux clients et générer des ventes.

1.3.14. Marketing vidéo

Le marketing vidéo est une stratégie de marketing numérique qui utilise des vidéos pour promouvoir un produit ou un service. Les publicités vidéo peuvent être diffusées sur des plateformes en ligne telles que YouTube, Vimeo, Facebook, Instagram, etc.

Les publicités pré-roll sont des vidéos promotionnelles qui apparaissent avant le contenu vidéo en ligne. Les publicités mid-roll sont des vidéos promotionnelles qui apparaissent pendant le contenu vidéo en ligne, tandis que les publicités post-roll sont des vidéos promotionnelles qui apparaissent après le contenu vidéo en ligne.

Ces publicités peuvent être utilisées pour sensibiliser les consommateurs, augmenter les ventes, renforcer la notoriété de la marque et fidéliser les clients. Il est important de créer des publicités vidéo de qualité, attrayantes et pertinentes pour le public cible afin d'atteindre les objectifs de marketing.

1.4. Stratégies e-marketing : définition, caractéristiques et composants, et étapes de développement

1.4.1. Définition d'une stratégie e-marketing

Une stratégie e-marketing est définie comme étant un plan global pour tous les objectifs commerciaux, les objectifs de marketing et les opportunités, car toutes les entreprises et les

propriétaires de sites auront accès à différents bénéficiaires de différentes manières (Jass, 2022), et elle est définie comme un plan à long terme avec lequel on travaille pour atteindre certains objectifs ou allouer certains bénéfices (Denga, 2023). Sa mise en œuvre nécessite des moyens appropriés pour réaliser les actions et les processus immédiats et nécessaires (Mulyana et Azka, 2022). Une stratégie d'e-marketing représente les étapes adoptées dans la commercialisation des produits basés sur l'Internet et ses différentes applications afin d'attirer les clients et de réaliser un volume important de ventes pour atteindre les objectifs de l'organisation (Gunawan et Sulaeman, 2020).

1.4.2. Caractéristiques et composants d'une stratégie e-marketing

La stratégie e-marketing présente un ensemble de caractéristiques qui peuvent être évoquées comme suit. Elle nécessite un balayage ou une analyse de l'environnement interne et externe. Aussi, elle nécessite une déclaration claire de la vision et des objectifs souhaités. Le développement de la stratégie passe par une nouvelle formulation et une bonne sélection des alternatives stratégiques disponibles (Fuertes et al., 2020). Par ailleurs, la stratégie marketing doit être réactive aux changements de structure et de marché (Li et al., 2021). Cela signifie que le succès de la stratégie d'e-marketing requiert flexibilité et évolutivité en fonction du développement de l'environnement numérique. En effet, il est nécessaire d'ajuster les stratégies e-marketing (Elsayed Fayed, 2021). Si la stratégie d'e-marketing est une stratégie qui combine la stratégie technologique de l'organisation avec la stratégie marketing de l'organisation, son processus de conception nécessite plusieurs considérations, telles que :

- La disponibilité des informations sur les caractéristiques formelles et objectives du produit (Kaur et al., 2022) ;
- La disponibilité de l'information sur les marchés cibles des nouveaux produits (Lasi, 2021).
- L'utilisation des expériences actuelles et passées dans le développement créatif de nouveaux services ;
- L'identification des activités concurrentielles attendues entre les concurrents et leur impact sur l'organisation (Namugenyi et al., 2019) ;
- Enfin, l'analyse de toutes les informations sur les stratégies qui ont été mises en œuvre avec succès pendant une période donnée.

1.4.3. Étapes de développement des stratégies e-marketing

La mise en place d'une stratégie e-marketing peut être résumée dans les étapes suivantes:

Phase de préparation : au cours de celle-ci, les informations nécessaires sur les clients potentiels - besoins, souhaits - marchés cibles et la nature de la concurrence sont recueillies par divers moyens traditionnels ou électroniques (Paralič et al., 2020). À ce stade, le plus important pour le vendeur ou le produit est d'étudier le marché afin de déterminer les besoins et les désirs des consommateurs pour présenter l'article ou le service à la demande sur le marché intérieur ou extérieur. Cela permet de fournir une base de données sur les souhaits des consommateurs locaux, potentiels ou réels, la taille du marché et la concurrence et, par conséquent, ces éléments aident à atteindre les objectifs de l'organisation de manière plus efficace et efficiente que ses concurrents.

Phase de contact : c'est ici que l'on contacte le client pour lui présenter les nouveaux produits mis en ligne sur le marché (Jaas, 2022). L'organisation utilisera divers moyens de communication pour échanger avec le client tout en faisant la promotion de ses produits et en essayant de le convaincre et de l'inciter à acheter ses produits, et la page web joue un rôle important dans ce domaine par le biais de ce que l'on appelle le : Site web ou page web. Elle permet également au consommateur de traiter l'information véhiculée par les moyens électroniques dans le processus de publicité et de promotion à travers le modèle AIDA dans la publicité et la promotion des produits d'entreprise, qui se compose des phases suivantes :

- Phase d'attraction de l'attention : l'objectif est d'attirer l'attention du client à l'aide de bannières publicitaires, d'outils et de vidéos, et d'un message électronique (Shabazz, 2005).
- Phase de collecte des données : à cette étape, on recueille les informations dont le client a besoin (Zeng et Pan, 2010), ce qui l'aide à se forger une opinion particulière sur le nouveau produit.
- La phase d'éveil du désir : elle se concentre sur l'éveil du désir du client à partir de la présentation d'une démonstration impressionnante et dans ce domaine, la technologie multimédia peut être utilisée (Najm et al., 2021).
- La phase de performance et d'action réelles : cette phase concerne en fin de compte la décision d'achat après avoir été convaincu du produit qui a été présenté (Nadube et Ordah, 2023).

- La phase d'échange : à ce stade, le processus d'acceptation et d'échange entre le vendeur et le consommateur est effectué (Houman Andersen, 2001). Cette étape est considérée comme l'essence du processus de marketing électronique, au cours de laquelle le client reçoit l'article ou le service, tandis que l'organisation reçoit le retour en espèces en utilisant des systèmes de paiement sécurisés qui assurent la confidentialité et la protection de la sécurité et de la crédibilité.
- Phase de post-vente : le marketing électronique ne se termine pas avec la fin de l'échange, mais il faut conserver le client en communiquant avec lui et en le servant par divers moyens, notamment les communautés virtuelles, les salons de discussion, le courrier électronique, les listes de questions fréquemment posées, les services de soutien technique et la modernisation (Alsayah, 2020).

1.5. Types des stratégies e-marketing

L'e-marketing repose sur un certain nombre de stratégies différentes. Parmi les caractéristiques cruciales de ces stratégies, on peut citer : la présence et le choix global, l'amélioration de la position concurrentielle, la qualité du service, l'adaptation du produit à la taille du client, la réduction des procédures d'approvisionnement avec une réponse rapide aux demandes, l'abondance des coûts, les opportunités commerciales innovantes et les produits et services renouvelables. Les stratégies e-marketing peuvent être distinguées comme suit :

1.5.1. Stratégies de marché cible

Pour répondre aux besoins des consommateurs ayant des besoins similaires ou parallèles, les stratégies de marché cible comprennent de nombreuses stratégies adoptées par le e-marketing (Al-Zyoud et al., 2021). Cet objectif est atteint en segmentant le marché et en ciblant des catégories dont les composants sont homogènes, en offrant différentes présentations par le biais des sites Web de l'organisation et en adoptant un ensemble de pratiques commerciales adaptées à chaque catégorie (Suh et Chow, 2021).

1.5.2. Stratégie de segmentation du marché

Elle implique que l'accès à tout consommateur se fait non seulement au niveau local mais aussi au niveau international à condition que le consommateur possède un ordinateur, ce qui procure une grande facilité dans les activités de promotion, de publicité mais aussi de distribution (Deepak et Jeyakumar, 2019). L'acheteur a ainsi accès aux informations nécessaires et peut

comparer les produits, les vendeurs, le paiement, etc. Cette stratégie permet également la possibilité d'un marketing personnalisé, c'est-à-dire d'atteindre chaque consommateur et d'essayer de le servir en fonction de ses attentes et de ses besoins grâce à l'e-marketing (Chanitphattana, 2022). Cette stratégie s'adresse principalement au consommateur final en utilisant les méthodes et technologies modernes de communication.

1.5.3. Stratégie de marketing ciblé

Il s'agit d'une stratégie dite de ciblage qui repose sur l'évaluation par la direction du marketing de l'attractivité de chaque secteur du marché qu'elle souhaite aborder (Soedarsono et al., 2020). Les résultats de l'évaluation détermineront, parmi les secteurs sélectionnés, celui qui sera ciblé et qui pourra produire la plus grande valeur rentable pour le client, et maintenir cette valeur dans le temps en établissant une relation client rentable et durable. Cette stratégie évalue l'attractivité de chaque secteur du marché, sélectionne un ou plusieurs secteurs à pénétrer et formule une offre de marketing mix ciblant ces secteurs (Ing et al., 2020).

1.5.4. Stratégie d'implantation sur le marché

Elle signifie que l'organisation a fait des offres qui permettent aux produits de l'organisation d'occuper une position claire, distinctive et désirable dans l'esprit des clients par rapport aux produits des organisations concurrentes (Gul et al., 2021). Elle s'effectue par la conception d'un site efficace et attrayant qui met l'accent sur la valeur du produit, en touchant l'esprit du public (Hassenzahl, 2018).

1.5.5. Stratégie d'excellence

Selon cette stratégie, l'organisation distingue ses produits des autres produits concurrents (Chakravarthy, 1986). L'excellence est définie comme les points importants des différences dans la présentation du produit ou du service (Cravens et al., 1992). Elle peut être également définit comme la création de quelque chose qui est reconnu dans chaque industrie comme quelque chose de distinct et d'unique (Porter, 2008). Atteindre l'excellence ou la différenciation nécessite un certain nombre d'activités coûteuses telles que la conception, la recherche, le développement, l'emballage, le conditionnement et la marque (Rathore, 2022). Toutes ces activités visent à distinguer le produit des produits concurrents afin de laisser une image positive au client et la différenciation dans le e-marketing.

1.5.6. Stratégie d'ajout d'une toute nouvelle industrie

L'organisation et son département e-marketing introduisent des produits entièrement nouveaux dans l'organisation ou sur des marchés auxquels ni l'organisation ni les marchés n'ont jamais eu affaire (Jaas, 2022).

1.5.7. Stratégie d'ajout de lignes de production à des produits

La base de cette stratégie est que l'organisation offre des produits qui ne sont pas nouveaux sur le marché mais nouveaux pour l'organisation, car elle essaie de les ajouter à ses lignes de production afin d'exploiter et d'investir les opportunités de marketing dans l'environnement Internet et les exigences et besoins du marché.

1.5.8. Stratégies d'expansion des lignes de produits existantes

Cette stratégie est mise en œuvre en élargissant la production par des ajustements, en ajoutant des propriétés, des éléments et des avantages aux produits existants, ou en apportant des ajustements à la façon dont les différentes organisations offrent leurs produits aux clients.

1.5.9. Stratégie de modification des caractéristiques implicites des produits

Cette stratégie redessine et développe les produits existants en termes de caractéristiques, d'éléments et d'avantages pour le client, tout en lui faisant prendre conscience, par sa visite sur le site web de l'organisation, que ce qui est proposé est un nouveau produit et que les améliorations apportées au produit sont clairement reconnues par le client.

1.5.10. Stratégie de repositionnement du produit sur le marché

Conformément à cette stratégie, l'organisation redéfinit la position de ses produits dans de nouveaux segments de marché en raison de la découverte de nouvelles utilisations de ses produits. Cela se fait par le biais d'une interaction avec les visiteurs sur le site web ou par le biais d'e-mails, de SMS et d'autres méthodes d'e-marketing.

1.5.11. Stratégie de tarification

Le choix de la stratégie de tarification appropriée dans les relations avec les clients et les visiteurs du site représente une étape critique importante car elle est liée à de nombreuses considérations internes et externes (Yan, 2010), divisées selon les types suivants :

1.5.11.1. Stratégie de tarification dynamique (Dynamic Pricing)

Elle est utilisée pour contrer les fluctuations des différents marchés, d'autant plus que le marketing électronique n'est pas spécifique à un marché et inclut différents marchés mondiaux aux caractéristiques économiques, culturelles, juridiques et autres dissemblables (Bin Abdul Lasi et Salim, 2020). Cela implique une méthode de traitement qui s'adapte aux caractéristiques de chaque marché. Cette stratégie peut permettre d'ajuster les prix en temps réel en fonction de la volonté et de la disposition du client à payer selon ses capacités financières et les conditions du moment (Škare et Gospić, 2015).

1.5.11.2. Stratégie de différenciation des prix

La stratégie de différenciation des prix est une approche qui consiste à fixer des prix pour un produit ou un service en fonction de la nature du marché, des niveaux de pouvoir d'achat des clients et de leurs besoins et désirs spécifiques. Cette stratégie vise à attirer un large éventail de visiteurs et à les convertir en clients en proposant des prix adaptés à leurs caractéristiques individuelles. Voici quelques exemples et éléments clés de cette stratégie :

- Tarification en fonction du pouvoir d'achat : l'entreprise adapte ses prix en fonction de la capacité financière de différents segments de clients. Par exemple, elle peut proposer des tarifs réduits pour les étudiants, les personnes âgées ou les personnes à faible revenu tout en maintenant des prix standard pour les autres segments.
- Personnalisation des offres : cette stratégie peut également inclure la personnalisation des offres en fonction des préférences individuelles des clients. Par exemple, une plateforme de streaming musical peut proposer des abonnements premium avec des fonctionnalités avancées pour ceux qui sont prêts à payer plus, tout en offrant une option de base pour les utilisateurs qui souhaitent un service plus simple à moindre coût.
- Segmentation du marché : l'entreprise identifie différents segments de marché ayant des caractéristiques et des besoins distincts, puis ajuste ses prix pour chaque segment. Par exemple, une compagnie aérienne peut proposer des tarifs différents pour les voyageurs d'affaires, les touristes et les voyageurs fréquents.
- Offres promotionnelles ciblées : la stratégie de différenciation des prix peut également inclure des offres promotionnelles ciblées sur des groupes spécifiques de clients. Par

exemple, une chaîne de restauration rapide peut offrir des repas à prix réduit pour les familles les week-ends pour attirer les clients avec des enfants.

L'objectif de cette stratégie est de maximiser la conversion des visiteurs en clients en leur proposant des prix adaptés à leurs besoins et à leur capacité financière, tout en augmentant la rentabilité globale de l'entreprise grâce à une segmentation stratégique du marché. Cela permet à l'entreprise de mieux répondre aux attentes de divers groupes de clients et d'optimiser ses revenus.

1.5.11.3. Stratégie de prix promotionnels

La stratégie de prix promotionnels est une composante importante du marketing en ligne, visant à stimuler les ventes de produits ou de services spécifiques via le site web de l'entreprise. Elle comprend plusieurs types (Drummond et Ensor, 2006) :

- Stratégie de prix critique : cette stratégie consiste à fixer un prix nettement inférieur à la normale pour un produit ou un service pendant une période limitée. Elle vise à attirer l'attention des clients en offrant une réduction significative, incitant ainsi à l'achat immédiat.
- Stratégie de prix d'événement spécial : cette approche implique la création d'événements spéciaux, tels que des ventes flash, des promotions saisonnières ou des offres spéciales liées à des occasions particulières (comme les soldes d'été ou le Black Friday). Ces événements génèrent un sentiment d'urgence et encouragent les clients à effectuer des achats pendant une période précise.
- Stratégie de prix de concurrence : dans cette stratégie, une entreprise ajuste ses prix pour être compétitive par rapport à d'autres acteurs du marché. Cela peut impliquer de réduire les prix pour rivaliser avec des concurrents directs ou de surveiller en permanence les prix du marché pour s'assurer que les offres restent compétitives.
- Stratégie de prix géographique : cette approche consiste à ajuster les prix en fonction de la localisation géographique des clients. Elle tient compte des différences de coûts, de demande et de concurrence dans différentes régions, ce qui permet à l'entreprise de proposer des tarifs adaptés à chaque marché.

Toutes ces stratégies ont pour objectif commun d'augmenter les volumes de vente en incitant les clients à acheter, que ce soit en offrant des réductions temporaires, en créant un sentiment

d'urgence ou en s'adaptant aux conditions du marché dans différentes régions. Elles sont utilisées pour influencer le comportement d'achat des consommateurs dans le cadre du e-marketing.

En plus des stratégies de prix promotionnels mentionnées précédemment, il existe d'autres types de prix promotionnels couramment utilisés dans le domaine du marketing en ligne. Voici quelques-uns de ces types :

- Prix d'appel : il s'agit d'un prix très bas ou attractif pour un produit ou un service spécifique, conçu pour attirer l'attention des clients et les inciter à visiter le site web de l'entreprise. Une fois sur le site, l'entreprise espère que les clients achèteront également d'autres produits à des prix plus élevés.
- Prix concurrentiel : cette stratégie implique de fixer des prix légèrement inférieurs à ceux de la concurrence pour des produits ou des services similaires. L'objectif est de capter une part du marché en attirant les clients qui recherchent des offres concurrentielles.
- Prix de volume : cette stratégie récompense les achats en grande quantité en offrant des réductions progressives en fonction du volume acheté. Elle encourage les clients à acheter davantage pour économiser sur le prix unitaire.
- Prix groupés : dans cette approche, plusieurs produits ou services sont regroupés et proposés à un prix global réduit par rapport à l'achat individuel de chaque élément. Les clients sont incités à acheter le pack complet pour réaliser des économies.
- Prix saisonniers : les prix sont ajustés en fonction des saisons ou des périodes spécifiques de l'année. Par exemple, des réductions peuvent être offertes sur les équipements d'hiver pendant l'été pour stimuler les ventes hors saison.
- Prix dynamiques : ces prix varient en temps réel en fonction de la demande, des niveaux de stock ou d'autres facteurs. Les clients peuvent voir les prix fluctuer au fur et à mesure qu'ils naviguent sur le site.
- Prix de fidélisation : les clients fidèles ou les membres d'un programme de fidélité peuvent bénéficier de réductions spéciales ou de prix préférentiels pour récompenser leur engagement continu envers la marque.

Ces différentes stratégies de prix promotionnels visent à influencer le comportement d'achat des clients en utilisant des mécanismes de tarification spécifiques pour atteindre des objectifs commerciaux particuliers.

1.5.12. Stratégie de modification des canaux de distribution

L'environnement Internet dans lequel les entreprises opèrent évolue rapidement, et les responsables de la distribution électronique doivent s'adapter en modifiant leurs canaux de distribution électronique (Molenaar, 2013). Cette adaptation peut être motivée par plusieurs raisons, notamment :

- Changement de l'environnement numérique : les évolutions technologiques et les comportements des consommateurs en ligne évoluent rapidement, ce qui peut nécessiter une révision des canaux de distribution pour s'aligner sur les nouvelles tendances et les opportunités émergentes.
- Changement des tendances de l'organisation : les objectifs et les priorités d'une entreprise peuvent changer avec le temps. Par exemple, une entreprise qui se concentrait initialement sur la vente directe aux consommateurs peut décider de développer des partenariats avec des revendeurs en ligne.
- Changement de la stratégie marketing : les objectifs marketing, tels que l'expansion dans de nouveaux marchés ou le ciblage de segments de clients spécifiques, peuvent nécessiter des ajustements dans les canaux de distribution pour atteindre efficacement ces objectifs.
- Changement des structures des canaux de distribution : les canaux de distribution eux-mêmes peuvent évoluer, avec de nouveaux acteurs entrant sur le marché ou des modifications dans les relations entre les fabricants, les distributeurs et les détaillants. Ces changements peuvent nécessiter des ajustements pour optimiser la chaîne de distribution.

Les stratégies de modification des canaux de distribution peuvent prendre différentes formes, telles que :

- Stratégie de distribution intensive : elle implique la distribution du produit ou du service via un maximum de canaux possibles pour atteindre un large public. Cela convient généralement aux produits de grande consommation.
- Stratégie de distribution sélective : elle consiste à choisir soigneusement les canaux de distribution qui correspondent le mieux à la nature du produit ou à la clientèle cible. Cela peut être approprié pour des produits nécessitant une certaine expertise ou une image de marque spécifique.

- Stratégie de distribution exclusive : elle se caractérise par la restriction de la distribution à un nombre limité de canaux, généralement des partenaires sélectionnés. Cette stratégie est souvent utilisée pour des produits haut de gamme ou de niche.

La modification des canaux de distribution est essentielle pour s'adapter à un environnement commercial en constante évolution et pour répondre aux besoins changeants des clients tout en maximisant l'efficacité et la rentabilité de la distribution électronique.

1.5.13. Stratégie de gestion des relations clients

L'utilisation d'Internet et des sites Web a rapproché les organisations de la stratégie de gestion de la relation client, car elle a modifié le comportement d'achat des consommateurs et a influencé leurs opinions et leurs emplacements à l'égard de l'organisation et de ses produits (Saputra et Sunardi, 2022). En outre, elle a contribué à l'activation de cette relation, c'est-à-dire que le développement des technologies et services de communication tels que le service web, le courrier électronique et les SMS a rendu l'organisation plus interactive avec le client (Reza Chowdhury, 2022).

2. Machine Learning : fondements théoriques

L'apprentissage automatique ou machine learning (ML) est une sous-catégorie de l'intelligence artificielle qui consiste à utiliser des algorithmes pour apprendre à partir de données et prendre des décisions ou des prédictions sans programmation explicite (Polce et Kunze, 2023). Les algorithmes de ML peuvent être entraînés sur de grands ensembles de données et peuvent améliorer leur performance au fil du temps en s'adaptant à de nouvelles données (Ray, 2019).

L'histoire de l'invention du machine learning remonte à la fin des années 1940 et au début des années 1950, lorsque les scientifiques ont commencé à explorer les idées de l'apprentissage automatique et de la théorie de l'information.

Le terme "machine learning" a été utilisé pour la première fois par le scientifique américain Arthur Samuel en 1959, pour décrire le processus par lequel les ordinateurs peuvent apprendre à partir de données, sans être explicitement programmés pour cela.

Dans les années 1960 et 1970, des scientifiques comme Marvin Minsky, John McCarthy et Claude Shannon ont commencé à explorer plus en profondeur les idées de l'apprentissage

automatique, en utilisant des réseaux de neurones artificiels et d'autres techniques pour créer des systèmes qui pourraient apprendre à partir de données.

Au cours des années 1980 et 1990, le développement de l'informatique et l'augmentation de la puissance de calcul ont permis aux scientifiques de développer des algorithmes machine learning plus complexes et plus performants.

Au fil des années, les progrès de la technologie de l'informatique ont permis de développer des algorithmes machine learning plus sophistiqués, capables de traiter de grandes quantités de données et d'apprendre à partir de modèles plus complexes. En particulier, l'avènement de l'apprentissage profond dans les années 2010 a considérablement amélioré la capacité des machines à apprendre et à s'adapter à des données de grande complexité, en utilisant des réseaux de neurones artificiels.

Aujourd'hui, le machine learning est utilisé dans une grande variété d'applications, allant de la reconnaissance vocale et de la traduction automatique à la prédiction des résultats des élections et des tendances du marché. Les avancées continues dans la technologie de l'informatique et dans la compréhension des méthodes machine learning promettent de continuer à étendre les capacités des machines à apprendre et à s'adapter à un large éventail de tâches.

Dans le domaine du e-marketing, le ML peut être utilisé pour améliorer les stratégies marketing en aidant les entreprises à mieux comprendre le comportement des clients, à adapter les efforts de marketing et à optimiser l'expérience client (Kapoor et Kapoor, 2021). Certaines applications courantes du ML en e-marketing comprennent la segmentation des clients (Wu et al., 2021), les recommandations personnalisées (Loukili et Messaoudi, 2023), (Loukili et al., 2023), la modélisation prédictive, l'analyse des sentiments (Loukili et al., 2023), la prédiction de churn (Loukili et al., 2022) et la détection de la fraude (Loukili et al., 2023).

En analysant de grandes quantités de données et en utilisant des algorithmes avancés, le ML peut aider les entreprises à identifier des modèles et des tendances dans le comportement des clients, à faire des recommandations personnalisées et à optimiser les efforts de marketing (Rathore, 2023). Cela peut conduire à une augmentation de l'engagement des clients, à une amélioration de la fidélité des clients et à une augmentation des ventes (Patel et Trivedi, 2020).

En général, le machine learning peut être divisé en quatre types différents, à savoir, l'apprentissage supervisé, semi-supervisé, non supervisé, l'apprentissage par renforcement. Cependant, il existe d'autres types du machine learning tels que : l'apprentissage profond, le transfer learning, et le traitement automatique du langage naturel.

2.1. Apprentissage supervisé

Les algorithmes d'apprentissage supervisé (Supervised Learning) sont des algorithmes d'apprentissage automatique qui apprennent à partir de données d'entraînement étiquetées (Mahesh, 2020). Les données d'entraînement sont constituées d'exemples d'entrée (également appelés caractéristiques) et de leurs étiquettes de sortie souhaitées correspondantes (également appelées cibles). L'objectif de l'apprentissage supervisé est de construire un modèle qui peut prédire les étiquettes de sortie pour de nouveaux exemples d'entrée en fonction des modèles appris à partir des données d'entraînement (Miyato et al., 2018). Les algorithmes du machine learning supervisé sont utilisés généralement pour résoudre des problèmes de classification, de régression, ou d'étiquetage de séquence.

2.1.1. Algorithmes de classification

Les algorithmes de classification prédisent une étiquette de classe pour un exemple d'entrée. Par exemple, un classificateur peut prédire si un e-mail est un spam ou non en fonction de son contenu. Les algorithmes de classification comprennent la régression logistique, les arbres de décision, les forêts aléatoires, les machines à vecteurs de support et les réseaux neuronaux.

La régression logistique est un algorithme de classification utilisé pour prédire un résultat binaire (Christodoulou el al., 2019). Il procède en ajustant une courbe aux données, en utilisant une fonction logistique pour prédire la probabilité d'une certaine classe (Shah et al., 2020).

Les arbres de décision est un modèle en forme d'arbre utilisé pour la classification et la régression (Thomas et al., 2020). Cet algorithme agit en créant une série de divisions binaires basées sur les valeurs des caractéristiques, résultant en une série de nœuds de "décision" qui mènent à une prédiction (Sutton, 2005).

Les forêts aléatoires est une méthode d'ensemble qui combine plusieurs arbres de décision pour faire une prédiction (Kocev et al., 2007). Il fonctionne en entraînant plusieurs arbres de

décision sur des sous-ensembles aléatoires des données (Prasad et al., 2006), puis en moyennant leurs prédictions pour faire une prédiction finale.

Les machines à vecteurs de support (SVM) est un algorithme de classification qui trouve l'hyperplan dans un espace de grande dimension qui sépare au mieux les classes (Chamasemani et Singh, 2011). SVM est souvent utilisé pour les problèmes où il y a une séparation claire entre les classes.

Les réseaux de neurones sont des modèles complexes inspirés de la structure et de la fonction du cerveau (Kasabov, 2019). Ils sont composés de plusieurs "neurones" interconnectés qui peuvent apprendre à reconnaître des modèles dans les données (Maind et Wankar, 2014). Les réseaux de neurones sont souvent utilisés pour des tâches telles que la classification d'images et le traitement du langage naturel.

2.1.2. Algorithmes de régression

Les algorithmes de régression prédisent une valeur numérique continue pour un exemple d'entrée (Tan et al., 2021). Par exemple, un modèle de régression peut prédire le prix d'une maison en fonction de sa taille, de son emplacement et d'autres caractéristiques. La régression linéaire, Ridge, Lasso, et Elastic net sont parmi les techniques de régression les plus courants.

- La régression linéaire : est un algorithme simple utilisé pour prédire une valeur continue (comme une probabilité). Il opère en ajustant une ligne droite ou un hyperplan aux données (Obulesu et al., 2018).
- La régression Ridge : est une variation de la régression linéaire qui ajoute un terme de régularisation à la fonction objective afin d'éviter le surapprentissage (overfitting) et d'améliorer la capacité de généralisation du modèle (Wyffels et al., 2008).
- Régression Lasso : est une autre variation de la régression linéaire qui ajoute un terme de régularisation basé sur la valeur absolue des coefficients, plutôt que le carré des coefficients comme dans la régression Ridge (McNeish, 2015).
- Elastic net : est une combinaison de la régression Ridge et Lasso, qui combine les termes de régularisation des deux méthodes (Gao et al., 2019).

2.1.3. Algorithmes d'étiquetage de séquence

Les algorithmes d'étiquetage de séquence (Sequence Labeling algorithms) prédisent une étiquette de classe pour chaque élément dans une séquence d'exemples d'entrée (Kumar et al.,

2018). Par exemple, un algorithme d'étiquetage de séquence pourrait prédire l'étiquette de partie du discours pour chaque mot dans une phrase.

- Modèles de Markov cachés (Hidden Markov Models, HMM) : sont des modèles probabilistes utilisés pour prédire une séquence d'étiquettes pour une séquence de données d'entrée (Tamposis et al., 2019). Les HMM supposent que les étiquettes sont générées à partir d'un ensemble d'états cachés, et que la probabilité de transition entre les états et d'émission d'une étiquette est donnée par un ensemble de paramètres appris (Qiao et al., 2015).
- Champs aléatoires conditionnels (Conditional Random Fields, CRF) : sont des modèles probabilistes similaires aux HMM, mais ils supposent que les étiquettes sont conditionnellement indépendantes donné les données d'entrée, plutôt que d'être générées à partir d'un ensemble d'états cachés (Klinger et Tomanek, 2007).
- Réseaux neuronaux récurrents (Recurrent Neural Networks, RNN) : sont des réseaux neuronaux qui ont une mémoire et peuvent traiter des données séquentielles (Tarwani et Edem, 2017). L es RNN sont souvent utilisés pour des tâches telles que la modélisation de langage et la traduction automatique.

2.2. Apprentissage non-supervisé

L'apprentissage non supervisé est un type machine learning où l'algorithme ne reçoit aucune donnée d'entraînement étiquetée ni aucune connaissance préalable sur la structure des données (Hosseinpour et al., 2014). L'objectif de l'apprentissage non supervisé est de découvrir des modèles ou des relations dans les données. Il existe plusieurs types d'algorithmes machine learning non supervisé, y compris le clustering, la réduction de dimensionnalité, La détection d'anomalies, les modèles génératifs, et les cartes auto-organisatrices.

2.2.1. Clustering

Il s'agit d'une technique de division des données en groupes, appelés clusters, en fonction de leur similarité ou de leur distance les uns des autres (Shetty et Singh, 2021). Des exemples d'algorithmes de clustering incluent : k-Means, DBSCAN, Expectation-Maximization, Affinity Propagation et le clustering hiérarchique.

K-means est un algorithme de clustering populaire qui fonctionne en attribuant de manière itérative chaque point de données au cluster ayant le centroïde (moyenne) le plus proche, puis

en mettant à jour les centroïdes en fonction de la moyenne des points de chaque cluster (Hu et al., 2020).

Clustering hiérarchique est une technique de clustering qui fonctionne en créant une hiérarchie de clusters, où chaque cluster est fusionné avec le cluster le plus similaire à chaque étape (Naeem et al., 2019). Le clustering hiérarchique peut être divisé en deux types principaux : agglomératif (bas vers le haut) et divisif (haut vers le bas) (Shetty et Singh, 2021).

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) est un algorithme de clustering basé sur la densité qui fonctionne en identifiant les points "centraux" entourés d'une forte densité de points, puis en étendant les clusters à partir de ces points centraux (Deng et al., 2019).

Expectation-Maximization (EM) est un algorithme destiné à l'apprentissage non supervisé, en particulier pour les problèmes où les données sont incomplètes ou comportent des valeurs manquantes (Cenitta et al., 2022). EM repose sur l'estimation itérative des valeurs manquantes et des paramètres du modèle sous-jacent jusqu'à convergence (Gold et Bentler, 2000). Il est souvent utilisé pour des tâches telles que le clustering et l'estimation de densité.

Affinity Propagation est un algorithme de clustering qui fonctionne en envoyant des "messages" entre les points de données, qui indiquent la similarité entre les points (Kazantseva et Szpakowicz, 2011). Les messages sont transmis de manière itérative jusqu'à ce qu'un consensus soit atteint sur les clusters (Taheri et Bouyer, 2020). Affinity Propagation est particulièrement utile pour les données qui ont un grand nombre de clusters avec des tailles et des densités inégales.

2.2.2. Réduction de dimensionnalité

La réduction de dimensionnalité est une technique pour réduire le nombre de dimensions (caractéristiques) dans les données tout en préservant autant d'informations que possible (Ayesha et al., 2020). Des exemples d'algorithmes de réduction de dimensionnalité comprennent l'analyse en composantes principales (ACP) et l'incorporation stochastique de voisins t-distribués (t-SNE).

Analyse en composantes principales est une technique de réduction de dimensionnalité linéaire qui fonctionne en projetant les données sur un sous-espace de dimensions inférieures de sorte que la variance totale des données soit maximisée (Barshan et al., 2011).

t-distributed Stochastic Neighbor Embedding (t-SNE) est un algorithme de réduction de dimensionnalité non linéaire qui fonctionne en minimisant la divergence entre les probabilités de points similaires dans l'espace de haute dimension et l'espace de basse dimension (Belkina et al., 2019).

Décomposition en valeurs singulières (SVD) est une technique d'algèbre linéaire utilisée pour la réduction de dimensionnalité et la compression de données (Jiang, 2022). SVD décompose une matrice en trois matrices : une matrice unitaire, une matrice diagonale et une autre matrice unitaire (Le Bihan et Mars, 2004). La matrice diagonale contient les "valeurs singulières" de la matrice d'origine, qui sont les racines carrées de ses valeurs propres. SVD peut être utilisé pour réduire la dimensionnalité d'une matrice en ne conservant que les plus grandes valeurs singulières et les vecteurs singuliers correspondants (Jaradat et al., 2021).

L'analyse discriminante linéaire (LDA) est un algorithme de réduction de dimensionnalité utilisé pour les tâches de classification (Zhao et al., 2019). LDA fonctionne en projetant les données sur un espace de dimensions inférieures de sorte que la séparation entre les classes soit maximisée (Li et al., 2011). LDA est particulièrement utile lorsque les classes sont bien séparées et que les dimensions sont corrélées (Deepika et al., 2023).

2.2.3. Détection d'anomalies

La détection d'anomalies représente une technique pour identifier les points de données inhabituels ou atypiques qui ne correspondent pas au modèle attendu des données (Trardi el al., 2022). Les algorithmes de détection d'anomalies sont souvent utilisés dans la détection de fraude et la cybersécurité (Jiang et al., 2019). Parmi ces algorithmes il y a les machines à vecteurs de support à une seule classe, Isolation Forest, Facteur d'anomalie local, et les autoencodeurs.

Isolation Forest est un algorithme qui permet de construire une forêt d'arbres de décision, où chaque arbre est formé sur un sous-ensemble de données sélectionné de manière aléatoire (Liu et al., 2008). Les points anormaux sont identifiés comme "isolés" dans la forêt, ce qui signifie qu'ils nécessitent plus de divisions pour être atteints dans l'arbre (Puggini et McLoone, 2018).

Les machines à vecteurs de support à une seule classe (SVM) est une méthode de classification qui est entraîné sur une seule classe de données, puis utilisé pour identifier les points de données qui n'appartiennent pas à cette classe (Amarappa et Sathyanarayana, 2014).

Facteur d'anomalie local (Local Anomaly Factor, LOF) est une approche utilisée pour la détection d'anomalies en apprentissage non supervisé (Xu et al., 2019). LOF fonctionne en mesurant la "densité" d'un point de données par rapport à ses voisins et en identifiant les points qui ont une densité significativement plus faible comme des anomalies (De Vries et al., 2010). LOF est particulièrement utile pour détecter les anomalies dans les données de haute dimension (Thudumu et al., 2020).

La détection d'anomalies basée sur les autoencodeurs est une méthode qui utilise des autoencodeurs, qui sont des réseaux de neurones qui apprennent à coder et décoder les données, pour la tâche de détection d'anomalies (Provotar et al., 2019). Un autoencodeur est entraîné sur les données normales, puis utilisé pour reconstruire les données d'entrée. Les anomalies sont identifiées comme les points de données qui sont mal reconstruits par l'autoencodeur. La détection d'anomalies basée sur les autoencodeurs est souvent utilisée pour des tâches telles que la détection de fraude ou la surveillance de systèmes pour des comportements inhabituels.

2.2.4. Modèles génératifs

Les modèles génératifs sont des modèles qui apprennent à générer de nouveaux points de données similaires aux données d'entraînement (Song et Ermon, 2020). Des exemples de modèles génératifs incluent les autoencodeurs, les autoencodeurs variationnels, et les réseaux antagonistes génératifs.

Les autoencodeurs sont des réseaux de neurones qui apprennent à encoder les données d'entrée en une représentation de dimension inférieure, puis à décoder la représentation en une approximation des données d'origine (Bank et al., 2020). Les autoencodeurs peuvent être utilisés pour des tâches telles que la réduction de la dimensionnalité ou le débruitage d'image.

Les autoencodeurs variationnels (Variational AutoEncoders, VAE) sont un type de modèle génératif utilisé pour des tâches telles que la génération d'images, la compression de données et l'apprentissage de représentations (Fan et al., 2022). Un VAE se compose d'un encodeur et d'un décodeur, qui sont entraînés à reconstruire les données d'entrée en minimisant la perte de

reconstruction (Moon et al., 2023). L'encodeur mappe les données d'entrée dans un espace latent, qui est une représentation de dimension inférieure des données, et le décodeur mappe l'espace latent de retour à l'espace d'origine (Ahmed et Longo, 2022). Les VAE sont entraînés en maximisant la vraisemblance des données sous le modèle, ce qui est fait en utilisant une technique appelée inférence variationnelle (Akrami et al., 2022).

Les réseaux antagonistes génératifs (Generative Adversarial Networks, GAN) sont un type de modèle génératif qui se compose de deux réseaux de neurones : un générateur et un discriminateur (Kar et al., 2023). Le générateur apprend à générer de nouveaux points de données similaires aux données d'entraînement, tandis que le discriminateur essaie de distinguer entre les points de données réels et faux (You et al., 2022). Le générateur et le discriminateur sont entraînés ensemble, de manière antagoniste, pour améliorer la qualité des données générées (Xie et al., 2022). Les GAN ont été utilisés pour des tâches telles que la génération d'images, la génération de texte et le transfert de style.

2.2.5. Self-organizing maps

Les cartes auto-organisatrices ou Self-Organizing Maps (SOM) sont des modèles de réseaux neuronaux qui peuvent être utilisés pour le regroupement et la visualisation de données de haute dimension (Zhong, 2022). Kohonen SOM et Growing SOM sont parmi les algorithmes machine learning utilisés dans ces modèles.

La carte auto-organisatrice de Kohonen (Kohonen SOM) est un ensemble de réseaux neuronaux qui apprennent à mapper les données d'entrée à une grille de dimension inférieure, de sorte que les points de données similaires soient mappés à des points proches dans la grille (Kovačević et al., 2022). Les SOM sont souvent utilisés pour la visualisation et pour des tâches telles que le regroupement et la détection d'anomalies.

La carte auto-organisatrice croissante (Growing Self-Organizing Map, GSOM) est une variante de l'algorithme traditionnel de la carte auto-organisatrice (SOM) qui est utilisée pour l'apprentissage non supervisé (Sarkar et al., 2022). Comme une SOM traditionnelle, une GSOM est un réseau neuronal qui apprend à mapper les données d'entrée à une grille de dimension inférieure, de sorte que les points de données similaires soient mappés à des points proches dans la grille (Mahdavi et al., 2022). Cependant, contrairement à une SOM traditionnelle qui a une taille fixe, une GSOM peut "grandir" en ajoutant de nouveaux nœuds à la grille au besoin. Cela

permet à une GSOM de s'adapter à la structure des données et d'éviter le surapprentissage. Les GSOM sont souvent utilisées pour des tâches telles que la visualisation et pour le regroupement et la détection d'anomalies.

2.3. Apprentissage semi-supervisé

L'apprentissage semi-supervisé est une technique d'apprentissage automatique dans laquelle un modèle est entraîné sur un ensemble de données partiellement étiquetées, avec certains points de données ayant leurs étiquettes et d'autres n'en ayant pas (Tu et Menzies, 2022). C'est une approche utile lorsqu'il y a une quantité limitée de données étiquetées disponibles, car elle permet au modèle d'apprendre à la fois à partir de données étiquetées et non étiquetées (Berthelot et al., 2019). Les méthodes basées sur les graphes, le Self-learning, le Co-training, et le Multi-view learning sont des méthodes qui peuvent être utilisés pour l'apprentissage semi-supervisé.

2.3.1. Méthodes basées sur les graphes

Les méthodes basées sur les graphes utilisent la structure des données pour propager les étiquettes des données étiquetées vers les données non étiquetées (Zhan et al., 2018). Des exemples de ces méthodes incluent la propagation d'étiquettes, la diffusion d'étiquettes, les arbres couvrants minimum, et les champs gaussiens.

La propagation d'étiquettes est une approche où les étiquettes sont propagées des points de données étiquetés aux points de données non étiquetés à l'aide d'un graphe pondéré, les poids entre les points de données étant déterminés par leur similarité dans l'espace des caractéristiques (Zhuʗ et Ghahramaniʗн, 2002). Les étiquettes des points de données non étiquetés sont ensuite mises à jour en fonction des étiquettes de leurs voisins. Ce processus est répété jusqu'à ce que les étiquettes des points de données non étiquetés convergent (Stikic et al., 2009).

La diffusion d'étiquettes (Label propagation) est similaire à la propagation d'étiquettes, mais au lieu d'utiliser un graphe pondéré pour propager les étiquettes, il utilise une fonction de noyau pour définir l'influence de chaque point de données étiqueté sur les points de données non étiquetés (Wang et al., 2013).

Arbre couvrant minimum (Minimum spanning tree, MST) est un autre algorithme semi-supervisé basé sur les graphes qui peut être utilisé pour la propagation d'étiquettes (Wang et Ji, 2020). Dans cette approche, un MST est construit à partir des points de données, et les étiquettes sont propagées le long des arêtes de l'arbre (Cheng et al., 2019).

Champs gaussiens (Gaussian fields) est une méthode semi-supervisée basée sur les graphes qui utilise une distribution gaussienne pour modéliser la relation entre les points de données dans l'espace des caractéristiques (Ng et al., 2018). Les étiquettes sont propagées en fonction de la probabilité que chaque point de données ait une étiquette particulière étant donné ses voisins (Gong et al., 2016).

2.3.2. Self-learning

L'auto-apprentissage ou le Self-learning est une méthode qui consiste à entraîner un modèle sur les données étiquetées, puis à utiliser le modèle entraîné pour prédire les étiquettes des données non étiquetées (Al Ghamdi et al., 2019). Les étiquettes prédites sont ensuite ajoutées à l'ensemble d'entraînement et le modèle est ré-entraîné. Les arbres de décision, les machines à vecteurs de support, les réseaux de neurones, et les forêts aléatoires sont des exemples des algorithmes Self-learning.

Les arbres de décision sont un type d'algorithme d'apprentissage automatique utilisé pour les tâches de classification et de régression (Somvanshi et al., 2016). Ils fonctionnent en construisant un modèle en forme d'arbre de décisions basé sur les caractéristiques des données. Chaque nœud interne de l'arbre représente une décision basée sur une caractéristique des données, et chaque nœud feuille représente une prédiction ou une étiquette de classe. L'arbre est construit en divisant les données sur la caractéristique qui maximise la séparation des classes ou réduit la variance de la variable cible.

Les machines à vecteurs de support (SVM) sont un type d'algorithme d'apprentissage supervisé qui peut être utilisé pour les tâches de classification et de régression. Ils fonctionnent en trouvant l'hyperplan dans un espace de caractéristiques haute dimension qui sépare de manière maximale les différentes classes ou cibles (Pisner et Schnyer, 2020). Les SVM sont particulièrement efficaces lorsque les données sont linéairement séparables, mais ils peuvent également être étendus à des problèmes non linéaires en utilisant des astuces de noyau (Nie et al., 2020).

Les réseaux de neurones sont un type d'algorithme d'apprentissage automatique inspiré de la structure et de la fonction du cerveau humain. Ils sont composés de plusieurs couches de nœuds interconnectés et sont entraînés à l'aide de grandes quantités de données étiquetées et de la rétropropagation (Dongare et al., 2012). Les réseaux de neurones peuvent être utilisés pour une large gamme de tâches, notamment la classification, la régression et le regroupement.

Les forêts aléatoires est une méthode d'apprentissage par ensemble qui combine plusieurs arbres de décision pour faire des prédictions (Altman et Krzywinski, 2017). Chaque arbre de la forêt est entraîné sur un sous-ensemble différent des données, et la prédiction finale est faite en moyennant les prédictions de tous les arbres (Kontschieder et al., 2015). Les forêts aléatoires sont connues pour leur capacité à traiter de grands ensembles de données complexes, et elles sont souvent utilisées pour les tâches de classification et de régression.

2.3.3. Co-training

Le co-training ou co-entraînement est une méthode consiste à entraîner deux modèles sur des vues différentes des données, puis à utiliser les prédictions des deux modèles pour étiqueter les données non étiquetées (Xia et al., 2020). Les données étiquetées sont ensuite ajoutées aux ensembles d'entraînement des deux modèles, et le processus est répété (Blum et Mitchell, 1998). Parmi les algorithmes utilisés dans cette méthode, on peut citer : les arbres de décision, SVM, les réseaux de neurones, et les forêts aléatoires.

2.3.4. Multi-view learning

L'apprentissage multi-vues (Multi-view learning) consiste à entraîner un modèle sur plusieurs vues des données, où chaque vue représente une caractéristique ou une représentation différente des données (Xia et al., 2020). Le modèle est alors capable d'apprendre à partir des données étiquetées et non étiquetées. L'analyse de corrélation canonique, l'apprentissage de multiples noyaux, et l'apprentissage profond sont des algorithmes d'apprentissage multi-vues.

L'analyse de corrélation canonique (Canonical Correlation Analysis, CCA) est une technique statistique qui vise à maximiser la corrélation entre deux vues des données. Elle est utilisée pour trouver une représentation partagée des données qui est utile pour une tâche particulière (Luo et al., 2015). CCA fonctionne en trouvant des combinaisons linéaires des caractéristiques de chaque vue qui sont maximale ment corrélées les unes avec les autres. Ces combinaisons

linéaires peuvent ensuite être utilisées comme caractéristiques pour une analyse ultérieure ou comme entrée pour un modèle d'apprentissage automatique.

L'apprentissage de multiples noyaux (Multiple Kernel Learning, MKL) est une approche d'apprentissage automatique qui consiste à utiliser plusieurs noyaux pour apprendre un modèle (Gönen et Alpaydın, 2013). Un noyau est une fonction qui mesure la similarité entre deux points de données. Dans MKL, le modèle est entraîné en utilisant plusieurs noyaux, chacun capturant un aspect différent des données (Schrouff et al., 2016). Les poids des noyaux sont ensuite optimisés pour améliorer les performances du modèle sur la tâche à accomplir. MKL peut être utilisé avec divers types d'algorithmes d'apprentissage automatique, y compris les machines à vecteurs de support (SVM) et les réseaux de neurones.

2.4. Apprentissage par renforcement

L'apprentissage par renforcement (Reinforcement Learning) est un type d'apprentissage automatique dans lequel un agent apprend en interagissant avec son environnement et en recevant des commentaires sous forme de récompenses ou de punitions (Kaelbling et al., 1996). L'objectif de l'agent est d'apprendre une politique qui maximise la récompense cumulative qu'il reçoit au fil du temps.

En apprentissage par renforcement, l'agent et l'environnement sont modélisés sous forme de processus de décision markovien (Markov Decision Process, MDP). Un MDP est défini par un ensemble d'états, un ensemble d'actions et un ensemble de transitions qui décrivent comment l'agent peut passer d'un état à un autre en prenant une action (Castro, 2020). L'agent reçoit une récompense pour chaque transition qu'il effectue, en fonction de l'état dans lequel il se trouve et de l'action qu'il prend.

La politique de l'agent est une fonction qui détermine l'action à prendre dans chaque état. L'objectif de l'agent est de trouver une politique qui maximise la somme attendue des récompenses qu'il reçoit au fil du temps. Cela se fait grâce à un processus appelé apprentissage par renforcement, dans lequel l'agent essaie différentes actions et reçoit des commentaires sous forme de récompenses ou de punitions, qu'il utilise pour mettre à jour sa politique. Il existe plusieurs types d'apprentissage par renforcement, dans cette optique on peut citer : l'apprentissage par renforcement basé sur un modèle, sans modèle, hors politique, sur politique, hiérarchique, et profond.

2.4.1. Apprentissage par renforcement basé sur un modèle

Dans l'apprentissage par renforcement basé sur un modèle, l'agent apprend un modèle de l'environnement et utilise ce modèle pour prédire les conséquences de ses actions. Il peut alors utiliser ce modèle pour planifier ses actions et optimiser sa politique (Ho et Lee, 2015). Les algorithmes d'apprentissage par renforcement basé sur un modèle incluent : Dyna et Monte Carlo Tree Search.

Dyna est algorithme qui combine la planification basée sur un modèle avec l'apprentissage sans modèle (Nagabandi et al., 2018). Il opère en apprenant un modèle de l'environnement et en utilisant ce modèle pour générer des expériences simulées, qui sont ensuite utilisées pour mettre à jour la fonction de valeur ou la politique (Gu et al., 2016).

Monte Carlo Tree Search est un algorithme de planification qui utilise des simulations Monte Carlo pour évaluer les résultats potentiels de différentes actions (Shuai et He, 2020). Il fonctionne en construisant un arbre d'actions et de résultats possibles et en utilisant des simulations Monte Carlo pour estimer la récompense attendue pour chaque action.

2.4.2. Apprentissage par renforcement sans modèle

Dans l'apprentissage par renforcement sans modèle, l'agent n'apprend pas de modèle de l'environnement. Au lieu de cela, il apprend directement la politique optimale par essais et erreurs (Saunders et al., 2017). Cela peut être plus efficace dans certains cas, car cela évite la nécessité d'apprendre un modèle de l'environnement, mais cela peut aussi être moins robuste, car l'agent a moins de compréhension des dynamiques sous-jacentes de l'environnement (Ibarz et al., 2021). Q-learning et SARSA sont des exemples d'apprentissage par renforcement sans modèle.

Q-learning est un algorithme de reinforcement learning sans modèle qui utilise une table Q pour stocker la récompense attendue pour chaque paire état-action (Zamfirache et al., 2022). Il fonctionne en mettant à jour la table Q en fonction de la récompense observée pour chaque action et en utilisant la table Q pour choisir l'action avec la récompense attendue la plus élevée.

SARSA (State-Action-Reward-State-Action) est un autre algorithme de reinforcement learning sans modèle qui utilise une table Q pour stocker la récompense attendue pour chaque paire état-action (Gautier, 2021). Il fonctionne en mettant à jour la table Q en fonction de la

récompense observée pour l'action actuelle et de la récompense attendue pour l'action suivante, telle que prédite par la table Q (Mamaghani et Hong, 2020).

2.4.3. Apprentissage par renforcement hors politique

Dans l'apprentissage par renforcement hors politique (Off-policy Reinforcement Learning), l'agent apprend une politique différente de celle qu'il suit actuellement. Cela peut être utile dans des situations où il est difficile ou impossible de suivre la politique optimale en tout temps, par exemple lorsque l'agent interagit avec un système du monde réel qu'il ne peut pas contrôler entièrement (Bloembergen et al., 2015). Parmi les algorithmes utilisés dans ce type d'apprentissage les réseaux de neurones profonds Q.

Réseaux de neurones profonds Q (DQN) : DQN est un algorithme d'apprentissage par renforcement profond qui utilise un réseau de neurones profonds pour approximer la fonction Q (Van Hasselt et al., 2016). Il fonctionne en utilisant une relecture d'expérience pour stocker et apprendre à partir d'expériences passées et en utilisant des réseaux cibles pour stabiliser le processus d'apprentissage (De Bruin et al., 2015).

Double DQN (DDQN) est une variante de DQN qui résout le problème des estimations de valeur sur-optimistes. Il fonctionne en utilisant deux réseaux de neurones distincts pour estimer la fonction d'action-valeur : un réseau est utilisé pour sélectionner l'action et l'autre réseau est utilisé pour évaluer l'action (Gong et al., 2020).

2.4.4. Apprentissage par renforcement sur politique

Dans l'apprentissage par renforcement sur politique (Policy-based Reinforcement Learning), l'agent apprend la politique optimale tout en la suivant. Cela peut être plus efficace dans certains cas, car l'agent prend toujours des actions susceptibles de conduire à la plus grande récompense, mais cela peut aussi être moins robuste, car l'agent ne peut pas apprendre à partir d'expériences qui divergent de la politique actuelle (Jia et Wang, 2020). Les méthodes de gradient de politique et l'optimisation de politique proximale sont des algorithmes d'apprentissage par renforcement sur politique.

Les méthodes de gradient de politique sont une classe d'algorithmes d'apprentissage par renforcement qui optimisent directement la politique en utilisant la descente de gradient (Desjardins et Chaib-Draa, 2011). Ils fonctionnent en estimant le gradient de la récompense

attendue par rapport aux paramètres de la politique et en utilisant ce gradient pour mettre à jour la politique (Zhou et al., 2020).

L'optimisation de politique proximale (Proximal Policy Optimization, PPO) est un algorithme de gradient de politique qui utilise une approche de région de confiance pour stabiliser le processus d'apprentissage (Cheng et al., 2021). PPO opère en contraignant le changement de la politique à une plage fixe, ce qui aide à éviter que la politique ne change trop rapidement et conduit à un apprentissage plus stable.

2.4.5. Apprentissage par renforcement hiérarchique

Dans l'apprentissage par renforcement hiérarchique (Hierarchical Reinforcement Learning), l'agent apprend à décomposer une tâche complexe en une hiérarchie de sous-tâches plus petites et plus gérables (Saha et al., 2020). Cela peut faciliter l'apprentissage pour l'agent et lui permettre de résoudre des tâches qui seraient difficiles à apprendre avec une approche classique d'apprentissage par renforcement. Parmi les algorithmes d'apprentissage par renforcement hiérarchique on peut citer : Option Critic et l'apprentissage féodal.

Option Critic est un algorithme d'apprentissage par renforcement hiérarchique qui apprend à décomposer une tâche en une hiérarchie d'options, chacune correspondant à une sous-tâche. Il fonctionne en apprenant une fonction de valeur séparée pour chaque option et en utilisant ces fonctions de valeur pour apprendre la politique optimale (Pateria et al., 2021).

L'apprentissage féodal est un autre algorithme d'apprentissage par renforcement hiérarchique qui décompose une tâche en une hiérarchie de sous-tâches (Dayan et Hinton, 1992). Il fonctionne en apprenant une politique séparée pour chaque niveau de la hiérarchie et en utilisant ces politiques pour apprendre la politique globale de la tâche (Vezhnevets et al., 2017).

2.4.6. Apprentissage par renforcement profond

L'apprentissage par renforcement profond (Deep Reinforcement Learning) est un type d'apprentissage par renforcement qui utilise des réseaux de neurones profonds pour apprendre la politique optimale (Mousavi et al., 2018). Dans l'apprentissage par renforcement profond, la fonction de valeur ou la politique de l'agent est représentée par un réseau de neurones profonds, qui est entraîné pour prédire la récompense attendue pour un état ou une action donnée

(Mousavi et al., 2017). Deep Deterministic Policy Gradients, A3C, Rainbow, et AlphaGo sont des exemples du deep reinforcement learning.

Deep Deterministic Policy Gradients (DDPG) est un algorithme d'apprentissage par renforcement profond qui combine les idées de Q-learning et des méthodes de gradient de politique (Sanghi et Sanghi, 2021). Il fonctionne en apprenant une politique déterministe qui mappe les états aux actions, et en utilisant un réseau critique pour évaluer la fonction de valeur d'action (Xu et al., 2022).

A3C ou Asynchronous Advantage Actor-critic est un modèle d'apprentissage par renforcement profond qui utilise plusieurs agents parallèles pour apprendre une politique. Il repose sur l'utilisation d'acteurs pour collecter des expériences et d'un critique pour évaluer la fonction de valeur de l'action (Babaeizadeh et al., 2016).

Rainbow est un algorithme d'apprentissage par renforcement profond qui combine plusieurs techniques de la littérature, notamment double DQN, relecture d'expérience priorisée et Q-learning de distribution. Il utilise ces techniques pour améliorer la stabilité et les performances du processus d'apprentissage (Fedus et al., 2020).

AlphaGo est un algorithme d'apprentissage par renforcement profond qui a été développé par DeepMind pour jouer au jeu de Go. Il fait appel à une combinaison d'apprentissage supervisé et d'apprentissage par renforcement pour apprendre la politique optimale (Halcomb et al., 2018).

2.5. Deep Learning

L'apprentissage profond ou le Deep Learning est un sous-domaine de l'apprentissage automatique qui s'inspire de la structure et de la fonction du cerveau (Lee, 2018). Les algorithmes de deep learning sont basés sur des réseaux de neurones, qui sont composés de plusieurs "neurones" interconnectés capables d'apprendre à reconnaître des motifs dans les données (Nikam, 2015). Les algorithmes de deep learning ont connu un grand succès dans une large gamme de tâches, telles que la reconnaissance d'images et de la parole, le traitement du langage naturel et la traduction automatique.

Il existe de nombreux types différents d'algorithmes de deep learning, et l'algorithme spécifique le mieux adapté à un problème donné dépendra des caractéristiques des données et des objectifs du modèle. Certains types courants d'algorithmes de deep learning comprennent :

les réseaux de neurones convolutionnels, les réseaux de neurones récurrents, les autoencodeurs, et les réseaux génératifs antagonistes.

2.5.1. Réseaux de neurones convolutionnels

Les réseaux de neurones convolutionnels (CNN) sont un type de réseau de neurones particulièrement bien adapté au traitement de données avec une topologie en grille, comme les images (Namatēvs, 2017). Ils sont composés de multiples couches de convolution et de regroupement, qui sont conçues pour extraire des caractéristiques des données d'entrée (Sideratos et al., 2020). Voici quelques-unes des techniques CNN les plus courantes : la descente de gradient stochastique, la rétropropagation, et Rectified Linear Unit.

La descente de gradient stochastique (Stochastic Gradient Descent, SGD) est une technique d'optimisation utilisé pour entraîner des réseaux de neurones en mettant à jour les poids et les biais du réseau en fonction de l'erreur entre la sortie prédite et la sortie réelle (Lee at al., 2019).

La rétropropagation (Backpropagation) est un algorithme utilisé pour calculer le gradient de la fonction de perte par rapport aux poids du réseau, permettant ainsi de mettre à jour les poids pendant l'entraînement (Yang et al., 2021).

ReLU ou Rectified Linear Unit est une fonction d'activation utilisée dans les CNN qui permet au réseau d'apprendre des relations non linéaires dans les données (Madhu et al., 2023).

2.5.2. Réseaux de neurones récurrents

Les réseaux de neurones récurrents (RNN) sont un type de réseau de neurones qui convient bien au traitement de données séquentielles, telles que les séries chronologiques ou le langage naturel (Carreon-Ortiz et al., 2023). Ils sont composés de plusieurs couches de cellules récurrentes, conçues pour capturer les dépendances entre les données d'entrée (Lai et al., 2018). Les algorithmes RNN les plus courants sont :

LSTM (Long Short-Term Memory), mémoire à long terme et courte mémoire, est type de réseau de neurones récurrent qui est capable de se souvenir des informations du passé pendant une période plus longue (Oruh et al., 2022), ce qui le rend adapté aux tâches telles que le traitement du langage naturel et la reconnaissance vocale.

GRU (Gated Recurrent Unit) ou unité récurrente à portes, est un autre type de réseau de neurones récurrent qui est similaire à LSTM mais qui est généralement plus rapide à entraîner et possède moins de paramètres (Zhou et al., 2016).

La propagation arrière à travers le temps (Backpropagation through Time, BPTT) est un algorithme utilisé pour entraîner les réseaux de neurones récurrents (RNN) (Ahmad et al., 2004). Il consiste à dérouler le RNN et à le traiter comme un réseau de neurones feedforward, puis à utiliser la rétropropagation pour calculer le gradient de la fonction de perte par rapport aux poids du réseau (Zhang et Li, 2019).

La propagation arrière tronquée dans le temps (Truncated backpropagation through time, TBPTT) est une variante de BPTT qui ne fait reculer qu'un nombre limité d'étapes temporelles plutôt que l'ensemble de la séquence de données (Tallec, 2019). Cela peut être plus efficace et réduire le risque de gradients qui disparaissent, mais cela peut également entraîner un modèle moins précis (Struye et Latré, 2020).

L'apprentissage récurrent en temps réel (Real-time Recurrent Learning, RTRL) est un autre algorithme utilisé pour entraîner les RNN (Mak et al., 1999). Il consiste à calculer le gradient de la fonction de perte par rapport aux poids du réseau en temps réel, permettant aux poids d'être mis à jour à chaque étape temporelle pendant l'entraînement (Mujika et al., 2018).

Les moindres carrés récursifs (Recursive Least Squares, RLS) : est un algorithme utilisé pour adapter les poids d'un RNN dans un contexte d'apprentissage en ligne (Chow et Cho, 1997). RLS consiste à mettre à jour de manière itérative les poids du réseau en fonction de l'erreur entre la sortie prédite et la sortie réelle, en utilisant une méthode des moindres carrés récursive (Zhou et al., 2018).

2.5.3. Autoencodeurs

Les autoencodeurs sont un type de réseau de neurones qui est entraîné à reconstruire les données d'entrée à partir d'une représentation de dimension inférieure, appelée représentation latente. Ils sont composés d'un encodeur et d'un décodeur, qui sont entraînés à compresser et à reconstruire les données d'entrée, respectivement (Wang et al., 2022). Les algorithmes utilisés dans les autoencodeurs sont les suivants :

Stochastic gradient descent (SGD) est un algorithme d'optimisation couramment utilisé pour entraîner des réseaux de neurones, y compris les autoencodeurs (Rajesh et al., 2023). Il consiste à mettre à jour de manière itérative les paramètres du modèle en fonction du gradient de la fonction de perte par rapport aux paramètres.

Backpropagation est un algorithme utilisé pour calculer le gradient de la fonction de perte par rapport aux paramètres du modèle dans un réseau de neurones (Rosenbaum, 2022). Il est utilisé en conjonction avec des algorithmes d'optimisation tels que SGD pour mettre à jour les paramètres du modèle pendant l'entraînement (Lee et Park, 2023).

Les fonctions d'activation sont utilisées pour introduire de la non-linéarité dans le modèle et sont souvent utilisées dans les couches cachées d'un autoencodeur (Halder et al., 2022). Les fonctions d'activation courantes incluent sigmoid, tanh et ReLU (unité linéaire rectifiée) (Banerjee et al., 2019).

Les fonctions de perte sont utilisées pour mesurer l'erreur entre la sortie prédite et la sortie réelle pendant l'entraînement (Kendall et Cipolla, 2017). Les fonctions de perte couramment utilisées dans les autoencodeurs comprennent l'erreur quadratique moyenne (MSE) et l'entropie croisée binaire.

Les techniques de régularisation sont utilisées pour éviter le surapprentissage, qui se produit lorsque le modèle est trop étroitement ajusté aux données d'entraînement et ne se généralise pas bien aux données invisibles (Kernbach et Staartjes, 2022). Les techniques de régularisation courantes comprennent la décroissance de poids et le dropout.

2.5.4. Réseaux génératifs antagonistes

Les réseaux génératifs antagonistes (GAN) sont un type de réseau neuronal qui se compose de deux réseaux : un générateur et un discriminateur (Kim et al., 2020). Le générateur est formé pour produire des échantillons synthétiques similaires aux données cibles, tandis que le discriminateur est formé pour distinguer entre les échantillons synthétiques et les échantillons réels (Wan et Jones, 2020). Le générateur et le discriminateur sont formés simultanément de manière antagoniste, le générateur essayant de tromper le discriminateur et le discriminateur essayant de classer avec précision les échantillons (Goodfellow et al., 2020). Les GAN utilisent plusieurs méthodes y compris :

L'optimisation alternée est une méthode d'entraînement qui consiste à optimiser de manière itérative l'un des deux réseaux dans un GAN (soit le générateur, soit le discriminateur) tout en maintenant l'autre réseau fixe (Liu et al., 2023). Ce processus est répété jusqu'à ce que les deux réseaux soient formés jusqu'à la convergence. L'optimisation alternée est utilisée dans les GAN pour former simultanément le générateur et le discriminateur de manière antagoniste, le générateur essayant de produire des échantillons similaires aux échantillons réels et le discriminateur essayant de distinguer avec précision les échantillons synthétiques des échantillons réels (Zhao et al., 2018).

Wasserstein GAN (WGAN) sont une variante des GAN qui utilisent la distance de Wasserstein comme fonction de perte pour le générateur et le discriminateur. Cette distance est une mesure de la distance entre deux distributions de probabilité, et elle présente l'avantage d'être plus stable et plus facile à optimiser que d'autres mesures de distance utilisées dans les GAN (Weng, 2019). Les WGAN sont moins sujets à l'effondrement de mode, un problème dans lequel le générateur produit des échantillons synthétiques qui sont tous similaires les uns aux autres et pas diversifiés (Wiatrak et al., 2019).

Improved Wasserstein GAN (IWGAN) ou les GAN améliorés de Wasserstein sont une amélioration des WGAN qui utilisent une technique de découpage de poids pour stabiliser davantage le processus d'entraînement (Ye et al., 2018). Dans les IWGAN, les poids du discriminateur sont découpés dans une certaine plage pour éviter que le discriminateur ne devienne trop puissant et domine le générateur. Cela peut aider à prévenir l'effondrement de mode et améliorer la diversité des échantillons générés.

Least squares GAN (LSGAN) ou les GAN à moindres carrés sont une variante des GAN qui utilisent une fonction de perte à moindres carrés pour le générateur et le discriminateur. Dans les LSGAN, le générateur est entraîné à produire des échantillons similaires aux échantillons réels, tandis que le discriminateur est entraîné à distinguer entre les échantillons synthétiques et réels en minimisant l'erreur à moindres carrés entre les étiquettes prédites et réelles (Ataş, 2023). Les LSGAN peuvent mieux fonctionner que d'autres GAN pour certaines tâches, telles que la génération d'images, en raison des propriétés de la fonction de perte à moindres carrés.

2.6. Transfer learning

Le transfer learning est une technique d'apprentissage automatique qui consiste à utiliser un modèle pré-entraîné pour une nouvelle tâche, dans le but d'améliorer les performances du nouveau modèle en exploitant les connaissances acquises à partir de la tâche d'origine (Subramanian et al., 2022). Il est particulièrement utile lorsque la nouvelle tâche a un domaine ou un contexte similaire à la tâche d'origine et lorsqu'il y a une pénurie de données étiquetées pour la nouvelle tâche (Zhuang et al., 2020). Il existe plusieurs types d'algorithmes de transfert d'apprentissage, notamment : le fine-tuning, l'extraction de caractéristiques, et le multi-task learning.

2.6.1. Fine-tuning

Le fine-tuning consiste à ajuster les poids du modèle pré-entraîné pour la nouvelle tâche, en utilisant une petite quantité de données étiquetées de la nouvelle tâche. Le modèle pré-entraîné sert de point de départ et la nouvelle tâche est utilisée pour "affiner" les poids du modèle (Araci, 2019). Il existe plusieurs algorithmes qui peuvent être utilisés pour le fine-tuning, tels que :

Adam (Adaptive Moment Estimation) est un algorithme d'optimisation qui combine les idées de SGD et de momentum (Wang et al., 2019). Il peut être utilisé pour affiner un modèle pré-entraîné en mettant à jour les paramètres du modèle en fonction des moyennes mobiles pondérées exponentiellement du gradient et du carré du gradient (Liu et al., 2020).

Resilient Propagation (RProp) est un algorithme d'optimisation qui utilise le signe du gradient pour mettre à jour les paramètres du modèle (Chen, 1999). RProp permet d'affiner un modèle pré-entraîné en mettant à jour les paramètres du modèle dans la direction qui réduit la perte (Ur Rehman et al., 2019).

L-BFGS ou Limited-memory BFGS est un algorithme d'optimisation qui approxime la matrice de Hesse en utilisant un nombre limité de gradients précédents (Loshchilov, 2017). L-BFGS est employé pour affiner un modèle pré-entraîné en mettant à jour les paramètres du modèle en fonction de l'approximation de la matrice de Hesse (Gupta et al., 2019).

2.6.2. Extraction de caractéristiques

L'extraction de caractéristiques implique l'utilisation du modèle pré-entraîné en tant qu'extracteur de caractéristiques fixe, où la sortie des couches du modèle pré-entraîné est utilisée en entrée d'un nouveau modèle qui est entraîné à effectuer la nouvelle tâche (El-Alami et al., 2022). Plusieurs algorithmes peuvent être utilisés pour l'extraction de caractéristiques, y compris :

Les machines à vecteurs de support (SVM) sont un type d'algorithme d'apprentissage supervisé qui peut être utilisé pour l'extraction de caractéristiques. Elles trouvent l'hyperplan dans l'espace des caractéristiques qui sépare de manière maximale les différentes classes (Lin, 2021).

Les forêts aléatoires sont une méthode d'apprentissage par ensemble qui peut être utilisée pour l'extraction de caractéristiques. Elles construisent un ensemble d'arbres de décision et font des prédictions en se basant sur le vote majoritaire des arbres individuels (Islam et Nahiduzzaman, 2022).

Les k plus proches voisins (k-Nearest Neighbor, k-NN) est un algorithme d'apprentissage par instance simple qui peut être utilisé pour l'extraction de caractéristiques. Il classe un nouvel échantillon en se basant sur la classe majoritaire de ses k voisins les plus proches dans l'espace des caractéristiques (Yang et al., 2020).

Naive Bayes est un classificateur probabiliste qui peut être utilisé pour l'extraction de caractéristiques (He et al., 2019). Il effectue des prédictions en se basant sur la probabilité d'une classe particulière étant donné les caractéristiques (Bakheet et Al-Hamadi, 2021).

2.6.3. Multi-task learning

Le multi-task learning ou apprentissage multi-tâches consiste à entraîner un seul modèle à effectuer plusieurs tâches simultanément, en utilisant une combinaison de couches partagées et spécifiques à chaque tâche (Gong et al., 2019). Les couches partagées sont pré-entraînées sur une ou plusieurs tâches et sont utilisées pour extraire des caractéristiques communes qui sont pertinentes pour toutes les tâches, tandis que les couches spécifiques à chaque tâche sont entraînées pour s'adapter aux caractéristiques spécifiques de chaque tâche (Liu et al., 2017). Plusieurs modèles qui peuvent être employés pour l'apprentissage multi-tâches, tels que :

Les réseaux de neurones sont un type de modèle d'apprentissage automatique qui peuvent être utilisés pour l'apprentissage multi-tâches (Ruder, 2017). Ils se composent de plusieurs couches de nœuds interconnectés et sont capables d'apprendre des relations complexes entre l'entrée et la sortie.

Le boosting de gradient (Gradient Boosting) est une méthode d'apprentissage ensembliste qui peut être utilisée pour l'apprentissage multi-tâches (Ferreira et al., 2022). Elle consiste à entraîner séquentiellement des modèles faibles et à combiner leurs prédictions pour former un modèle plus fort (Zhang et Haghani, 2015).

Transfert learning avec des réseaux adversaires est une approche consiste à utiliser un réseau antagoniste génératif pour transférer les connaissances d'un modèle pré-entraîné à un nouveau modèle (Soleimani et Nazerfard, 2021).

2.7. Traitement Automatique du Langage Naturel

Le traitement automatique du langage naturel ou le Natural Language Processing (NLP) est un sous-ensemble de la linguistique, de l'informatique et de l'intelligence artificielle associé aux interactions entre les ordinateurs et le langage humain, et plus précisément à la manière dont les ordinateurs sont programmés pour traiter et interpréter d'énormes quantités de données en langage naturel (Chowdhary et Chowdhary, 2020). Il vise à créer des logiciels qui produisent et comprennent des langages naturels (Yildirim, 2022). Au lieu de se contenter de programmer ou d'utiliser des langages artificiels tels que Java ou C, un utilisateur peut avoir des discussions naturelles avec sa machine (Darwish et al., 2021).

Le NLP couvre de nombreux cas d'utilisation y compris : la classification de textes, la reconnaissance de texte, le résumé automatique, la traduction automatique, les chatbots, voicebots ou callbots, les assistants intelligents, etc. (Boonstra, 2021).

Le NLP rassemble l'IA et le traitement du langage. La dernière génération de technologies NLP est basée sur des réseaux neuronaux artificiels ou des modèles statistiques simples du machine learning (Lauriola et al., 2022). Les modèles d'apprentissage ont généralement dû être entraînés sur de grands volumes de texte. L'objectif peut viser plusieurs types de traitements automatiques : speech-to-text et text-to-speech, reconnaissance d'entités nommées (noms de personnes, de lieux...), analyse de sentiments (sentiment positif, négatif, ou neutre), résumé de

texte, extraction d'aspects (ciblant l'intention du texte) ou modélisation de thèmes (Millstein, 2020).

De façon générale, le traitement automatique du langage naturel peut être réparti en deux principales classes de modèles machine learning (Torfi et al., 2020) :

- Les modèles machine learning orientés NLU (natural language understanding) qui s'attachent à saisir le sens d'une langue et d'un discours dans son contexte (Ni et al., 2020),
- Les modèles machine learning orientés NLG (natural language generation) qui ont pour but de générer un texte à la manière d'un humain (Ji et al., 2023).

3. Méthodes d'évaluation des modèles machine learning

L'évaluation des modèles machine learning est une étape importante dans le processus de développement de tout système machine learning. Elle permet de mesurer les performances et l'exactitude du modèle par rapport aux données d'entraînement et de test. Le choix de la méthode d'évaluation dépend généralement du type d'apprentissage, de la taille et de la complexité des données.

3.1. Evaluation des algorithmes d'apprentissage supervisé

Les algorithmes d'apprentissage supervisé sont entraînés sur des données étiquetées, où les données d'entrée sont accompagnées d'une étiquette de sortie correspondante. Il existe deux types principaux d'apprentissage supervisé : la classification et la régression.

3.1.1. Classification

La classification est le processus de prédiction d'une étiquette discrète en fonction des données d'entrée. L'évaluation des algorithmes de classification implique généralement de diviser les données disponibles en un ensemble d'entraînement et un ensemble de test (Bhavsar et Ganatra, 2012). L'ensemble d'entraînement est utilisé pour entraîner le modèle d'apprentissage automatique et l'ensemble de test est utilisé pour évaluer ses performances. Les mesures pour évaluer les performances des algorithmes de classification incluent :

- Précision : la fraction des vrais positifs prédits par le modèle parmi toutes les prédictions positives. La précision est calculée comme suit :

$$\text{Précision} = \frac{\text{nombre de vrais positifs}}{\text{nombre de vrais positifs} + \text{nombre de faux positifs}} \quad (1)$$

- Rappel : la fraction des vrais positifs prédits par le modèle parmi tous les cas positifs réels. Le rappel est calculé comme suit :

$$\text{Rappel} = \frac{\text{nombre de vrais positifs}}{\text{nombre de vrais positifs + nombre de faux négatifs}} \quad (2)$$

- F1 score : la moyenne harmonique de la précision et du rappel, souvent utilisée comme une seule mesure pour évaluer les performances d'un modèle. Le score F1 est calculé comme suit:

$$\text{F1 score} = 2 * \frac{\text{Précision} * \text{Rappel}}{\text{Précision+ Rappel}} \quad (3)$$

Il est important de considérer l'équilibre entre la précision et le rappel lors de l'évaluation des performances d'un modèle de classification. Un modèle à haute précision peut avoir un rappel faible et vice versa (Sohony et al., 2018). Le score F1 prend en compte à la fois la précision et le rappel et est une mesure utile pour comparer les performances de modèles avec différents compromis précision-rappel (Grandini et al., 2020).

Ces mesures peuvent être calculées à l'aide d'une matrice de confusion, qui est une table qui montre le nombre de vrais positifs, de vrais négatifs, de faux positifs et de faux négatifs prédits par un modèle de classification. Elle est couramment utilisée pour évaluer les performances des algorithmes de classification et peut fournir un aperçu des types d'erreurs que le modèle commet. La matrice de confusion comporte quatre cellules, correspondant chacune à une combinaison des étiquettes de classe prédites et réelles du modèle :

- Vrai positif (VP) : le nombre d'instances où le modèle a correctement prédit la classe positive.
- Vrai négatif (VN) : le nombre d'instances où le modèle a correctement prédit la classe négative.
- Faux positif (FP) : le nombre d'instances où le modèle a incorrectement prédit la classe positive (également connu sous le nom d'"erreur de type I").
- Faux négatif (FN) : le nombre d'instances où le modèle a incorrectement prédit la classe négative (également appelée "erreur de type II").

Ces quatre valeurs peuvent être utilisées pour calculer diverses mesures d'évaluation pour les algorithmes de classification, telles que la précision, le rappel et le score F1.

Il est crucial de considérer l'importance relative des différents types d'erreurs lors de l'évaluation des performances d'un modèle de classification. La matrice de confusion peut

également être utile pour comprendre les caractéristiques de l'ensemble de données et identifier les problèmes potentiels avec le modèle ou les données (Boyce et al., 2002). Par exemple, un modèle qui fait systématiquement des faux positifs peut souffrir d'un déséquilibre de classes ou être sur-ajusté aux données d'entraînement.

3.1.2. Régression

Dans les problèmes de régression, il s'agit du processus de prédiction d'une valeur continue (par exemple, le prix des actions) basée sur des données d'entrée. L'évaluation des algorithmes de régression peut être effectuée en utilisant des métriques telles que l'erreur moyenne absolue, l'erreur quadratique moyenne et l'erreur quadratique moyenne racine, qui mesurent la différence entre les valeurs prédites et les valeurs réelles (Priya Varshini et Anitha Kumari, 2020).

- Erreur moyenne absolue (MAE) : La différence absolue moyenne entre les valeurs prédites et les valeurs réelles. Le MAE est calculé comme suit :

$$\mathrm{MAE} = \frac{1}{\mathrm{n}}\sum|y - y'| \tag{4}$$

Où n est le nombre d'échantillons, y est la valeur réelle et y' est la valeur prédite.

- Erreur quadratique moyenne (MSE) : La différence quadratique moyenne entre les valeurs prédites et les valeurs réelles. Le MSE est calculé comme suit :

$$\mathrm{MSE} = \frac{1}{\mathrm{n}}\sum|y - y'|^2 \tag{5}$$

Où n est le nombre d'échantillons, y est la valeur réelle et y' est la valeur prédite.

- Erreur quadratique moyenne racine (RMSE) : La racine carrée de l'erreur quadratique moyenne. Le RMSE est calculé comme suit :

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}} \tag{6}$$

Le RMSE est couramment utilisé comme mesure de l'erreur globale du modèle.

- Le coefficient de détermination, également connu sous le nom de R carré ou du coefficient de corrélation de Pearson au carré, est une autre métrique couramment utilisée pour évaluer la performance des algorithmes de régression. Le R carré est une mesure de la force de la relation entre les valeurs prédites et les valeurs réelles (Onyutha, 2022). Il varie de 0 à 1, une valeur de 1 indiquant un ajustement parfait et une valeur de 0 indiquant aucune relation. Le R carré est calculé comme suit :

$$\mathrm{R}^2 = 1 - \frac{\text{somme des erreurs au carré}}{\text{somme totale des erreurs au carré}} \tag{7}$$

Où la somme des erreurs au carré est la somme des différences au carré entre les valeurs prédites et les valeurs réelles, et la somme totale des erreurs au carré est la somme des différences au carré entre la moyenne des valeurs réelles et les valeurs réelles individuelles.

Le R carré peut être une métrique utile pour évaluer la performance des algorithmes de régression, en particulier lorsque la variable cible a une grande échelle ou lors de la comparaison de la performance de différents modèles (Chicco et al., 2021). Cependant, il est important de noter que le R carré est sensible au nombre de variables indépendantes dans le modèle et peut ne pas toujours refléter avec précision la véritable force de la relation (Gurumurthy et Mukherjee, 2020).

3.2. Evaluation des algorithmes d'apprentissage non supervisé

Les algorithmes d'apprentissage non supervisé n'ont pas de données étiquetées et sont utilisés pour identifier des modèles et des relations dans les données. Il existe deux types principaux d'apprentissage non supervisé : le clustering et l'association.

3.2.1. Clustering

Le clustering est le processus de regroupement des points de données en clusters en fonction de leur similarité. L'évaluation des algorithmes de regroupement peut être difficile car il n'y a pas d'étiquettes de référence pour comparer les prédictions du modèle (Badillo et al., 2020). Les techniques d'évaluation de la performance des algorithmes de clustering comprennent :

- Visualisation : l'utilisation de graphiques, de graphiques ou d'autres aides visuelles pour comprendre les modèles ou clusters identifiés par le modèle.
- Coefficient de silhouette : une mesure de la compacité et de la séparation des clusters identifiés par le modèle, avec une valeur plus élevée indiquant une meilleure performance (Bagirov et al., 2023).

3.2.2. Association

L'association est le processus de recherche de relations entre les variables dans les données. L'évaluation des algorithmes d'association peut être effectuée en utilisant des métriques telles que le lift et la conviction, qui mesurent la force des relations identifiées par le modèle (Abdirad et Mathur, 2021).

3.3. Evaluation des algorithmes d'apprentissage semi supervisé

Les algorithmes d'apprentissage semi-supervisé sont entraînés sur un mélange de données étiquetées et non étiquetées. Ce type d'apprentissage peut être utile lorsque seulement une quantité limitée de données étiquetées est disponible. Il existe deux principaux types d'apprentissage semi-supervisé : la classification et le clustering.

- L'évaluation des algorithmes de classification semi-supervisée peut être effectuée en utilisant les mêmes techniques que l'apprentissage supervisé, telles que la précision et le score F1 (Ligthart et al., 2021).
- L'évaluation des algorithmes de clustering semi-supervisés peut être effectuée en utilisant des techniques telles que la visualisation et le coefficient de silhouette (Morette et al., 2020).

3.4. Evaluation des algorithmes d'apprentissage par renforcement

Les algorithmes d'apprentissage par renforcement sont entraînés à prendre des actions dans un environnement afin de maximiser une récompense. Ces algorithmes sont souvent utilisés dans les systèmes de contrôle et les jeux. L'évaluation des algorithmes d'apprentissage par renforcement implique généralement la mesure de la récompense obtenue par le modèle au fil du temps (Dulac-Arnold et al., 2019). Les techniques d'évaluation de la performance des algorithmes d'apprentissage par renforcement comprennent :

- La récompense : la mesure du succès dans la tâche, généralement obtenue en prenant certaines actions dans l'environnement (Schoettler et al., 2020).
- La comparaison à une approche de base ou heuristique : la performance du modèle est souvent évaluée en comparant la récompense totale obtenue au fil du temps à une approche de base ou heuristique (Zhang et al., 2022). Cela peut fournir des informations sur l'efficacité de l'algorithme d'apprentissage par renforcement par rapport à une approche simple ou basée sur des règles.
- La stabilité et convergence : il est important de prendre en compte la stabilité et la convergence de l'algorithme d'apprentissage par renforcement. La stabilité fait référence à la capacité de l'algorithme à atteindre un niveau de performance satisfaisant sans osciller ou diverger (Chen et al., 2022). La convergence fait référence à la capacité de l'algorithme à s'améliorer au fil du temps et à atteindre une solution optimale (Zaman et Gharehchopogh, 2022). Les techniques d'évaluation de la stabilité et de la convergence des algorithmes

d'apprentissage par renforcement comprennent la représentation graphique de la récompense au fil du temps et l'analyse de la courbe d'apprentissage.

- L'évaluation humaine : dans certains cas, il peut être utile de faire évaluer la performance de l'algorithme d'apprentissage par renforcement par un expert humain (Cowley et al., 2022). Cela peut être particulièrement utile pour les tâches où la fonction de récompense est complexe ou subjective.

Conclusion

Dans ce chapitre, les définitions du e-marketing sont exposées, mettant en évidence l'évolution du marketing traditionnel vers le marketing numérique. Les différentes techniques du marketing digital sont abordées pour donner une compréhension approfondie des outils disponibles.

De plus, les stratégies e-marketing sont présentées en détail, en couvrant divers aspects. Ces stratégies permettent aux entreprises de développer des approches efficaces pour atteindre leurs objectifs marketing dans le monde numérique en constante évolution et révolution.

Le chapitre se poursuit en abordant les fondements théoriques du machine learning, avec un accent particulier sur les différentes méthodes d'apprentissage supervisé, non supervisé, semi-supervisé et par renforcement. Les concepts clés tels que les algorithmes de classification, de régression, de clustering et de réduction de dimensionnalité sont expliqués pour fournir une base solide de compréhension.

Enfin, le chapitre se termine par une partie consacrée aux méthodes d'évaluation des modèles machine learning, en examinant les différentes approches d'évaluation pour les algorithmes d'apprentissage supervisé, non supervisé, semi-supervisé et par renforcement. L'évaluation joue un rôle crucial dans l'identification de la performance des modèles et dans la prise de décisions informées dans le domaine du e-marketing.

À travers ce chapitre, une base conceptuelle solide du e-marketing et du machine learning a été établie, permettant la compréhension des principes fondamentaux et les techniques clés associées à ces domaines. Cette compréhension approfondie servira de fondement pour la revue de littérature à venir et pour les contributions originales présentées dans les chapitres suivants.

Chapitre 2

Application du machine learning dans le e-marketing, tendances et développements futurs : état de l'art

Sommaire

INTRODUCTION.....57
REVUE DE LITTERATURE PORTANT SUR L'APPLICATION DU MACHINE LEARNING DANS LE E-MARKETING 57
TENDANCES ET DEVELOPPEMENTS FUTURS DU MACHINE LEARNING DANS LE MARKETING DIGITAL.....87
CONCLUSION.....95

Chapitre 2. Application du machine learning dans le e-marketing, tendances et développements futurs : état de l'art

Le domaine du marketing digital a connu une évolution rapide ces dernières années, avec l'avènement de nouvelles technologies telles que l'intelligence Artificielle et le machine learning. Les 7P du marketing digital sont considérés comme des éléments clés pour développer une stratégie marketing en ligne efficace (Pistol et Bucea-Manea, 2017). En outre, le machine learning peut aider à optimiser les stratégies e-marketing en fournissant des analyses de données plus approfondies et plus précises, automatisant certains processus marketing, et permettant ainsi de prendre des décisions plus éclairées.

Dans ce contexte, un aperçu de la littérature existante sur l'utilisation du machine learning dans le cadre des 7P de l'e-marketing est donné dans la première partie de ce chapitre. Les principales études et leurs résultats sont analysés, et les lacunes de la littérature sont identifiées. Les 7P, comprenant le produit (Product), le prix (Price), le placement (Place), la promotion (Promotion), les personnes (People), le processus (Process) et les preuves matérielles (Physical Evidence) (Ho et al., 2022), sont examinés. Dans la deuxième partie, des tendances et développements futurs du machine learning dans le e-marketing sont explorés.

1. Revue de littérature portant sur l'application du machine learning dans le e-marketing

Cette revue de littérature est basée sur les 7P de l'e-marketing comme montré dans la figure 1.

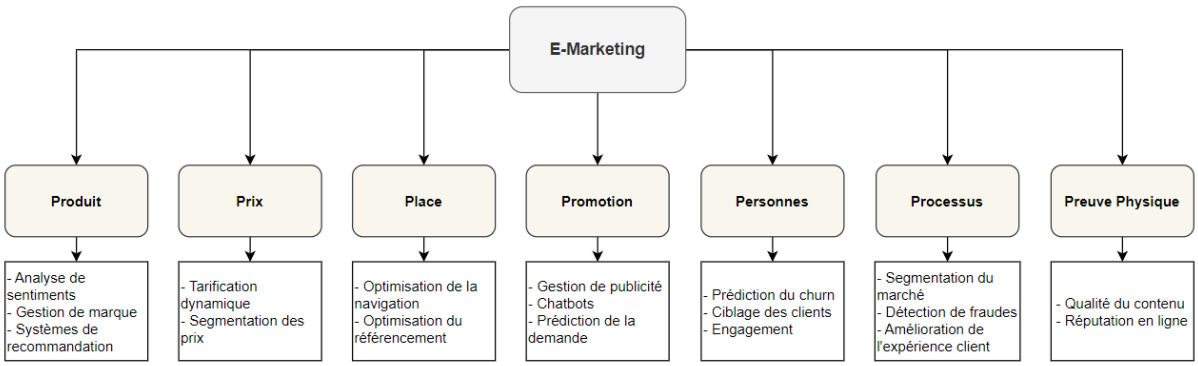


Figure 1. « Les 7P de l'e-marketing ».

1.1. Produit

« Le produit » est un élément important du mix marketing des 7P. Le produit ou service proposé en ligne est considéré comme la pierre angulaire de la stratégie e-marketing (Shaltoni, 2016). Il doit être de qualité et répondre aux besoins des clients. Le prix doit être déterminé pour maximiser les bénéfices tout en restant compétitif sur le marché. La distribution en ligne est essentielle et les canaux de distribution appropriés doivent être choisis pour atteindre le public cible (Scott, 2009), tels que les réseaux sociaux, les marketplaces en ligne, les sites web, etc.

La littérature existante applique les technologies ML pour analyser les activités de marketing liées au produit, notamment les recommandations de produits, la gestion des marques et des marques déposées, et la prédiction des décisions d'achat. Ces applications sont décrites ci-dessous.

1.1.1. Analyse de sentiments

L'analyse des sentiments des clients est devenue un défi majeur pour les entreprises ces dernières années. Le machine learning est largement utilisé pour analyser les commentaires des clients sur les réseaux sociaux et les forums en ligne. Cette analyse permet de comprendre les opinions et les sentiments des clients à l'égard de l'entreprise et de ses produits. Les algorithmes de traitement du langage naturel (NLP) peuvent être utilisés pour extraire des informations sur les sentiments des clients et les sujets de discussion. Cette analyse peut aider les entreprises à répondre aux besoins des clients et à améliorer leur image de marque.

Il existe plusieurs techniques machine learning pour l'analyse des sentiments des clients, notamment :

La classification binaire : cette technique consiste à classifier les commentaires des clients en deux catégories, positives ou négatives. Les algorithmes utilisent des ensembles de données annotés pour entraîner des modèles qui peuvent prédire si un commentaire est positif ou négatif.

L'analyse de polarité : cette technique permet de classifier les commentaires en plusieurs catégories de polarité, allant de très négatif à très positif. Les algorithmes machine learning utilisent des techniques de traitement du langage naturel pour extraire les sentiments des commentaires et les classer en fonction de leur polarité.

L'analyse de sujet : cette technique consiste à identifier les sujets de discussion dans les commentaires des clients. Les algorithmes machine learning utilisent des techniques de NLP pour extraire les sujets principaux des commentaires et les classer en fonction de leur importance.

Ces dernières années, de nombreuses études se sont intéressées à l'analyse des sentiments et à l'exploration des opinions (Keikhosrokiani et Pourya Asl, 2022).

Rain (2013) a mis en œuvre et a étendu les travaux actuels en matière de traitement du langage naturel et d'analyse du sentiment à des ensembles de données d'évaluation d'Amazon. Ils ont utilisé des classificateurs bayésiens naïfs et des listes de décision pour étiqueter une évaluation donnée comme positive ou négative.

Xu et al. (2015) ont expérimenté différents algorithmes de sélection de caractéristiques et d'apprentissage supervisé pour prédire les évaluations (nombre d'étoiles) de l'ensemble de données Yelp à partir du texte des avis. Ils ont évalué les performances des différents modèles sur la base de mesures de précision et de rappel. Cette étude montre que l'algorithme binarisé de Naive Bayes, combiné à la sélection des caractéristiques, à la suppression des mots vides et au stemming, est le plus performant dans leur contexte d'analyse des sentiments.

Chen et al. (2019) ont proposé la méthodologie TF-IDF et FPCDA phrase FE pour l'analyse du sentiment des critiques de produits. En tenant compte des différentes longueurs des critiques de produits, les modèles locaux des vecteurs de caractéristiques ont été découverts en utilisant l'algorithme de bi-clustering OPSM. Le Prefix Span a été créé pour reconnaître les phrases fréquentes et les phrases pseudo-consécutives avec un niveau élevé de discrimination et d'information sur l'ordre des mots. En outre, la séparation et la capacité de discrimination des mots ont été utilisées pour améliorer la capacité de la polarité des sentiments à distinguer les différents sentiments. L'extraction des caractéristiques du texte est venue ensuite. La progression de l'expérience et les résultats analogiques ont démontré une amélioration des performances de l'extraction d'opinion sur les revues de produits. Malheureusement, la précision de la classification est inférieure lors de l'extraction des caractéristiques textuelles.

Mukherjee et Bhattacharyya (2012) ont utilisé une analyse des sentiments spécifique aux caractéristiques pour étudier l'évaluation des produits. Pour déterminer la relation entre les caractéristiques et les opinions auxquelles elles sont liées, une technique d'analyse des

dépendances est appliquée. Ils ont créé un système qui rassemble des expressions d'opinion définissant diverses caractéristiques prospectives à partir de revues et les extraites.

Nagarajan et Gandhi (2019) utilisent des techniques d'hybridation pour catégoriser les flux de données twitter sur la base d'une analyse de sentiment. La catégorisation de l'analyse du sentiment utilise l'algorithme génétique, l'optimisation par essaim particulaire et l'algorithme d'arbre de décision. 600 tweets sont collectés pour la classification de l'analyse du sentiment avec l'utilisation de la génération de caractéristiques et des outils de sécurité basés sur l'URL. En fusionnant la technique de réduction dimensionnelle avec la régression logistique dans l'analyse des composantes principales, ils ont développé l'extraction des caractéristiques des données. Les spammeurs sur Twitter sont repérés à l'aide d'une approche basée sur les statistiques linéaires.

La procédure de catégorisation de la polarité des sentiments est utilisée par Fang et Zhan (2015) pour décrire un système d'analyse des sentiments pour les critiques de produits. Trois phases composent l'ensemble du processus. Les techniques de catégorisation choisies sont la méthode bayésienne naïve, la machine à vecteur de support et la forêt aléatoire. La phase 1 du processus d'évaluation consiste à supprimer le contenu objectif et à extraire le contenu subjectif des données. Effectuez le marquage POS sur le contenu extrait après l'extraction. Choisir entre les phases de sentiment de négation de l'adjectif (NOA) et de négation du verbe pendant la phase 2 (NOV). Calculer également le score de sentiment pour les jetons de sentiment. Le vecteur de caractéristiques pour le sentiment est construit à l'aide de la formule de score de sentiment. La catégorisation de la polarité des sentiments de la phase 3 est la dernière étape.

Yang et al. (2020) ont créé des modèles de la famille CNN-LSTM (Convolutional Neural Network Long Short-Term Memory) à deux canaux intégrés lexicalement pour l'analyse des sentiments, qui sont des modèles centrés sur l'apprentissage profond pour l'analyse des sentiments. Pour créer un échantillon de données d'entrée d'une taille fiable et développer la proportion de données de sentiment dans chaque critique, la méthodologie du sentiment padding a été utilisée. Le sentiment padding a permis de résoudre le problème de disparition du gradient qui pouvait survenir lors de l'utilisation d'un padding "0" entre la couche d'entrée et la première couche cachée. Des composants de lexique de première qualité ont été développés pour l'analyse des sentiments afin d'être utilisés dans l'opération de remplissage des sentiments.

De nombreuses études ont montré que la fourniture d'un modèle parallèle "à deux canaux" et d'informations sur le lexique des sentiments améliorait la précision de l'analyse des sentiments.

Selon l'approche proposée par Sindhu et Mukherjee (2021), la crédibilité d'un évaluateur est déterminée par son lien étroit ou professionnel avec la catégorie de produits évaluée.

Haque et al. (2018) ont utilisé une approche d'apprentissage supervisé sur un ensemble de données Amazon à grande échelle pour le biaiser et obtenir une grande précision. Ils ont utilisé une approche manuelle et une approche d'apprentissage actif pour étiqueter les ensembles de données. Dans le processus d'apprentissage actif, ils ont utilisé différents classificateurs pour atteindre une précision satisfaisante. Les résultats de cette étude ont montré que le modèle de machine à vecteur de support a donné les meilleurs résultats de classification.

Jagdale et al. (2019) ont également utilisé un ensemble de données provenant d'Amazon qui contient des avis sur différents produits tels que des appareils photo, des ordinateurs, des smartphones, des tablettes, etc. Ils ont utilisé des algorithmes d'apprentissage automatique pour classer les avis qui sont positifs ou négatifs. Les résultats de leur étude ont montré que les techniques d'apprentissage automatique donnent les meilleurs résultats pour classer les avis sur les produits.

Les résultats de l'analyse des sentiments peuvent être utilisés pour améliorer les stratégies de marketing et de communication. Les entreprises peuvent utiliser les commentaires négatifs pour identifier les problèmes et y remédier, améliorer la qualité de leurs produits et services, et ainsi fidéliser leurs clients. De plus, les commentaires positifs peuvent être utilisés pour promouvoir l'image de marque de l'entreprise et renforcer la confiance des clients.

1.1.2. Gestion de marque

Étant donné que les marques et les marques déposées jouent un rôle majeur dans le marketing et améliorent l'équité de marque et la compétitivité mondiale des entreprises (Trappey et al., 2020), les chercheurs s'intéressent à la gestion des marques et des marques déposées.

Certains chercheurs explorent les photos de marque publiées sur les réseaux sociaux pour prédire la fidélité à la marque (Kaiser et al., 2020) et mesurer la manière dont les marques sont représentées (Liu et al., 2020). En plus du contenu visuel d'une marque, les chercheurs explorent

le contenu textuel des marques pour analyser la réputation de la marque (Ducange et al., 2019), le bouche à oreille électronique digne de réponse (eWOM) des marques (Vermeer et al., 2019), la valence et le volume pertinents pour la marque (Kübler et al., 2020), la concurrence entre les marques (Tirunillai et Tellis, 2014) et les traits de personnalité de la marque (Chen et al., 2015). Carpineto et Romano (2020) se concentrent sur le domaine de recherche de marque pour détecter les sites de commerce électronique présents dans la liste des résultats de recherche de marque et les sites frauduleux, ce qui est utile pour des actions ciblées contre la contrefaçon.

Pour les marques déposées, Trappey et al. (2020) développent des modèles avancés d'évaluation de la similarité des marques déposées couvrant les similarités dans l'orthographe, la prononciation et l'image pour protéger les marques déposées.

1.1.3. Systèmes de recommandation

La recommandation de produits est une technique qui vise à suggérer des produits complémentaires ou similaires à un client en se basant sur ses achats précédents, ses préférences et ses interactions avec la marque. Cette technique est souvent utilisée pour améliorer l'expérience d'achat des clients et augmenter les ventes en proposant des produits pertinents et adaptés à leurs besoins.

Le machine learning permet de réaliser cette recommandation de manière plus efficace en analysant de grandes quantités de données comportementales des clients pour détecter des modèles et des similarités. Les algorithmes machine learning peuvent ainsi identifier des produits complémentaires ou similaires à ceux achetés par le client, en se basant sur des données telles que les historiques d'achat, les évaluations de produits et les commentaires des clients.

Il existe plusieurs types d'algorithmes de recommandation de produits, mais les plus couramment utilisés sont :

La recommandation basée sur la similarité : cette méthode recommande des produits similaires à ceux déjà achetés ou consultés par le client. Elle se base sur des données telles que les historiques de navigation et les achats précédents pour détecter des similarités entre les produits.

La recommandation basée sur la popularité : cette méthode recommande les produits les plus populaires ou les mieux notés parmi l'ensemble des produits proposés. Elle est souvent utilisée pour les nouveaux clients ou ceux qui n'ont pas encore effectué d'achat sur le site.

La recommandation basée sur le filtrage collaboratif : cette méthode recommande des produits en se basant sur les comportements d'achat de clients similaires. Elle se base sur des données telles que les historiques d'achat et les évaluations de produits pour identifier des groupes de clients ayant des comportements similaires.

Par exemple, un site de vente en ligne peut utiliser le machine learning pour recommander des produits à ses clients en fonction de leurs historiques d'achat et de navigation. Si un client a acheté une paire de chaussures de sport, le système de recommandation peut lui suggérer des chaussettes de sport ou des vêtements de sport similaires. Si un client a consulté plusieurs produits de beauté, le système de recommandation peut lui suggérer des produits similaires de la même marque ou de marques similaires. Voici quelques études qui ont utilisé l'apprentissage automatique pour créer des systèmes de recommandation :

Kuo et Cheng (2022) ont proposé un système de recommandation personnalisé basé sur le contenu qui intègre l'architecture du système de recommandation traditionnel basé sur le contenu existant pour le commerce électronique avec l'ajout d'un ajusteur de retour d'information. Les résultats ont montré que le système proposé, qui est basé sur une approche plus objective, peut déterminer les préférences des clients sur la base de leur comportement d'achat répété, évitant ainsi les jugements subjectifs.

Huang (2022) a effectué une analyse comparative de trois algorithmes de recommandation populaires sur un ensemble de données du site web d'Alibaba. Les résultats expérimentaux ont révélé que l'algorithme de recommandation par réseau neuronal convolutif basé sur l'apprentissage profond était plus performant que les autres algorithmes.

Dans une autre étude, Chehal et al. (2022) ont proposé une approche visant à améliorer les recommandations en tirant parti de la fonction d'évaluation des produits. Cette méthode consiste à extraire les caractéristiques des produits à partir des commentaires fournis par les utilisateurs et à ne pas recommander des produits présentant des caractéristiques non pertinentes afin d'améliorer la liste des recommandations.

Al-Ghamdi et al. (2021) ont réalisé une étude comparative de divers algorithmes de filtrage collaboratif afin d'évaluer leurs performances, à savoir : K-Nearest Neighbor, Slope One, co-clustering et Non-negative Matrix Factorization. Les résultats ont montré que l'algorithme du plus proche voisin (K-Nearest Neighbor) pour le filtrage collaboratif basé sur les éléments était plus performant que les autres modèles.

Xuecong et al. (2021) ont mis en œuvre une solution pour construire un système de recommandation pour le commerce électronique sur la plateforme d'informatique en nuage afin de gérer plus efficacement les informations massives sur les marchandises et les informations sur les utilisateurs. La solution proposée leur permet d'accroître la capacité d'exploration des données massives et d'analyse de la veille stratégique et d'obtenir un calcul de haute performance à un coût réduit.

Wang et al. (2021) ont proposé une approche de collecte et de prétraitement des données en temps réel provenant de multiples plateformes de commerce électronique, qui génèrent des données sur les utilisateurs, et de collecte de toutes les données personnalisées sur les utilisateurs pour préparer l'exploration de données suivante, puis d'utilisation de la technologie d'exploration de données dans les big data pour recommander automatiquement des produits personnalisés afin de satisfaire les besoins et les goûts personnalisés des clients.

Afin d'identifier les préférences des clients, Osmond et al. (2021) ont proposé une méthode d'analyse conjointe améliorée. Pour évaluer leur méthode, ils l'ont comparée à d'autres algorithmes de prédiction à savoir : le modèle linéaire généralisé, l'arbre de décision, la forêt aléatoire, les arbres boostés par le gradient et la machine à vecteur de support. Les résultats expérimentaux ont montré que ce réseau d'analyse conjointe est performant et fonctionne bien pour choisir un produit parmi plusieurs attributs afin d'obtenir le meilleur prix, malgré ses limites.

Une recommandation plus précise permet à l'entreprise d'améliorer l'expérience de ses clients en leur proposant des produits pertinents et adaptés à leurs besoins, augmentant ainsi la satisfaction et la fidélité des clients.

1.2. Prix

« Le prix » est un autre élément du mix marketing des 7P qui avait attiré l'attention des chercheurs.

Wolkenfelt et Situmeang (2020) avaient examiné l'effet des structures de tarification des applications sur l'évaluation des produits en ce qui concerne l'assertivité et les sentiments exprimés dans les critiques. Ils avaient constaté que le modèle de tarification affectait le sentiment des consommateurs dans les critiques et que la suppression des obligations de paiement initial avait un impact positif sur le sentiment des consommateurs.

Casamatta et al. (2022) avaient étudié les différences de prix mois par mois entre les hôtes professionnels et non professionnels sur Airbnb (une plateforme de location de logements en ligne qui permet aux particuliers de louer leurs propriétés à court terme). Ils avaient constaté que l'écart de prix existait pendant la saison de pointe et que la plateforme augmenterait son profit si la part des hôtes professionnels était suffisamment grande.

Certains chercheurs s'étaient concentrés sur la prévision des prix sur le marché de l'électricité. Sueyoshi et Tadiparthi (2005), Sueyoshi et Tadiparthi (2008) s'étaient concentrés sur le marché américain de gros de l'électricité pour examiner comment les changements de prix se produisaient dans différents environnements économiques (Sueyoshi et Tadiparthi, 2005) et comprendre les changements de prix dynamiques pour une prise de décision efficace (Sueyoshi et Tadiparthi, 2008).

Yang et al. (2022) avaient étudié la prévision des prix de l'électricité. Ils avaient développé un nouveau modèle de prévision offrant les avantages d'un prétraitement de données adaptatif, d'une méthode d'optimisation avancée, d'un modèle à noyau et d'une stratégie optimale de sélection de modèles, ce qui avait permis une prévision des prix de l'électricité plus efficace.

1.2.1. Tarification dynamique

La tarification dynamique est une stratégie de tarification basée sur la variation des prix en fonction des tendances du marché et de la demande. Cette stratégie permet aux entreprises de maximiser leur revenu en ajustant les prix en temps réel en réponse aux fluctuations de la demande.

Le machine learning peut être utilisé pour aider les entreprises à mettre en œuvre des stratégies de tarification dynamique en analysant les données de marché et de demande. Les algorithmes machine learning peuvent être entraînés à partir des données historiques des ventes et des prix pour identifier les tendances et les modèles dans la demande des consommateurs.

En utilisant ces données, les algorithmes peuvent prédire la demande future et aider les entreprises à déterminer les prix optimaux pour leurs produits en temps réel. Les entreprises peuvent ajuster les prix en fonction de la demande et de l'offre, en utilisant des techniques d'optimisation pour maximiser les profits (Loukili et al., 2023).

Les algorithmes machine learning peuvent également prendre en compte les variables qui influencent la demande, telles que les promotions, les événements saisonniers, les préférences des consommateurs et les tendances du marché. En utilisant ces variables, les entreprises peuvent personnaliser leur tarification en fonction des segments de clients et des moments de la journée.

La tarification dynamique est une application clé du machine learning dans le marketing digital, car elle permet aux entreprises de maximiser leur revenu en ajustant les prix en temps réel en fonction des tendances du marché et de la demande. Les algorithmes machine learning peuvent aider les entreprises à prédire la demande future, à personnaliser leur tarification en fonction des segments de clients et des moments de la journée, et à optimiser les profits en temps réel.

1.2.2. Segmentation des prix

Un autre aspect qui peut être intégré à l'élément "Prix" du marketing 7P dans le contexte du marketing digital est la segmentation des prix. La segmentation des prix implique de proposer différents prix pour différents segments de clients en fonction de leurs besoins, de leurs comportements d'achat et de leur capacité à payer.

Les entreprises peuvent utiliser des données clients pour segmenter leurs clients et proposer des prix différents à chaque segment. Par exemple, une entreprise de voyage en ligne peut proposer des prix différents pour les billets d'avion en fonction de la destination, de la date de voyage, du moment de l'achat et de l'historique d'achat du client.

La segmentation des prix peut aider les entreprises à maximiser leurs bénéfices en proposant des prix différents pour différents segments de clients. Cela peut également améliorer l'expérience client en proposant des prix plus adaptés aux besoins et aux capacités de chaque client.

Cependant, comme pour la tarification dynamique, il est important que les entreprises utilisent la segmentation des prix de manière responsable et éthique pour éviter de tromper ou de maltraiter les clients. Les entreprises doivent être transparentes quant à leur utilisation de la segmentation des prix et doivent communiquer clairement les raisons des prix différents proposés à chaque segment de clients.

1.3. Place

« Place » fait référence aux canaux de marketing dans le mix marketing des 7P. Les chercheurs ont également remarqué l'influence de la place sur les activités de marketing.

Danaher et al. (2020) ont enquêté sur les achats en magasin et en ligne de trois marques de vente au détail dans la catégorie de vêtements et ont révélé l'existence d'un groupe important de clients qui achètent à travers plusieurs marques et canaux d'achat. Ils ont montré que trois formes de dépenses publicitaires (e-mail, catalogues et recherche payante) peuvent influencer les ventes en magasin et en ligne.

Bhowmick et Mitra (2019) se sont concentrés sur la proximité géographique pour prédire la popularité des entreprises dans les réseaux sociaux basés sur la localisation. Ils ont défini deux types de consommateurs visitant une entreprise, locale et étrangère, et ont constaté que les clients locaux jouent un rôle majeur dans la formation de la popularité des entreprises dans les villes locales.

De plus, Leguina et al. (2020) ont révélé que le modèle d'attribution est un instrument utilisé pour évaluer le retour sur investissement de différents canaux et explorer l'impact des principales propriétés des parcours utilisateur sur les modèles d'attribution.

1.3.1. Optimisation de la navigation sur les sites web

L'optimisation de la navigation sur les sites web est un aspect clé du marketing en ligne, car une navigation fluide et intuitive peut aider les utilisateurs à trouver facilement les informations qu'ils recherchent et à effectuer des actions telles que l'achat de produits ou la demande de

services. Le machine learning peut être utilisé pour optimiser la navigation sur le site web en offrant des suggestions de produits, en améliorant la recherche et en offrant une navigation plus intuitive.

Les algorithmes machine learning peuvent être utilisés pour analyser les données comportementales des utilisateurs, telles que l'historique des achats, les pages visitées et les interactions avec le site web. Ces données peuvent être utilisées pour comprendre les intérêts et les préférences des utilisateurs, ainsi que pour offrir des suggestions de produits en fonction de leur comportement et de leurs préférences.

Le machine learning peut également être utilisé pour améliorer les résultats de recherche en offrant des suggestions de recherche plus précises et pertinentes en fonction des préférences de l'utilisateur. L'optimisation de la recherche peut également inclure l'analyse des requêtes de recherche les plus courantes pour aider les entreprises à mieux comprendre les besoins des utilisateurs.

Enfin, le machine learning peut être utilisé pour offrir une navigation plus intuitive en analysant les parcours des utilisateurs sur le site web et en offrant des suggestions de navigation plus claires et plus simples.

Les avantages de l'optimisation de la navigation sur le site web comprennent une augmentation de la satisfaction utilisateur, une augmentation de la fidélisation utilisateur et une augmentation des conversions.

En conclusion, le machine learning peut être utilisé pour optimiser la navigation sur le site web en offrant des suggestions de produits, en améliorant la recherche et en offrant une navigation plus intuitive. L'optimisation de la navigation sur le site web peut aider les entreprises à augmenter la satisfaction utilisateur, la fidélisation utilisateur et les conversions.

1.3.2. Optimisation du référencement

L'optimisation du référencement est un domaine clé du e-marketing qui utilise des techniques pour améliorer la visibilité et le classement des sites web dans les résultats des moteurs de recherche. Le référencement est important pour les entreprises qui souhaitent augmenter leur visibilité en ligne et attirer plus de trafic sur leur site web.

Le machine learning peut être utilisé pour optimiser le référencement en identifiant les mots-clés et les stratégies de contenu qui sont les plus efficaces pour améliorer le classement dans les moteurs de recherche. Les algorithmes machine learning peuvent analyser les données de recherche pour identifier les tendances, les modèles de comportement des utilisateurs, les intentions de recherche et les préférences des utilisateurs. Ces informations peuvent être utilisées pour adapter les stratégies de référencement et de contenu pour mieux répondre aux besoins des utilisateurs et améliorer le classement dans les moteurs de recherche.

En utilisant des techniques du machine learning, les spécialistes du marketing peuvent également identifier les pages de leur site web qui ont le plus de potentiel pour le référencement et concentrer leurs efforts sur ces pages pour maximiser leur impact. Ils peuvent également identifier les erreurs techniques sur leur site web qui peuvent nuire à leur référencement et les corriger pour améliorer leur visibilité en ligne.

L'optimisation du référencement est un domaine clé du e-marketing et l'utilisation de techniques machine learning peut aider les entreprises à améliorer leur visibilité en ligne, attirer plus de trafic sur leur site web et atteindre leur public cible.

1.4. Promotion

La « promotion » est considérée comme un élément clé du marketing digital et peut être effectuée via divers canaux tels que les réseaux sociaux, les publicités en ligne, les campagnes de marketing par e-mail, les contenus sponsorisés, etc (Key, 2017).

1.4.1. Gestion de publicité

Ces derniers temps, la publicité en ligne s'est rapidement développée pour la promotion. Les chercheurs cherchent à savoir comment améliorer l'efficacité de la publicité et comment prédire les taux de clics, qui sont des composantes critiques de toute plateforme de publicité en ligne (Bai et al., 2019).

Certains chercheurs ont examiné l'impact des attributs de publicité observables tels que Goodrich et al. (2015), Schwartz et al. (2017), Lee et al. (2018), Peng et al. (2020), ainsi que l'utilisation antérieure de la marque (Simmonds et al., 2020). Ils ont également découvert les intérêts personnels immédiats des blogueurs (Fan et Chang, 2010, Fan et Chang, 2011), optimisé la technique de microciblage en réponse directe (Miralles-Pechuán et al., 2018) et analysé les

effets des appels publicitaires (Couwenberg et al., 2017, Matz et al., 2019), du développement d'histoires (Quesenberry et Coolsen, 2019) et des informations visuelles pour améliorer l'efficacité de la publicité.

D'autres chercheurs ont appliqué différentes méthodes d'apprentissage automatique pour la prédiction des clics à travers la publicité en ligne, telles que Chapelle et al. (2015), King et al. (2015), Xu et al. (2018), Bai et al. (2019) et Yang et al. (2020).

La publicité frauduleuse a également reçu une attention significative. Les chercheurs ont utilisé les technologies d'apprentissage automatique pour détecter les publicités frauduleuses en tirant parti de la flexibilité de la tolérance envers les différents serveurs de publicité (Haider et al., 2018), de nouveaux graphes hétérogènes pondérés (Hu et al., 2020), de fonctionnalités robustes contre l'évasion des attaquants (Kanei et al., 2020), d'insights de l'analyse des médias sociaux et de l'informatique bio-inspirée (Aswani et al., 2018), d'apprentissage basé sur l'attention des données des médias personnels (Wang et al., 2020, Hou et al., 2021), et de fonctionnalités sémantiques dans les articles de marketing de contenu (Liang et al., 2019).

1.4.2. Chatbots

Les chatbots galvanisés par l'IA et le ML sont l'une des dernières innovations dans le domaine du e-marketing et des affaires (Luo et al., 2019). Les chatbots sont définis comme "un programme informatique, qui simule le langage humain à l'aide d'un système de dialogue basé sur le texte (Zumstein et Hundertmark, 2018)." Ils sont conçus pour imiter une conversation interpersonnelle, caractérisée par un haut degré de personnalisation (Letheren & Glavas, 2017). Ils peuvent être utilisés pour une variété de tâches, notamment la réponse aux questions des clients, l'assistance technique, la prise de commandes, la réservation de rendez-vous et la collecte de données.

Le machine learning peut être utilisé pour améliorer les performances des chatbots intelligents en permettant aux algorithmes de s'adapter et d'apprendre de nouvelles informations à partir des interactions utilisateur. Les chatbots intelligents basés sur l'apprentissage automatique peuvent s'améliorer continuellement à mesure qu'ils interagissent avec les utilisateurs et collectent des données, ce qui leur permet de fournir des réponses plus précises et plus rapides aux requêtes des utilisateurs.

Les chatbots peuvent également être intégrés à des systèmes d'intelligence artificielle plus larges pour automatiser des processus commerciaux, tels que la collecte et l'analyse de données, l'identification des tendances du marché et la prise de décision.

Les avantages des chatbots comprennent une disponibilité 24 heures sur 24, une réponse rapide aux requêtes des clients et une réduction des coûts de personnel en automatisant les processus de support client. Les chatbots intelligents peuvent également améliorer l'expérience client en offrant des réponses rapides et précises aux questions des clients, en évitant les temps d'attente et en offrant une assistance personnalisée.

Androutsopoulos et al. (1995) ont présenté différentes méthodes d'inférence en langage naturel dans les bases de données. Ils ont également expliqué les avantages et les inconvénients de ces méthodes. Leur système n'utilisait pas d'intelligence artificielle.

Dans son étude, Joachims (1998) a tenté d'explorer les machines à vecteurs de support (SVM) pour catégoriser les textes. Il a expliqué que les SVM sont plus performants pour la catégorisation des textes en raison de l'espace d'entrée de haute dimension, de moins de caractéristiques non pertinentes, de vecteurs de documents clairsemés et de la capacité à séparer linéairement les problèmes de catégorisation des textes. Kassabgi (2017) a montré une approche algorithmique très simple pour classer les textes en utilisant des bayes naïves multinomiales. Il a également expliqué une approche simple d'apprentissage automatique utilisant un réseau neuronal pour classer les textes.

Tensorflow est un excellent outil d'apprentissage automatique développé par Google. Polosukhin (2016) a expliqué un modèle de classification de texte en utilisant Tensorflow. Mishu et Rafiuddin (2016) ont comparé les performances de Naive Bayes, Multinomial Naive Bayes, Bernoulli Naive Bayes, Logistic Regression, Stochastic Gradient Descent, SVC, Linear SVC et Back Propagation Network.

Den Broeck et al. (2019) ont étudié l'efficacité de la publicité des chatbots en examinant si et comment l'utilité perçue et l'utilité d'un chatbot consulté sur Facebook affectaient le caractère intrusif perçu de la publicité initiée par le chatbot dans une étape ultérieure. Zarouali et al. (2018) ont testé les déterminants cognitifs et affectifs qui influencent l'efficacité des chatbots à travers un modèle CAT adapté sur Facebook pour les marques. Luo et al. (2019) ont constaté que la divulgation des chatbots avant une conversation machine-client réduisait le taux d'achat, bien

que les chatbots non divulgués soient aussi efficaces que les travailleurs compétents pour engendrer l'achat de clients. Wu et al. (2020) ont proposé une architecture de système d'agent conversationnel intelligent basé sur les connaissances pour soutenir les services à la clientèle dans les ventes et le marketing du commerce électronique. Le système a été construit dans un environnement réel et a intégré diverses technologies, y compris l'exploration du Web, le traitement du langage naturel, les bases de connaissances et l'IA.

Les chatbots intelligents sont un outil crucial du e-marketing qui devient de plus en plus populaire, et l'utilisation de techniques machine learning peut aider les entreprises à améliorer la précision et l'efficacité des chatbots, à automatiser des processus commerciaux et à améliorer l'expérience client.

1.4.3. Prédiction de la demande

La prédiction de la demande peut aider les entreprises à prendre des décisions éclairées concernant leur production, leur approvisionnement et leur marketing. Les algorithmes de ML permettent de prédire les ventes futures en analysant les relations entre les données historiques des ventes et les variables de marché telles que les dépenses publicitaires, les promotions et les événements saisonniers.

Les organisations sont continuellement à la recherche de meilleurs modèles et techniques d'exploration des données pour rester compétitives et augmenter leurs revenus (Cheriyan et al., 2018). Il existe une vaste littérature sur les prévisions de ventes utilisant des algorithmes d'apprentissage automatique. Cette partie passe en revue les différentes études qui ont été menées dans le domaine de la prévision des ventes.

Certains des modèles qui ont été mis en œuvre pour la prévision des ventes comprennent des modèles statistiques tels que la régression, la moyenne mobile intégrée autorégressive (ARIMA), la moyenne mobile autorégressive (ARMA). Néanmoins, la prédiction des ventes est une question compliquée et dépend de nombreux facteurs internes et externes, c'est pourquoi l'approche statistique a ses limites (Hill et al., 1994).

Krishna et al. (2018) se sont concentrés sur la prévision des ventes au niveau du produit. Une variété de techniques de prévision telles que la régression multiple, la régression polynomiale et la régression Ridge, ainsi qu'une variété d'algorithmes d'amplification tels que

AdaBoost et l'amplification de l'arbre de gradient ont été considérées pour développer des approches appropriées afin d'atteindre une précision maximale.

Sun et al. (2008) ont appliqué l'algorithme Extreme Learning Machine, qui se réfère à un type de réseau neuronal, pour établir la relation entre le volume des ventes et les facteurs affectant la demande en utilisant un ensemble de données en temps réel et ont révélé que leur modèle était plus satisfaisant que tout autre modèle de prédiction des ventes utilisant des réseaux neuronaux à rétropropagation.

Dans une étude comparative, Arif et al. (2019) ont discuté d'une nouvelle méthode qui utilise l'apprentissage automatique pour permettre des prédictions précises. La méthode proposée recueille et analyse les données historiques de l'entrepôt et ap-plique divers algorithmes tels que k-Nearest Neighbor (k-NN), Gaussian Naïve Bayes et Decision Tree Classifier pour détecter la meilleure tech-nique de prédiction de la demande. L'analyse des modèles adoptés a révélé que Gaussian Na-ïve Bayes avait la plus grande précision pour les données antérieures.

Dans une autre étude, (Cheriyan et al., 2018) ont appliqué différentes techniques d'apprentissage automatique et d'exploration de données dans le but de sélectionner l'approche prédictive la plus performante pour la prévision des ventes avec une plus grande précision. Ils ont toutefois indiqué que les modèles de prévision traditionnels sont extrêmement complexes lorsqu'ils traitent des big data et de la précision des prévisions de ventes, et que pour surmonter ces problèmes, diverses techniques d'exploration de données peuvent être utilisées.

Chu et Zhang (2003) ont comparé les modèles non linéaires aux modèles linéaires pour la prévision des ventes et ont constaté que les réseaux neuronaux donnent de bons résultats avec des séries chronologiques corrigées des variations saisonnières, tandis que les modèles de régression donnent de bons résultats avec des variables saisonnières fictives.

Dans une autre étude réalisée par Liu et al. (2013), des techniques statistiques ont été utilisées pour prédire les ventes, à savoir le lissage exponentiel et l'analyse bayésienne. De même, la prédiction des ventes a été effectuée à l'aide de réseaux neuronaux artificiels (ANN) et de réseaux neuronaux évolutifs (ENN).

Thomassey et al. (2005) ont proposé une approche de prévision hybride combinant la logique floue, les procédures évolutives et les réseaux neuronaux.

Singh et al, (2017) ont adopté la dimension big data pour prédire les ventes de Walmart. Ils ont utilisé différents outils et techniques, notamment les API Python et Scala du cadre Spark.

Harsoor et Patil (2015) ont réalisé des prédictions de ventes basées sur des applications Big Data à savoir MapReduce et Hadoop. Catal et al. (2019) ont proposé une prédiction des ventes en utilisant des techniques de régression telles que la régression linéaire, la régression bayésienne, le réseau neuronal, la forêt décisionnelle et l'arbre décisionnel boosté. Ainsi que des méthodes d'analyse de séries temporelles telles que : Méthode de la moyenne, ARIMA saisonnier et non saisonnier, méthode ETS (erreur, tendance, saisonnier) saisonnière et non saisonnière, méthode de la dérive et méthode naïve, sur les données historiques des ventes de Walmart. La régression par arbre de décision boostée s'est avérée la plus performante.

Ensafi et al. (2022) ont utilisé des modèles de prévision de séries temporelles, notamment la méthode des moyennes mobiles intégrées autorégressives saisonnières (SARIMA), le triple lissage exponentiel, le réseau neuronal convolutif (CNN) et la méthode LSTM. En vue de prévoir les ventes d'un magasin de meubles au détail à partir de données historiques. En conséquence, la méthode LSTM a obtenu de meilleurs résultats que les autres méthodes. En outre, une bonne performance des modèles Prophet et CNN a été observée.

Behera et al. (2022) ont utilisé des techniques de régression, à savoir Random Forest, Ridge Regression, Linear Regression, Lasso Regression, Decision Tree, Supplementary Tree et XGBoost, pour prévoir les futures ventes hebdomadaires d'un magasin. Sur la base des mesures de performance MAE, MAE pondéré (WMAE) et RMSE. Les résultats ont montré que le modèle de forêt aléatoire était plus performant que les autres modèles.

Schmidt et al. (2022) ont mené une étude comparative utilisant plusieurs modèles d'apprentissage automatique et des modèles de réseaux neuronaux récurrents (RNN). Ils ont utilisé les données de vente réelles d'un restaurant pour tester les effets de la tendance et de la saisonnalité. Cependant, ils ont constaté que les meilleurs résultats pour la prévision à un jour donné provenaient des modèles linéaires. Les modèles RNN LSTM et Temporal Fu-sion Transformer (TFT), ainsi que les modèles d'ensemble, ont également obtenu de bons résultats. Pour les prédictions de ventes à une semaine, les modèles non-RNN ont sous-performé et les modèles RNN ont eu plus de succès. Les modèles RNN ont sous-performé sur les ensembles de données sans tendances ni saisonnalité, alors que les modèles d'apprentissage automatique

simples ont obtenu des résultats significatifs en séparant linéairement chaque instance d'apprentissage.

Les prévisions de vente basées sur l'apprentissage automatique peuvent également être utilisées pour optimiser la gestion des stocks et réduire les coûts associés à la production et à l'approvisionnement. En analysant les tendances de la demande, les entreprises peuvent ajuster leur production et leur approvisionnement pour éviter les surstocks ou les ruptures de stock, ce qui peut entraîner des coûts supplémentaires. En outre, la prédiction des ventes peut être utilisée pour adapter les stratégies de marketing et de communication. Les entreprises peuvent utiliser les prévisions pour identifier les produits les plus populaires et cibler les clients potentiels, ce qui peut augmenter les ventes et la rentabilité.

La prédiction des ventes est un domaine clé de l'application du machine learning dans le marketing digital, car elle peut aider les entreprises à prendre des décisions plus efficaces, et à améliorer leur rentabilité en prévoyant les ventes futures de manière plus précise et en adaptant leurs stratégies de marketing et de communication en conséquence.

1.5. Personnes

« Personnes » désignent les clients qui achètent des produits ou des services et les autres clients dans l'environnement des services (Rafiq et Ahmed, 1995). La littérature existante applique les technologies de ML pour analyser les activités de marketing liées aux personnes, y compris la prédiction du désabonnement, la prédiction du ciblage des clients et l'engagement. Ces applications sont présentées ci-dessous.

1.5.1. Prédiction du désabonnement des clients

La gestion du désabonnement est une priorité absolue pour la plupart des entreprises, et la prédiction du désabonnement joue un rôle central dans les programmes de gestion du désabonnement (Ascarza, 2018). La prédiction du churn des clients a suscité l'intérêt des chercheurs.

Huang et al. (2015) ont mis en évidence les 3V qui sont le volume, la variété et la vitesse. La prédiction peut être réalisée efficacement en utilisant une grande quantité de données d'entraînement à partir d'un grand nombre de caractéristiques et en augmentant la vélocité du traitement des données. Le résultat de cette méthode a atteint un taux de précision de 0,96.

Kumar et al. (2017) ont comparé trois algorithmes du machine learning pour la prédiction du churn à savoir la régression logistique, la forêt aléatoire et la forêt aléatoire équilibrée. Ils ont utilisé l'ACP pour la sélection des caractéristiques. En conséquence, il a été constaté que le modèle de régression logistique avait le meilleur score AUC de 0,86 par rapport aux autres modèles.

Zhang et al. (2018) ont mentionné que l'analyse de l'attrition des clients est très utile pour les entreprises de télécommunications afin de conserver les clients qui pèsent lourd. Le modèle de prédiction de désabonnement des clients avec plus de précision est fortement pondéré pour la décision de rétention des clients. Dans cette étude, le modèle SVM est employé parce qu'il est meilleur pour la précision. Il peut résoudre les échantillons dans un espace de faible dimension qui est linéaire dans un espace bidimensionnel. Cependant, la limite du modèle proposé est qu'il est très complexe de mesurer les clients désabonnés.

Bharat (2019) a adopté un modèle qui se concentre sur le schéma d'activité du client. Il est construit sur l'activité du client en recherchant la durée moyenne du temps d'inactivité et les fréquences d'inactivité. Cette méthode peut également être utilisée dans d'autres domaines pour la prédiction du désabonnement des clients.

Andrews et al. (2019) ont utilisé un ensemble de données de 10 000 enregistrements de clients d'une entreprise de télécommunications. Ils ont utilisé des modèles d'apprentissage profond et des méthodes d'approbation croisée à 10 chevauchements pour évaluer la précision de la prédiction. Les résultats ont permis d'obtenir un score AUC de 0,89.

Ullah et al. (2019) se sont concentrés sur le secteur des télécommunications et ont proposé un modèle de prédiction de résiliation pour identifier les clients susceptibles de résilier et explorer les facteurs conduisant à la résiliation des clients.

Calzada et al. (2020) ont mené une étude comparative de deux méthodes de prédiction du churn, à savoir Time-Order-Graph et Aggregated-Static-Graph, et d'un classificateur forestier basé sur trois métriques de seuil pour évaluer la qualité de la prédiction du classificateur forestier de similarité par rapport à chaque métrique de centralité.

Schaeffer et Sanchez (2020) ont proposé et testé une méthode de prédiction de la fidélité des clients en utilisant les enregistrements mensuels des transactions des clients dans un

contexte interentreprises. Ils ont constaté que les techniques actuelles d'apprentissage automatique peuvent prédire de manière adéquate quels clients seront perdus.

Wang et al. (2020) ont présenté une analyse comparative des approches de classification populaires utilisées pour surmonter la perte de clientèle dans le secteur des télécommunications. Sur la base des performances des modèles, ils ont sélectionné les modèles les plus performants, à savoir Light GBM, XGBoost, Random Forest et Decision Tree. Ils ont ensuite développé un modèle d'ensemble à l'aide de la technique de vote souple, qui a obtenu le score AUC le plus élevé.

Nam (2021) a également comparé deux méthodes de prédiction, à savoir SMOTE et Deep Belief Network (DBN), avec des approches d'apprentissage sensibles aux coûts, telles que Focal Loss et Weighted Loss. Cette étude a révélé que les performances globales de Focal Loss et de Weighted Loss étaient supérieures à celles de SMOTE et de DBN.

Siddika et al. (2021) ont utilisé des techniques machine learning et de deep learning pour déterminer les clients qui abandonnent et les facteurs saillants qui les poussent à abandonner. Le modèle Random Forest a surpassé les autres algorithmes, suivi par les modèles d'apprentissage profond CNN et MLP. Les résultats de l'expérience ont mis en évidence les modèles de prédiction qui identifient les churners probables avec une précision optimale et les principaux facteurs qui influencent le décrochage des clients.

Zadoo et al. (2022) ont réalisé un examen du taux de désabonnement des clients dans le secteur des télécommunications, en utilisant une combinaison de techniques d'apprentissage supervisé : pour la prédiction de l'attrition des clients, et des techniques d'apprentissage non supervisé de regroupement : pour la segmentation de la clientèle.

Vakeel et al. (2022) ont proposé un modèle d'apprentissage automatique qui comprend des algorithmes de stimulation pour identifier le taux de désabonnement et une technique de mélange gaussien pour regrouper les clients désabonnés. Les résultats obtenus avec le modèle de mélange gaussien adopté pour le regroupement des clients ont donné un score de silhouette de 0,36.

1.5.2. Ciblage des clients

Les messages marketing sont plus efficaces s'ils atteignent les bons clients (Chaniago et Ariyani, 2023). L'allocation de différentes ressources marketing à différents consommateurs est importante pour les entreprises. Plusieurs études ont prédit les clients de ciblage et appliqué le marketing direct.

Les chercheurs prédisent le ciblage des clients du point de vue de la réponse des clients (Lessmann et al., 2021), du sexe (Ameer et al., 2019, Hsieh et al., 2015, Jeon et al, 2020), la préférence (Abernethy et al., 2008, Luaces et al., 2015, Fiore et al., 2017), l'hétérogénéité (Chen et al., 2017, Dew et al., 2020) et la loyauté (Buckinx et al., 2007).

En outre, Chen et al. (2020) ont utilisé une expérience de terrain randomisée pour explorer l'efficacité du marketing de proximité direct auprès des patients grâce à la reconnaissance faciale. L'objectif était de mettre en œuvre des programmes de marketing personnalisés dans le domaine des soins de santé.

La reconnaissance faciale est une technologie importante pour le marketing ciblé et a suscité l'intérêt des chercheurs (Nikolaev et al., 2023). Les caractéristiques telles que la couleur de la peau, le sexe, l'apparence, l'émotion et les marques faciales peuvent être extraites des images faciales à l'aide de différentes technologies de ML (Ramya et al., 2022). Les chercheurs ont donc exploré plusieurs applications de la reconnaissance faciale pour le marketing ciblé interactif.

Su et al. (2020) ont appliqué l'analyse de l'expression faciale pour développer un modèle de recommandation de vêtements personnalisés. Pour cela, ils ont combiné l'intensité et la durée de l'expression pour améliorer le rappel de la recommandation.

Yolcu et al. (2020) ont présenté un système basé sur l'apprentissage profond pour surveiller les intérêts des clients. Ils ont utilisé des caractéristiques locales basées sur les parties avec des informations faciales holistiques pour une reconnaissance robuste des expressions faciales.

Masui et al. (2020) ont analysé les expressions faciales et les réponses physiologiques à travers des vidéos faciales pour évaluer l'efficacité de la publicité dans l'environnement naturel et l'intention d'achat.

Álvarez-Pato et al. (2020) ont présenté un nouveau système d'analyse sensorielle, incluant la reconnaissance des émotions faciales, pour prédire l'acceptation par les consommateurs d'échantillons alimentaires.

Renigier-Biłozor et al. (2021) ont proposé une méthode qui corrèle les états émotionnels des acheteurs avec la visualisation des attributs sélectionnés des propriétés sur le marché immobilier. Ils ont vérifié le classement d'importance déterminé par les émotions faciales inconscientes.

Ab Wahab et al. (2021) ont proposé un modèle hybride CNN et KNN pour la reconnaissance des expressions faciales afin de reconnaître les émotions des utilisateurs. La capacité de la technologie de reconnaissance faciale à déterminer les émotions des utilisateurs et à influencer les décisions d'achat fait partie d'une stratégie de marketing numérique importante pour le marketing. La reconnaissance faciale continue donc d'être un domaine de recherche prometteur pour le marketing personnalisé.

1.5.3. Engagement

Les plateformes de médias sociaux, telles que Facebook, Twitter et Instagram, ont créé de nouveaux modes d'interaction, de communication et d'engagement (Arora et al., 2019). Les gens affichent leurs opinions et utilisent les termes "aimer", "partager" ou "commenter" pour échanger des informations (Hai Ming et al., 2022) par l'intermédiaire de ces plateformes. Le niveau d'engagement permet d'évaluer l'efficacité des efforts de marketing des entreprises dans les médias sociaux (Yost et al., 2021). Les énormes volumes d'opinions, exprimés en temps réel, présentent un grand intérêt en tant qu'application marketing novatrice (Lies, 2019). C'est pourquoi des études approfondies ont été menées sur la ML dans l'engagement et l'interaction.

Comme les clients préfèrent publier des avis pour partager leurs expériences après avoir consommé, les chercheurs examinent les avis des clients dans différents domaines et sous différentes perspectives. Ils se concentrent sur les critiques de films (Cheng et Huang, 2020), les voyageurs (Kwok et al., 2020), les détaillants de mode rapide (Pantano et al., 2019) et les produits de santé (Chatterjee et al., 2021), ainsi que sur l'utilité des critiques en ligne (Srivastava et Kalro, 2019). D'autres chercheurs s'intéressent aux "lurkers", qui ne publient pas d'avis après l'achat (Li et Zhang, 2020), et aux avis "spoilers" pour trouver de nouvelles idées (Ryoo et al., 2020).

Toutefois, à mesure que le commerce électronique se développe, la prévalence des faux avis en ligne augmente également (Wu et al., 2020). Les faux avis sont utilisés pour manipuler la réputation des produits (Yu et al., 2019) et les décisions d'achat des consommateurs. Plusieurs chercheurs tentent de détecter les faux avis en appliquant des méthodes de ML (Zhang et al., 2016, Khurshid et al., 2019), telles qu'un nouveau modèle de relation individu-groupe-marchand (Yu et al., 2019), des n-grammes de caractères (Cagnina et Rosso, 2017), l'algorithme d'échantillonnage de Gibbs (Dong et al., 2018) et des caractéristiques basées sur les statistiques (Hazim et al., 2018) pour la détection.

D'autres chercheurs explorent activement le marketing d'influence sur les médias sociaux. Ils mesurent l'indice d'influence des médias sociaux sur les plateformes de médias sociaux populaires, prédisent la prédiction des meilleurs persuadeurs (Fang et Hu, 2018), explorent le rôle des récits (Feng et al., 2020) et étudient l'influence des superstars sur les spectateurs (Montañés et al., 2014).

1.6. Processus

« Processus » fait référence aux procédures, aux mécanismes et au déroulement des activités. Les processus en ligne, tels que le processus de commande, de paiement et de livraison, sont également essentiels pour garantir une expérience utilisateur satisfaisante. Les chercheurs dans ce contexte s'intéressent à la manière dont les gestionnaires procèdent à la segmentation marketing.

1.6.1. Segmentation du marché

La segmentation de marché est une technique largement utilisée en marketing pour diviser un marché en différents sous-groupes de consommateurs ayant des besoins, des comportements et des attentes similaires (Tsiotsou et Ratten, 2010, Bruwer et al., 2017). La segmentation de la clientèle consiste à regrouper les clients en segments qui ne se chevauchent pas, de sorte que les clients d'un même segment ont des besoins et des préférences similaires (Blumberg et Thompson, 2022). Cette méthode permet aux entreprises de mieux comprendre leur audience cible et de développer des stratégies de marketing plus efficaces en fonction des besoins spécifiques de chaque segment.

Le machine learning permet de réaliser cette segmentation de manière plus précise et plus efficace, en analysant de grandes quantités de données comportementales et démographiques

des consommateurs pour détecter des modèles et des similarités (Alkhayrat et al., 2020). Les algorithmes machine learning peuvent ainsi classer automatiquement les clients en segments plus pertinents et personnalisés, permettant aux entreprises de mieux cibler leur audience et d'adapter leur stratégie marketing en conséquence (Shah et al., 2020).

Les méthodes de segmentation traditionnelles se basent généralement sur des données démographiques telles que l'âge, le sexe, le niveau de revenu ou le lieu de résidence. Cependant, le machine Learning permet d'aller au-delà de ces critères classiques en incluant des données comportementales telles que les préférences d'achat, les centres d'intérêts ou les habitudes de consommation. Cela permet aux entreprises de mieux comprendre les motivations et les besoins des consommateurs, et d'adapter leur stratégie marketing.

Par exemple, une entreprise de vente de vêtements pour femmes peut utiliser le machine learning pour segmenter son marché en fonction des goûts et des préférences vestimentaires de ses clients. En analysant les données de comportement d'achat et d'interaction des clients sur le site de l'entreprise, le machine learning peut détecter des similarités entre certains groupes de clients et créer des segments de marché plus pertinents tels que "les clientes intéressées par les vêtements de sport" ou "les clientes intéressées par la mode écologique".

L'analyse de la littérature montre que plusieurs chercheurs explorent la segmentation de la clientèle en e-marketing sous différents angles (Ahani et al., 2019, Alkhayrat et al., 2020, Aouad et al, 2023). En se basant sur les commentaires des clients (Nilashi et al., 2019), en évaluant la segmentation du marché en évaluant la satisfaction des voyageurs (Ahani et al., 2019b), ou en appliquant la gestion de la relation client (Florez-Lopez et Ramon-Jeronimo, 2009), le comportement des clients (Koehn et al., 2020), les préférences (Jagabathula et al., 2018) et les besoins (Kühl et al., 2019, Timoshenko et Hauser, 2019) des clients, et les facteurs psychographiques et cognitifs (Mostafa, 2009) pour procéder à la segmentation e-marketing/clientèle.

Smirnov et Huchzermeier (2020) ont integré la prévision de l'arrivée des clients, l'estimation du temps de service et la dotation en personnel dans un processus de planification du travail.

Dans une étude comparative, Koul et Philip (2021) ont examiné différentes techniques de segmentation de la clientèle. Les résultats ont montré que le regroupement est la meilleure

technique de segmentation de la clientèle. En particulier, l'algorithme K-Means s'est avéré être la méthode la plus efficace et la plus simple parmi les méthodes de regroupement.

Daoud et al. (2015) ont réalisé une étude de cas portant sur l'application de l'analyse RFM et des méthodes de regroupement, à savoir SOM et k-Means, dans le secteur du commerce électronique. Ils ont segmenté 730 clients en huit clusters sur la base des variables Récence, Fréquence et Monétaire, cherchant à mesurer les valeurs des clients et à identifier les clients les plus rentables afin de développer des stratégies marketing efficaces pour chaque cluster. Les résultats ont montré que le groupe 7 est le plus important d'après les valeurs moyennes de R, F et M.

Dans une autre étude, Maryani et al. (2018) ont également cherché à effectuer une segmentation de la clientèle en traitant 82 648 transactions "Nine Reload Credit". Ils ont employé un modèle d'exploration de données basé sur la segmentation RFM et utilisant l'algorithme de clustering k-means et les outils Rapidminer 5.2 qui ont été utilisés pour déterminer le nombre de clients dans chaque cluster en termes de R, F et M.

Shihab et al. (2019) ont réalisé une analyse comparative entre le clustering agglomératif, le clustering k-means et une version avancée du clustering k-means basée sur la segmentation RFM. Les résultats de cette étude ont révélé que le clustering agglomératif nécessite un temps de traitement plus long pour les grands ensembles de données que le k-means et le RFM k-means. La méthode RFM k-means s'est donc avérée être la meilleure, car elle nécessite un temps de traitement relativement faible et produit des résultats de regroupement plus élevés que les autres méthodes.

Rizki et al. (2020) ont segmenté la fidélité des clients du magasin Goldfinger. Ils ont également utilisé la segmentation RFM et le regroupement K-Means pour identifier les clients fidèles et potentiels.

Zhao et Li (2021) ont proposé une amélioration du modèle RFM intitulée RFAC pour segmenter les clients d'une entreprise de commerce électronique. Ils ont utilisé la méthode de pondération de l'entropie pour déterminer les poids des valeurs des caractéristiques et l'algorithme de clustering k-means++ pour regrouper les classes de clients. Les résultats ont montré que cette technique de segmentation de la clientèle est capable d'identifier avec précision la valeur des clients électroniques.

Shen (2021) a utilisé des méthodes d'apprentissage automatique non supervisées pour la segmentation de la clientèle. L'auteur a utilisé le modèle RFM pour créer les variables comportementales, la méthode TF-IDF pour les descriptions de produits afin de générer des catégories de produits, et l'algorithme de regroupement k-Means pour regrouper les clients. Pour l'analyse des produits achetés, il a utilisé l'extraction de règles d'association par l'algorithme Apriori. Il a également appliqué l'analyse en composantes principales (ACP) et les méthodes T-Distributed Stochastic Neighbor Embedding (T-sne) pour réduire la dimension des données à des fins de visualisation. Enfin, cette étude lui a permis de formuler des recommandations solides pour l'entreprise sur la base des résultats obtenus.

Grâce à une segmentation plus précise, l'entreprise peut proposer des produits et des offres plus adaptés aux besoins spécifiques de chaque segment, améliorant ainsi la satisfaction des clients et la performance de la stratégie Marketing de l'entreprise.

1.6.2. Détection de fraudes

La détection de fraude est également pertinente dans le contexte du marketing 7P. Elle représente est un défi majeur pour les entreprises de toutes tailles et de tous secteurs.

Les fraudes peuvent prendre de nombreuses formes, telles que l'usurpation d'identité, la fraude à la carte de crédit, la fraude d'assurance, la fraude de retour de marchandise et la fraude fiscale. Les entreprises peuvent subir des pertes financières considérables en raison de ces fraudes. C'est pourquoi, la détection de fraudes est devenue un domaine clé où les techniques machine learning sont de plus en plus utilisées.

Le machine learning peut être utilisé pour analyser de grandes quantités de données et identifier des modèles de comportement suspects. Les algorithmes machine learning peuvent être formés pour identifier des transactions frauduleuses en fonction de caractéristiques telles que les montants inhabituels, les lieux de transaction inhabituels, les tendances de comportement, les caractéristiques de l'appareil et les comportements de navigation (Loukili et al., 2023).

Il existe plusieurs algorithmes machine learning qui peuvent être utilisés pour la détection de fraudes, notamment :

Les arbres de décision : ces algorithmes peuvent être utilisés pour prendre des décisions basées sur des règles préétablies pour détecter les transactions frauduleuses.

Les réseaux de neurones : ces algorithmes peuvent être utilisés pour analyser de grandes quantités de données pour identifier des tendances et des modèles de comportement suspects.

Les algorithmes d'apprentissage non supervisé : ces algorithmes peuvent être utilisés pour découvrir des schémas de comportement inhabituels qui pourraient être des indicateurs de fraudes.

En utilisant le machine learning pour la détection de fraudes, les entreprises peuvent réduire les pertes financières causées par les fraudes. En outre, cela peut aider à améliorer la sécurité des clients et à renforcer la confiance des clients dans l'entreprise.

1.6.3. Amélioration de l'expérience client

L'amélioration de l'expérience client est l'un des aspects importants de l'élément "Processus" du marketing 7P. Les entreprises peuvent mettre en place des processus conviviaux et optimisés pour garantir une expérience utilisateur fluide et agréable, car les clients s'attendent à des expériences personnalisées et adaptées à leurs besoins. Le machine learning peut être utilisé pour améliorer l'expérience client en offrant des recommandations et des offres personnalisées, en améliorant la navigation sur le site Web, en offrant une assistance en temps réel et en améliorant la satisfaction globale du client.

Les algorithmes machine learning peuvent être utilisés pour analyser les données comportementales des clients, telles que l'historique des achats, les interactions avec le site Web et les commentaires des clients, afin de comprendre les besoins et les préférences des clients. Ces données peuvent ensuite être utilisées pour offrir des recommandations et des offres personnalisées qui répondent aux besoins spécifiques des clients.

Le machine learning peut également être utilisé pour améliorer la navigation sur le site Web en offrant des suggestions de produits, en optimisant les résultats de recherche et en offrant une navigation plus intuitive. En offrant une assistance en temps réel, telle que la fonction de chat en direct, les entreprises peuvent aider les clients à résoudre rapidement leurs problèmes en répondant à leurs questions.

Enfin, le machine learning peut être utilisé pour améliorer la satisfaction globale du client en offrant une expérience client plus cohérente et personnalisée, en identifiant et en résolvant rapidement les problèmes de service à la clientèle et en offrant une expérience client plus fluide et intuitive.

Les avantages de l'amélioration de l'expérience client comprennent une augmentation de la satisfaction client, une augmentation de la fidélisation client et une augmentation des conversions.

Le machine learning permet d'optimiser l'expérience client en offrant des recommandations et des offres personnalisées, en améliorant la navigation sur le site Web, en offrant une assistance en temps réel et en améliorant la satisfaction globale du client. L'amélioration de l'expérience client peut aider les entreprises à augmenter la satisfaction client, la fidélisation client et les conversions.

1.7. Preuve physique

La « Preuve physique » en ligne se réfère aux éléments tangibles qui soutiennent le produit ou le service proposé en ligne, tels que le site web et l'application mobile (Moon et Armstrong, 2020).

1.7.1. Qualité du contenu

La qualité du contenu proposé sur les sites web et les réseaux sociaux est un aspect crucial en e-marketing. Les consommateurs sont souvent à la recherche d'informations pertinentes et utiles sur les produits ou les services avant de prendre une décision d'achat.

Les entreprises peuvent offrir une preuve physique solide en créant du contenu de qualité qui répond aux besoins et aux intérêts de leurs clients potentiels. Cela peut inclure des descriptions de produits détaillées, des vidéos de démonstration, des évaluations et des critiques de produits, des guides d'achat, des infographies et des articles de blog.

En fournissant des informations utiles et fiables, les entreprises peuvent renforcer la confiance et la crédibilité de leur marque auprès des consommateurs. Cela peut également aider à améliorer le référencement de leur site web et à attirer plus de trafic qualifié. En fin de compte, une preuve physique solide peut contribuer à stimuler les ventes et la fidélité des clients.

1.7.2. Réputation en ligne

Un autre aspect qui peut être intégré à l'élément preuve physique du marketing 7P est la réputation en ligne de l'entreprise. Les consommateurs font souvent des recherches en ligne avant d'acheter un produit ou un service, et les avis et les commentaires en ligne peuvent avoir un impact important sur leur décision d'achat.

Pour offrir une preuve physique solide, il est important que l'entreprise surveille et gère sa réputation en ligne. Cela peut inclure la réponse aux commentaires et aux avis négatifs, la mise en place de programmes de satisfaction client, et la création de contenu en ligne pour renforcer la confiance et la crédibilité de l'entreprise.

L'entreprise peut également utiliser des témoignages de clients satisfaits pour renforcer sa preuve physique. Les témoignages peuvent être publiés sur le site web de l'entreprise, sur ses réseaux sociaux ou sur des sites tiers spécialisés dans les avis et les commentaires de clients.

La littérature existante applique les technologies machine learning pour étudier les activités marketing liées aux preuves matérielles. Le marketing de contenu consiste à fournir des informations tangibles et à persuader le public à l'aide d'un contenu solide pour inciter les clients à acheter des biens ou des services (Mendonça, 2022).

Poecze et al. (2019) ont mesuré le succès de différents types de posts sur Facebook pour optimiser la communication de la marque sur les médias sociaux.

Tsao et al. (2019) ont développé une nouvelle méthode basée sur le ML pour mesurer les constructions marketing à travers l'analyse passive des données textuelles générées par les consommateurs.

Rietveld et al. (2020) ont exploré comment les appels de messages émotionnels et informatifs influencent l'engagement des clients sur Instagram. Ils ont constaté que les images positives à haut niveau d'éveil et les images négatives à faible niveau d'éveil stimulaient l'engagement des clients. Les appels informatifs ne stimulaient pas l'engagement des clients, à l'exception des appels informatifs liés à la marque.

Vazquez et al. (2014) ont classé les contenus générés par les utilisateurs en fonction des différentes étapes du parcours décisionnel du consommateur et des éléments du marketing mix.

Chapelle et al. (2015) ont exploré comment modéliser la prédiction des réponses dans la publicité display, qui est une forme de publicité en ligne où les annonceurs paient les éditeurs pour placer des publicités graphiques sur leurs pages web.

Jai et al. (2021) ont étudié l'effet de différents types d'informations sensorielles visuelles sur l'activation cérébrale précédant les décisions d'achat.

Salminen et al. (2019) ont comparé l'approche ML pour étiqueter et classer automatiquement différents types d'articles de presse en ligne pour l'efficacité du marketing de contenu.

En mettant en place une bonne gestion de la réputation en ligne et en utilisant des témoignages de clients satisfaits, l'entreprise peut renforcer sa preuve physique et ainsi inspirer confiance et crédibilité auprès de ses clients potentiels.

2. Tendances et Développements Futurs du Machine Learning dans le Marketing Digital

2.1. Traitement du langage naturel (NLP)

Le Traitement du Langage Naturel (NLP) est une branche de l'intelligence artificielle qui permet aux machines d'analyser, de comprendre et de générer le langage humain (Hirschberg et Manning, 2015). Dans le contexte du marketing numérique, le NLP a un large éventail d'applications potentielles, notamment l'analyse de sentiment, la traduction de langues, les chatbots, l'optimisation de contenu, et bien plus encore (El-Ansari et Beni-Hssane, 2023). Par exemple, le NLP peut être utilisé pour analyser les commentaires des clients, améliorer l'optimisation des moteurs de recherche, et automatiser le service client (Aldunate, 2022).

L'analyse de sentiments permet aux spécialistes du marketing de comprendre le sentiment des clients et de suivre la réputation de la marque, tandis que la traduction de langues peut aider à toucher un public mondial (Farzadnia et Vanani, 2022). Les chatbots peuvent fournir des recommandations personnalisées et automatiser le service client, et les techniques de NLP, telles que la génération de langage naturel, peuvent être utilisées pour créer du contenu personnalisé (Patel et Trivedi, 2020). Cependant, il existe également des défis et des limites à l'utilisation du NLP dans le marketing numérique, tels que des biais potentiels dans les algorithmes (Benbya et al., 2020) et des limites dans la traduction de machine pour certains types de contenu (Wu et al., 2022).

2.2. Réseaux adversoriels génératifs

Les réseaux adversoriels génératifs (Generative Adversarial Networks, GAN) sont un type de modèle d'apprentissage automatique composé de deux réseaux neuronaux, un générateur et un discriminateur, qui travaillent ensemble pour créer de nouvelles données (Li et al., 2019). Le générateur crée des données synthétiques, telles que des images ou des descriptions de produits, qui sont destinées à être similaires aux données réelles, tandis que le discriminateur essaie de distinguer entre les données réelles et synthétiques (Jain et al., 2022).

Les GAN ont un large éventail d'applications potentielles dans le e-marketing. Une application consiste à créer des publicités personnalisées pour chaque client. Le générateur peut être utilisé pour créer des données synthétiques, telles que des images ou des descriptions de produits, adaptées aux préférences et caractéristiques de chaque client (Gan et al., 2021). Cela peut conduire à des publicités plus efficaces qui résonnent avec le client et génèrent des conversions plus élevées.

Une autre application des GAN dans le marketing numérique est d'améliorer les résultats de recherche visuelle. Les GAN peuvent générer de nouvelles images similaires à la requête de recherche, fournissant un ensemble de résultats de recherche plus précis et pertinent (Zhang et al., 2020). Cela peut être particulièrement utile pour les entreprises de commerce électronique qui vendent des produits avec des caractéristiques visuelles uniques, comme les vêtements ou les meubles.

Les GAN peuvent également être utilisés pour générer de nouveaux designs ou variations de produits, ce qui peut être utile dans les tests A/B et le développement de produits. Par exemple, une entreprise pourrait utiliser les GAN pour générer différentes variations d'une image ou d'une description de produit et les tester avec différents publics pour déterminer la plus efficace (Sattigeriet al., 2020).

Les GAN sont un domaine en évolution rapide, et il y a de nombreux futurs développements potentiels qui pourraient avoir des applications dans le marketing numérique. Un développement potentiel est l'utilisation des GAN dans la réalité virtuelle et augmentée, permettant aux entreprises de créer des expériences plus réalistes et immersives pour les clients (Rosenberg, 2022). Un autre développement potentiel est l'utilisation des GAN dans le

traitement du langage naturel, permettant aux entreprises de générer un texte plus naturel et personnalisé (Evholt et Larsson, 2020).

2.3. Hyperpersonnalisation

Hyper-personalization est une technique qui consiste à personnaliser les messages et le contenu marketing pour chaque client individuel en fonction de ses intérêts, préférences et comportements (Vavliakis et al., 2019). Dans le marketing digital, l'hyper-personnalisation peut aider à augmenter l'engagement et les conversions en offrant un contenu ciblé et pertinent (Tyagi et Bhatnagar, 2021).

L'hyper-personnalisation a de nombreuses applications potentielles dans le marketing digital, telles que des recommandations de produits personnalisées, des campagnes d'e-mails personnalisées et de la publicité ciblée (Vinaykarthik, 2022). En utilisant les données des clients et les algorithmes d'apprentissage automatique, l'hyper-personnalisation peut aider les spécialistes du marketing à livrer le bon message à la bonne personne au bon moment (Valdez Mendia et Flores-Cuautle, 2022).

Il existe plusieurs techniques disponibles pour mettre en œuvre l'hyper-personnalisation dans le marketing digital, telles que le filtrage collaboratif, le filtrage basé sur le contenu et le filtrage hybride (Vavliakis et al., 2018). Ces techniques peuvent être utilisées pour analyser les données des clients et générer des recommandations et du contenu personnalisés.

À mesure que l'hyper-personnalisation est de plus en plus adoptée dans le marketing digital, il y a un besoin croissant de techniques et d'outils plus sophistiqués. Les développements futurs peuvent inclure l'utilisation de l'intelligence artificielle pour générer un contenu encore plus personnalisé, ainsi que le développement de modèles plus avancés pour prédire le comportement et les préférences des clients.

2.4. Analyse prédictive

L'analyse prédictive est une technique qui consiste à utiliser des données historiques et des algorithmes statistiques pour faire des prédictions sur des événements futurs (Yun et al., 2022). En marketing numérique, l'analyse prédictive peut être utilisée pour anticiper le comportement des clients, tel que la probabilité d'achat ou de désabonnement, et pour informer les stratégies et les campagnes de marketing (Siegel, 2013).

L'analyse prédictive a de nombreuses applications potentielles en e-marketing, telles que la segmentation des clients, l'optimisation du contenu et l'évaluation des prospects. En analysant les données clients et en prévoyant leur comportement futur, les marketeurs peuvent prendre des décisions plus éclairées sur la manière de s'engager avec leur public.

Il existe plusieurs techniques disponibles pour mettre en œuvre l'analyse prédictive en marketing numérique, telles que l'analyse de régression (Sakas et al., 2022), les arbres de décision (Armas-Arias et al., 2022) et les réseaux neuronaux (Bara et al., 2022). Ces techniques peuvent être utilisées pour analyser les données clients et générer des prédictions sur leur comportement futur.

À mesure que l'analyse prédictive continue d'évoluer, il y a un besoin croissant de techniques et d'outils plus avancés. Les développements futurs peuvent inclure l'utilisation des techniques d'intelligence artificielle plus sophistiquées pour générer des prédictions encore plus précises, ainsi que le développement de modèles plus développés pour prédire le comportement et les préférences des clients.

L'analyse prédictive présente de nombreux avantages potentiels pour les marketeurs numériques, et à mesure que la technologie continue d'évoluer, elle est susceptible de devenir un outil de plus en plus important pour comprendre et engager les clients.

2.5. Optimisation de la recherche vocale

La recherche vocale consiste à utiliser des commandes vocales pour rechercher des informations, des produits ou des services en ligne, au lieu de taper des requêtes de recherche dans un moteur de recherche (Schalkwyk et al., 2010). Avec l'essor des assistants virtuels comme Siri, Alexa, Cortana et Google Assistant, la recherche vocale est devenue un moyen de plus en plus populaire pour les gens d'accéder à des informations en ligne (Tulshan et Dhage, 2019).

La recherche vocale a des implications significatives pour le marketing numérique, car elle oblige les spécialistes du marketing à optimiser leur contenu et leurs sites Web pour les requêtes de recherche vocale. Cela comprend la compréhension de la manière dont les gens utilisent la recherche vocale, les types de requêtes qu'ils sont susceptibles d'utiliser, et la façon de structurer le contenu pour qu'il soit facilement découvrable par la recherche vocale (Peyravi et al., 2020).

Il existe plusieurs techniques disponibles pour optimiser le contenu et les sites Web pour la recherche vocale, telles que l'utilisation de mots clés longue traîne (Gokkoeva, 2020), la fourniture de données structurées (Lambrecht et Peter, 2022) et la création de contenu qui répond aux questions courantes (Britchenko et al., 2019). Les spécialistes du marketing peuvent également utiliser des outils tels que les analyses de recherche vocale de Google pour analyser les données de recherche vocale et identifier des opportunités d'optimisation.

À mesure que la technologie de recherche vocale continue d'évoluer, il y a un besoin croissant de techniques et d'outils avancés pour l'optimisation de la recherche vocale. Les développements futurs peuvent inclure l'utilisation du traitement du langage naturel et des modèles machine learning et deep learning plus avancés afin de mieux comprendre et répondre aux requêtes de recherche vocale, ainsi que l'intégration de la recherche vocale avec d'autres canaux marketing, tels que les chatbots intelligents et les applications de messagerie (Krishna et al., 2023).

L'optimisation de la recherche vocale devient un aspect de plus en plus important du marketing numérique, car de plus en plus de personnes utilisent la recherche vocale pour accéder à des informations en ligne. En comprenant comment optimiser le contenu et les sites Web pour la recherche vocale, les spécialistes du marketing peuvent améliorer leur visibilité et leur engagement avec les clients qui utilisent cette technologie.

2.6. Enchères automatisées

Les enchères automatisées (Automated Bidding) se réfèrent à l'utilisation d'algorithmes d'apprentissage automatique pour automatiser le processus d'enchères sur les stocks publicitaires dans les enchères en temps réel (Shandilya et al., 2023). En analysant les données sur le comportement de l'utilisateur et la performance de la publicité, les enchères automatisées peuvent ajuster les enchères en temps réel pour optimiser les résultats souhaités, tels que la maximisation des conversions ou la minimisation du coût par clic (Christopher et al., 2022).

Les enchères automatisées offrent plusieurs avantages aux spécialistes du marketing numérique, notamment une efficacité accrue, une précision améliorée et la capacité à mieux adapter les campagnes. En automatisant le processus d'enchères, les spécialistes du marketing peuvent se concentrer sur d'autres aspects de la gestion de campagne, tels que le développement créatif et le ciblage d'audience (Kiran et Arumugam, 2022).

Il existe plusieurs techniques disponibles pour implémenter des enchères automatisées, telles que l'utilisation de systèmes basés sur des règles (Vij et al., 2015), des algorithmes d'apprentissage automatique ou deep learning (Jiao et al., 2020) et des outils tiers tels que Google Smart Bidding (Shandilya et al., 2023). Les spécialistes du marketing peuvent également utiliser les données sur le comportement de l'audience et la performance de la publicité pour informer les stratégies d'enchères et améliorer les résultats de la campagne.

À mesure que l'apprentissage automatique et l'intelligence artificielle continuent d'évoluer, il y a un besoin croissant de techniques d'enchères automatisées plus avancées qui peuvent s'adapter au comportement de l'utilisateur et à la performance de la publicité. Les développements futurs peuvent inclure l'utilisation des réseaux neuronaux pour améliorer la précision des enchères, ainsi que l'intégration des enchères automatisées avec d'autres canaux marketing, tels que les médias sociaux.

2.7. Vision par ordinateur

La vision par ordinateur (Computer Vision) est un domaine de l'intelligence artificielle qui vise à permettre aux machines de reconnaître et d'interpréter les données visuelles du monde qui les entoure (Bayoudh et al., 2021). Cela inclut les images, les vidéos et d'autres types de médias visuels.

En e-marketing, la vision par ordinateur présente un large éventail d'applications potentielles. Par exemple, elle peut être utilisée pour analyser et classer des images et des vidéos, identifier des objets et des logos dans des images et des vidéos, et suivre les mouvements et les comportements des utilisateurs dans des vidéos (Nanneet al., 2020).

Il existe plusieurs techniques disponibles pour mettre en œuvre la vision par ordinateur en marketing numérique, notamment les algorithmes d'apprentissage automatique, les logiciels de reconnaissance d'images et les réseaux de neurones à apprentissage profond. Les spécialistes du marketing peuvent utiliser ces techniques pour analyser les données visuelles, identifier des modèles et des tendances, et générer des informations qui peuvent informer leurs stratégies de marketing (Sümer et al., 2018).

À mesure que la technologie de la vision par ordinateur continue d'évoluer, il y a un besoin croissant de techniques plus avancées qui peuvent gérer de plus grands volumes de données et fournir des informations plus précises. Les développements futurs pourraient inclure l'utilisation

de technologies de réalité augmentée et de réalité virtuelle pour créer des expériences plus immersives pour les utilisateurs, ainsi que l'intégration de la vision par ordinateur avec d'autres canaux de marketing, tels que les médias sociaux et les applications mobiles (Philp et al., 2022).

La vision par ordinateur devient un aspect de plus en plus important du marketing numérique, car les spécialistes du marketing cherchent à mieux comprendre et à interagir avec leur public à travers les médias visuels. En comprenant comment mettre en œuvre les techniques de vision par ordinateur, les spécialistes du marketing peuvent obtenir des informations précieuses sur le comportement et les préférences des utilisateurs, et utiliser ces informations pour optimiser leurs campagnes de marketing pour de meilleurs résultats.

2.8. Réalité augmentée

La réalité augmentée (RA) est une technologie qui superpose des contenus numériques sur le monde réel, généralement en utilisant l'appareil photo d'un smartphone ou d'un autre dispositif (Berryman, 2012). Cela peut inclure des textes et des graphiques, des modèles 3D et des animations.

La réalité augmentée a de nombreuses applications potentielles dans le domaine du marketing digital. Par exemple, elle peut être utilisée pour créer des démonstrations de produits immersives, permettre aux utilisateurs d'essayer virtuellement des vêtements ou des accessoires, et fournir des expériences interactives qui engagent les utilisateurs avec une marque (Rauschnabel et al., 2022).

Pour mettre en œuvre la réalité augmentée dans le e-marketing, les marketeurs peuvent utiliser divers outils et techniques, notamment des plateformes de développement de RA, des applications mobiles et des expériences de RA basées sur le web (Herpich et al., 2017). Ils peuvent également exploiter des technologies telles que le traitement du langage naturel pour améliorer l'expérience utilisateur et fournir un contenu plus personnalisé (Jang et Park, 2022).

À mesure que la technologie de la réalité augmentée continue d'évoluer, il y a un besoin croissant de techniques et de plateformes plus avancées capables de traiter des volumes de données plus importants et de fournir des expériences plus sophistiquées pour les utilisateurs. Les développements futurs peuvent inclure l'intégration de la RA avec d'autres canaux de marketing, tels que les médias sociaux et les moteurs de recherche, ainsi que l'utilisation de la RA dans les environnements de vente au détail physique pour améliorer l'expérience d'achat.

La réalité augmentée est un outil puissant pour les marketeurs qui cherchent à engager les utilisateurs avec des expériences immersives et interactives qui peuvent stimuler l'engagement et les conversions. En comprenant comment mettre en œuvre la RA dans leurs stratégies de marketing, les marketeurs peuvent créer un contenu convaincant qui résonne avec leur public et les aide à atteindre leurs objectifs commerciaux.

2.9. Edge Computing

L'informatique en périphérie ou le Edge Computing est un paradigme de calcul distribué qui consiste à traiter les données et à exécuter des applications plus près de la source des données, au lieu de s'appuyer sur des centres de données centralisés (Bilal et al., 2018). Cette approche peut réduire la latence et l'utilisation de la bande passante, ce qui la rend idéale pour les applications en temps réel telles que le marketing numérique (Kong et al., 2022).

Le Edge Computing a de nombreuses applications dans le marketing numérique, telles que la diffusion de contenu personnalisé aux clients en temps réel en fonction de leur emplacement, de leur historique de navigation et d'autres facteurs contextuels. Le Edge Computing peut également être utilisé pour traiter les données des dispositifs IoT et des capteurs, permettant aux spécialistes du marketing de mieux comprendre le comportement et les préférences des clients (Sachdev, 2020).

L'un des principaux avantages du Edge Computing dans le marketing numérique est sa capacité à traiter les données et à exécuter des applications en temps réel, ce qui peut améliorer l'expérience client et l'efficacité des campagnes de marketing. Le Edge Computing peut également réduire la latence et améliorer la confidentialité et la sécurité des données, car les données n'ont pas besoin d'être transmises sur de longues distances vers des centres de données centralisés.

À mesure que le Edge Computing continue d'évoluer, plusieurs développements futurs pourraient avoir un impact sur son utilisation dans le marketing numérique. Il s'agit notamment de progrès dans le matériel de Edge Computing, tels que des dispositifs Edge plus puissants et de meilleures options de connectivité, ainsi que de l'intégration du Edge Computing avec d'autres technologies telles que l'IA et la blockchain. En outre, la montée des réseaux 5G et de l'Internet des objets (IoT) devrait stimuler davantage l'adoption du Edge Computing dans le marketing

numérique, les entreprises cherchant à exploiter des données et des informations en temps réel pour créer des campagnes de marketing plus personnalisées et plus efficaces.

Le Edge Computing est une technologie prometteuse qui peut aider les spécialistes du marketing à diffuser des campagnes de marketing plus personnalisées et plus efficaces. En traitant les données et en exécutant les applications plus près de la source, le Edge Computing peut réduire la latence, améliorer la confidentialité et la sécurité des données et permettre une prise de décision en temps réel. À mesure que le Edge Computing continue d'évoluer, il est susceptible de jouer un rôle de plus en plus important dans l'avenir du marketing digital.

Conclusion

En conclusion, la littérature existante sur l'application du machine learning dans le domaine du marketing digital offre un aperçu approfondi des nombreuses possibilités offertes par cette approche dans les différents aspects des 7P de l'e-marketing. Les études examinées démontrent le potentiel de l'apprentissage automatique pour améliorer la prédiction de la rétention des clients, la segmentation des clients, l'analyse des sentiments des avis produits, et bien d'autres aspects essentiels du marketing digital. Cependant, malgré les avantages et les opportunités offerts par le machine learning, plusieurs défis persistent. La confidentialité des données est un enjeu majeur, nécessitant une gestion rigoureuse pour garantir la protection des informations personnelles des utilisateurs.

La mise en œuvre du machine learning dans le marketing digital peut également être complexe, nécessitant des ressources techniques et une expertise appropriée pour développer, mettre en place et évaluer les modèles. De plus, la capacité à interpréter et à expliquer les décisions prises par les modèles machine learning reste un défi important pour assurer la confiance et la transparence. Malgré ces défis, les tendances et les développements futurs du machine learning dans le marketing digital ouvrent de nouvelles perspectives passionnantes. Des avancées telles que le traitement du langage naturel, les réseaux adversoriels génératifs, l'hyperpersonnalisation, l'analyse prédictive, l'optimisation de la recherche vocale, la vision par ordinateur, la réalité augmentée et l'edge computing promettent d'améliorer encore davantage les stratégies et les performances du marketing digital.

En somme, cette revue de littérature met en évidence l'importance croissante du machine learning dans le domaine du marketing digital et souligne les opportunités ainsi que les défis

auxquels sont confrontées les entreprises. La poursuite de la recherche et de l'innovation dans ce domaine permettra d'exploiter pleinement le potentiel du machine learning pour une stratégie de marketing digital plus efficace, personnalisée et axée sur les besoins des consommateurs.

Chapitre 3

Contributions à la littérature

Sommaire

INTRODUCTION....98
PARTIE I. LE MACHINE LEARNING POUR LA PREDICTION DU CHURN DES CLIENTS AVEC OPTIMISATION DES HYPERPARAMETRES....98
PARTIE II. AMELIORATION DE LA FIDELISATION DES CLIENTS GRACE AU DEEP LEARNING ET AUX TECHNIQUES DE DONNEES DESEQUILIBREES....114
PARTIE III. LE MACHINE LEARNING POUR LES SYSTEMES DE RECOMMANDATION DANS LE E-COMMERCE....127
PARTIE IV. SE DEFENDRE CONTRE LE VOL NUMERIQUE : UNE APPROCHE MACHINE LEARNING POUR PREDIRE LA FRAUDE AU PAIEMENT ELECTRONIQUE....138
PARTIE V. ANALYSE DES SENTIMENTS A TRAVERS LES COMMENTAIRES SUR LES PRODUITS POUR LA RECOMMANDATION DANS LE E-COMMERCE : UNE APPROCHE MACHINE LEARNING....152
PARTIE VI. PREDIRE LES VENTES FUTURES EN UTILISANT LES ALGORITHMES MACHINE LEARNING....164
PARTIE VII. LE MACHINE LEARNING AU SERVICE DE LA TARIFICATION DYNAMIQUE DANS LE E-MARKETING....174
CONCLUSION....184

Chapitre 3. Contributions à la littérature

Le présent chapitre constitue une synthèse des contributions scientifiques réalisées dans le cadre de cette thèse. Plusieurs domaines clés du marketing digital ont été abordés où le machine learning a été appliqué avec succès pour améliorer les stratégies et les résultats des entreprises.

La première partie de ce chapitre se concentre sur l'utilisation du machine learning pour la prédiction du churn des clients. La deuxième partie aborde l'amélioration de la fidélisation des clients grâce au deep learning et aux techniques de données déséquilibrées. La troisième partie se concentre sur les systèmes de recommandation dans le e-commerce. Dans la quatrième partie, l'analyse des sentiments des clients sera abordée. La cinquième partie se penchera sur l'utilisation du machine learning pour la prédiction des ventes. Ensuite, la sixième partie abordera la détection de la fraude au paiement électronique. Enfin, la septième partie mettra en lumière l'application du machine learning dans la tarification dynamique dans le marketing digital.

Partie I. Le machine learning pour la prédiction du churn des clients avec optimisation des hyperparamètres

1. Contexte

En raison de la mondialisation et des avancées dans l'industrie des télécommunications, la croissance exponentielle du nombre d'opérateurs sur le marché conduit à une augmentation de la concurrence. Dans cet environnement concurrentiel, il est devenu essentiel de maximiser continuellement les profits. Par conséquent, différentes approches ont été développées, telles que l'acquisition de nouveaux clients, la vente incitative aux clients existants et le renforcement de la période de fidélité des clients existants. Cependant, la stratégie de rétention des clients existants est considérée comme la plus simple, la moins coûteuse par rapport aux autres stratégies et la plus rentable. Les entreprises qui mettent en œuvre cette stratégie sont poussées à minimiser le taux de résiliation des clients, qui est principalement dû à l'insatisfaction à l'égard des services fournis aux consommateurs et au mécanisme de soutien. Pour résoudre ce problème, il est nécessaire de prédire quels clients sont susceptibles de résilier leur abonnement. Les modèles de prédiction de résiliation visent à détecter les signaux d'alerte précoce de la résiliation. Grâce à l'approche prédictive rendue possible par l'apprentissage automatique, il est possible de lutter efficacement contre la résiliation. En étant capable d'anticiper les

comportements des utilisateurs, les entreprises peuvent mesurer la relation client et agir préventivement pour éviter la résiliation. De cette manière, les entreprises peuvent s'engager dans des actions marketing plus rentables que l'acquisition de nouveaux clients. Du point de vue de l'apprentissage automatique, la prédiction de résiliation est considérée comme un problème d'apprentissage supervisé de classification. Cette prédiction peut être effectuée en se basant sur plusieurs algorithmes de classification d'apprentissage automatique.

Dans cette étude, un processus d'apprentissage automatique a été mis en place en se basant sur un ensemble de données provenant de Kaggle (Telco Customer Churn). Ensuite, les performances des différents modèles utilisés, à savoir le k-plus proche voisin, la régression logistique, la forêt aléatoire et la machine à vecteurs de support, ont été comparées. Le pipeline utilisé comprend les étapes suivantes :

- Définition du problème
- Collecte des données
- Analyse exploratoire des données et nettoyage des données
- Visualisation des données
- Ingénierie des caractéristiques
- Division de l'ensemble d'entraînement / ensemble de test
- Entraînement, prédiction et évaluation
- Ajustement des hyperparamètres
- Comparaison des mesures de performance

2. Matériaux

2.1. Description du problème

La prédiction de la résiliation des clients est jugée cruciale pour les organisations de télécommunications afin d'assurer une rétention efficace de leurs clients. Cependant, il est moins coûteux de conserver les clients existants que d'attirer de nouveaux clients. C'est pourquoi les grandes entreprises de télécommunications cherchent à développer des modèles permettant de prévoir les clients les plus susceptibles de résilier et d'agir en conséquence. Dans ce contexte, des modèles d'apprentissage automatique seront mis en place pour prédire la probabilité de résiliation d'un client en analysant ses caractéristiques, telles que ses informations

démographiques, ses informations de compte et les informations sur les services auxquels il est abonné. Ensuite, les performances de ces modèles seront comparées en termes de différents paramètres tels que l'exactitude, la précision, le rappel, le score F1 et la courbe ROC. Le modèle présentant les meilleures performances et susceptible de réduire la résiliation sera sélectionné, ce qui contribuera à augmenter la satisfaction des clients et les revenus de l'entreprise.

2.2. Description de l'ensemble de données

L'ensemble de données utilisé dans cette étude est disponible sur le site web Kaggle (le dataset est une version adaptée des données originales d'IBM). Il est composé de vingt colonnes de variables qui indiquent les caractéristiques des clients d'une entreprise de télécommunications, ainsi que de la variable cible "Churn", comme le montre le tableau 1 et le tableau 2. La colonne "Churn", qui représente la variable cible, indique si le client a résilié son contrat au cours du dernier mois ou non. La classe "No" contient les clients qui n'ont pas résilié, tandis que la classe "Yes" contient les clients qui ont résilié leur contrat avec l'entreprise. L'objectif de cette analyse est de trouver la relation entre les caractéristiques des clients et la variable de résiliation.

Tableau 1. « L'ensemble de données du churn des clients ».

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	DeviceProtection	TechSupp
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	No	
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	Yes	
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	No	
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	Yes	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	No	
5	9305-CDSKC	Female	0	No	No	8	Yes	Yes	Fiber optic	No	...	Yes	
6	1452-KIOVK	Male	0	No	Yes	22	Yes	Yes	Fiber optic	No	...	No	
7	6713-OKOMC	Female	0	No	No	10	No	No phone service	DSL	Yes	...	No	
8	7892-POOKP	Female	0	Yes	No	28	Yes	Yes	Fiber optic	No	...	Yes	
9	6388-TABGU	Male	0	No	Yes	62	Yes	No	DSL	Yes	...	No	

10 rows × 21 columns

Tableau 2. « Liste des caractéristiques du dataset utilisé ».

Caractéristiques	Type
PhoneService	Services aux clients
MultipleLines	
InternetServices	
DeviceProtection	
OnlineBackup	
OnlineSecurity	
TechSupport	
StreamingMovies	
StreamingTV	
Tenure	Informations sur le compte du client
Contract	
PaperlessBilling	
PaymentMethod	
MonthlyCharges	
TotalCharges	
CustomerID	Informations démographiques du client
Gender	
SeniorCitizen	
Partner	
Dependents	
Churn	Variable cible

3. Modèles de classification d'apprentissage automatique utilisés

Plusieurs algorithmes de classification d'apprentissage automatique sont utilisés pour prédire la résiliation des clients dans l'industrie des télécommunications. Dans ce qui suit, quatre techniques de classification utilisées dans cette étude pour la prédiction de l'attrition des clients sont brièvement présentées.

- Le k-plus proches voisins est un algorithme de classification supervisée utilisé dans diverses applications, y compris la prédiction de la résiliation des clients. Il vise à classer les points cibles ou la classe inconnue, en fonction de leur distance par rapport aux points constituant un échantillon d'entraînement dont la classe est connue a priori. La classe de la nouvelle instance est définie en fonction des votes majoritaires de ses voisins. Il s'agit d'un algorithme simple et non paramétrique qui ne nécessite que la définition d'une valeur d'entrée k représentant le nombre de voisins participant à la classification.
- La régression logistique est un algorithme de classification statistique probabiliste qui fonctionne mieux dans les problèmes de classification binaire. Il est utilisé pour interpréter la relation entre une variable dépendante binaire et un ensemble de variables indépendantes. Pour prédire la résiliation des clients, les variables indépendantes pondérées sont introduites dans la fonction logistique et produisent une probabilité comprise entre zéro et un. Ainsi, si la probabilité de résiliation d'un client dépasse 0,5, ce client est prédit comme un résiliant.
- Random Forest est un algorithme de classification ensembliste basé sur le concept d'arbre de décision, dans lequel la modélisation ensembliste est basée sur l'utilisation de deux modèles analytiques différents mais liés. Pour générer de nouveaux ensembles de données, RF applique l'approche du bagging et ajuste un modèle d'arbre de décision à chacun des nouveaux ensembles de données. Concrètement, le modèle Random Forest divise l'ensemble de données initial en sous-groupes aléatoires d'échantillons, puis génère un arbre de décision sur chaque sous-groupe lors de la phase d'apprentissage du modèle. Ensuite, lors de la phase de prédiction, chaque arbre de décision produit sa prévision. Le résultat de la prédiction correspond à la valeur de prédiction majoritaire.
- La machine à vecteurs de support (SVM) est un algorithme d'apprentissage supervisé qui effectue une analyse des données pour identifier des motifs. Il est utilisé pour résoudre des problèmes de classification, de régression ou de détection d'anomalies. Pour la prédiction de la résiliation des clients, ce modèle considère chaque observation comme des points dans un espace multidimensionnel et identifie les plans de séparation optimaux parmi les instances de différentes classes dans l'ensemble de données d'entraînement. Lors de l'étape de prédiction, il représente les nouvelles observations dans l'espace donné et les classe dans une classe en fonction de leur proximité avec l'espace de séparation. La technique du noyau est utilisée pour passer de la classification linéaire à la classification non linéaire. Les fonctions

de noyau permettent une représentation non linéaire des données d'entrée dans des espaces de caractéristiques multidimensionnels de manière efficace sur le plan computationnel.

4. Méthodologie

La figure suivante illustre le pipeline de cette étude pour la prédiction du churn des clients.

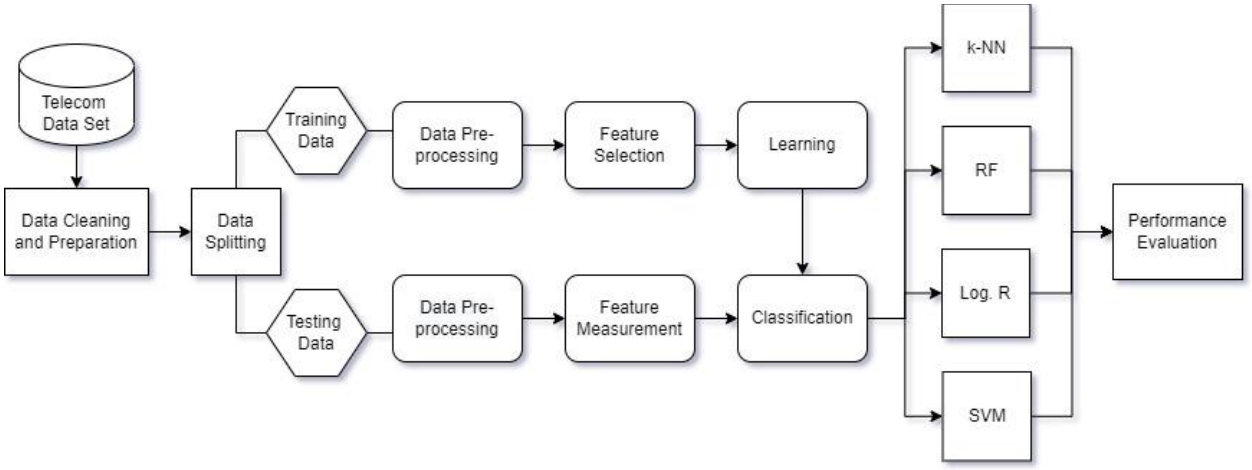


Figure 2. « Le pipeline de l'étude de prédiction du churn des clients en utilisant les algorithmes machine learning ».

4.1. Prétraitement des données

• **Nettoyage des données** : cette étape consiste à supprimer les données manquantes. Le tableau 3 montre que l'ensemble de données contient 11 valeurs manquantes pour la variable "TotalCharges". Pour simplifier, toutes les lignes de données manquantes ont été supprimées.

Tableau 3. « Résultats de vérification des données manquantes ».

```
customerID          0
gender              0
SeniorCitizen       0
Partner             0
Dependents          0
tenure              0
PhoneService        0
MultipleLines       0
InternetService     0
OnlineSecurity      0
OnlineBackup        0
DeviceProtection    0
TechSupport         0
StreamingTV         0
StreamingMovies     0
Contract            0
PaperlessBilling    0
PaymentMethod       0
MonthlyCharges      0
TotalCharges       11
Churn               0
dtype: int64
```

• **Ingénierie des caractéristiques** : les étapes suivantes, basées sur les types de données, sont entreprises pour prétraiter les caractéristiques :

- Suppression des variables non pertinentes : la colonne "CustomerID" a été supprimée car elle n'est pas significative.
- Encodage des étiquettes : les caractéristiques suivantes : "gender", "Partner", "Dependents", "PaperlessBilling", "PhoneService", "Churn", sont catégorielles et ont chacune deux valeurs (oui, non). Par conséquent, elles sont transformées en entiers binaires.
- Encodage one-hot : les caractéristiques suivantes : "InternetService", "MultipleLines", "OnlineSecurity", "OnlineBackup", "TechSupport", "DeviceProtection", "StreamingMovies", "StreamingTV", "PaymentMethod", "Contract", sont catégorielles, mais non ordinales et contiennent plus de deux valeurs. Une nouvelle variable est générée pour chaque valeur, avec un entier binaire indiquant si la valeur apparaît dans une entrée de données (1,0).
- Normalisation des données : les caractéristiques suivantes : "MonthlyCharges", "TotalCharges", "Tenure", sont numériques. Elles sont mises à l'échelle entre 0 et 1 à l'aide du scaler min-max, qui vise à mettre les valeurs sur une moyenne de 0 et un écart-type de 1.
- Transformation de toutes les caractéristiques catégorielles disponibles dans l'ensemble de données en caractéristiques numériques.
- Création d'une nouvelle variable "N_services" en additionnant le nombre de services auxquels le client est abonné ("OnlineSecurity", "OnlineBackup", "DeviceProtection", "StreamingMovies", "StreamingTv", "TechSupport").

• **Visualisation des données** : la figure suivante (figure 3) montre un diagramme à barres du pourcentage d'observations correspondant à chaque classe de la variable de réponse "Churn" : non et oui. Comme le montre la figure ci-dessous, l'ensemble de données contient 73,4% de non-résiliants et 26,6% de clients ayant résilié le service.

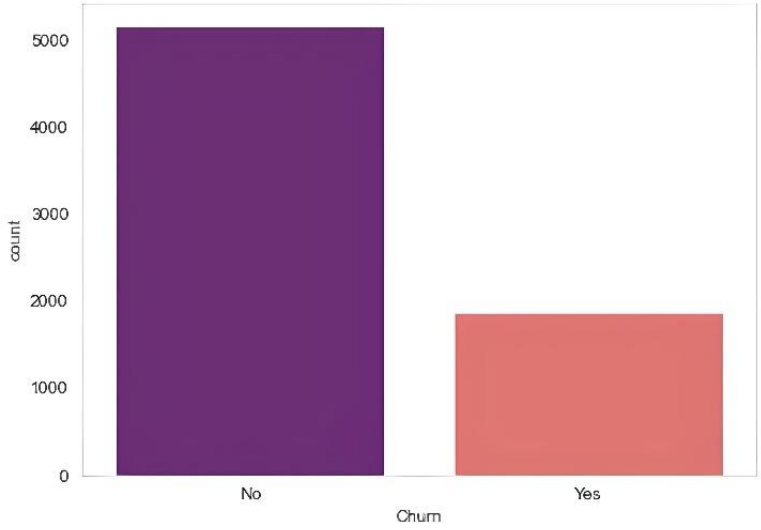


Figure 3. « Répartition des clients résiliants et non résiliants ».

Sur la base du graphique des variables "TotalCharges", "MonthlyCharges" et "tenure" par rapport à la variable "Churn", comme indiqué dans la figure 4, on peut constater que la variable "TotalCharges", qui est la somme des "MonthlyCharges", fournit des informations plus précises sur la résiliation au niveau individuel. De plus, les résiliants ont des "MonthlyCharges" plus élevés, avec une médiane d'environ 80 $ et une plage interquartile beaucoup plus basse que les non-résiliants, avec une médiane d'environ 65 $. De même, sur la base du graphique de la variable "tenure", les résiliants ont une durée de vie très courte, avec une médiane d'environ 10 mois, tandis que la médiane pour les non-résiliants est d'environ 38 mois.

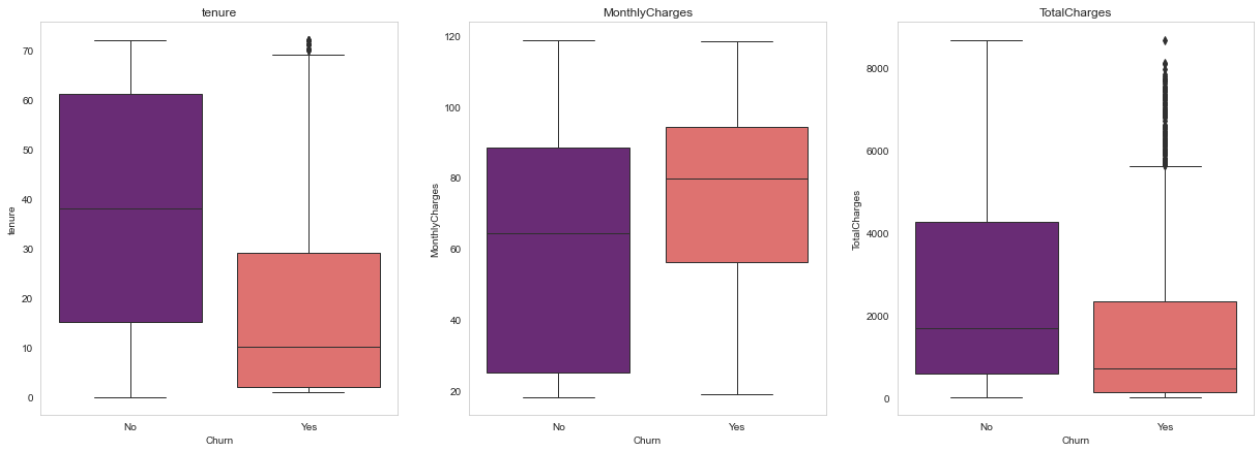


Figure 4. « Répartition des variables "TotalCharges", "MonthlyCharges" et "tenure" par rapport à la variable "Churn" ».

En se basant sur les graphiques des variables "Partner", "Dependents", "SeniorCitizen" et "Contract", comme indiqué dans la figure 5, les observations suivantes peuvent être faites :

- Les clients sans partenaires ont un taux de résiliation plus élevé.

- Les clients indépendants ont un taux de résiliation plus élevé par rapport aux clients dépendants.
- Les personnes âgées ont un taux de résiliation beaucoup plus élevé que les personnes non âgées.
- Les clients avec des contrats mensuels ont tendance à résilier par rapport aux clients avec des contrats d'un ou deux ans.

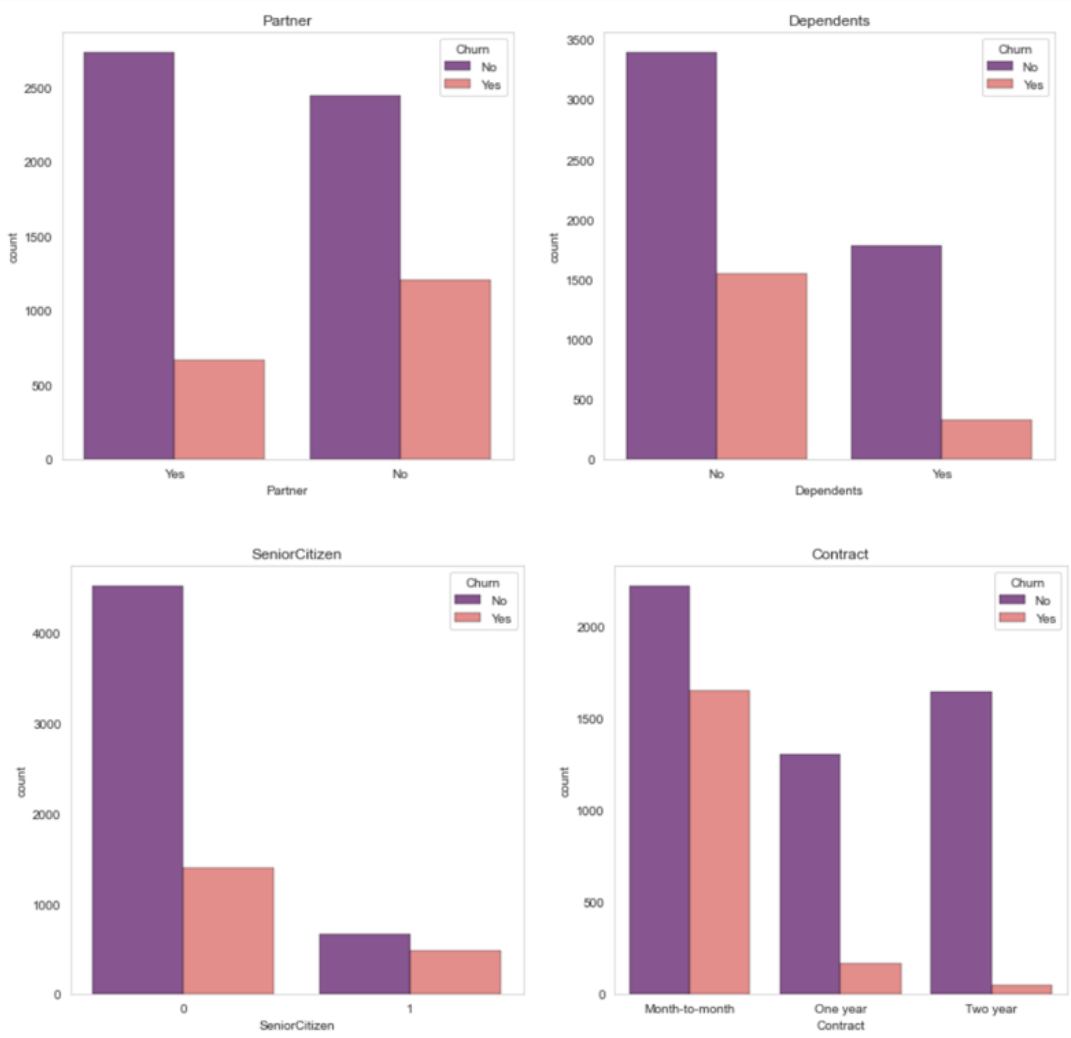


Figure 5. « La répartition des variables "Partner", "Dependents", "SeniorCitizen" et "Contract" par rapport à la variable "Churn" ».

La figure 6 est le graphique de la variable "N_services" qui représente le nombre de services auxquels le client est abonné. On peut constater que plus le nombre de services est élevé, plus le taux de résiliation est bas.

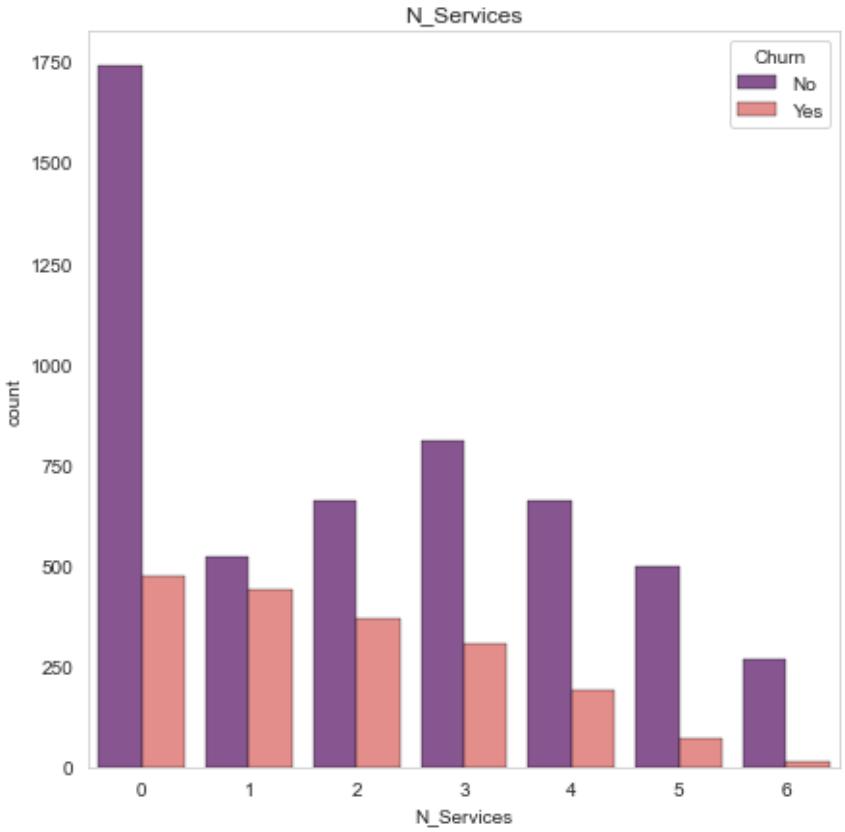


Figure 6. « Le nombre de services par rapport à la variable "Churn" ».

• **Corrélation :** la figure 7 affiche le diagramme de corrélation pour analyser la corrélation de la résiliation avec chacune des autres caractéristiques.

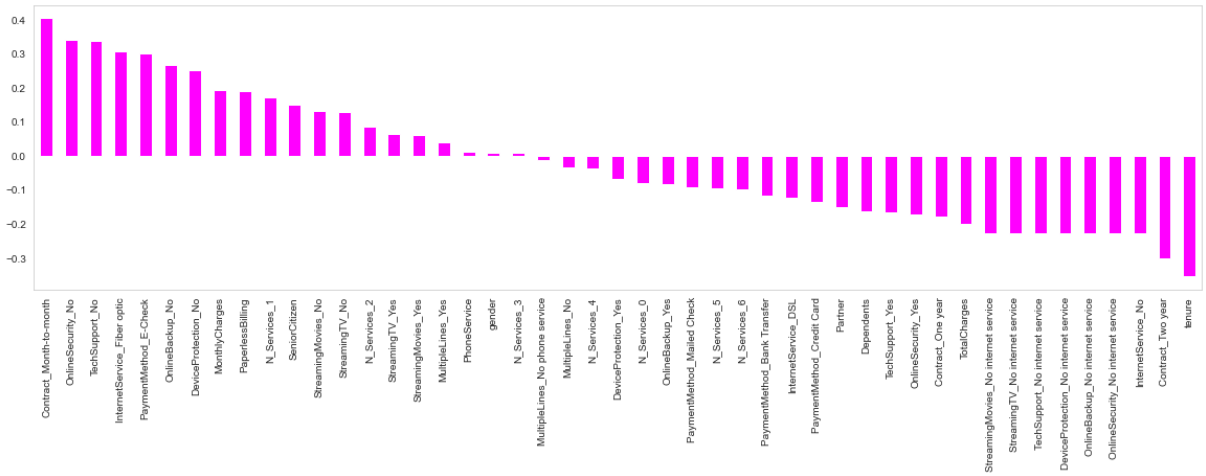


Figure 7. « La corrélation entre la variable cible et les autres variables ».

4.2. Ensembles d'entraînement et de test

Après les étapes de prétraitement des données, l'ensemble de données sur la résiliation est ensuite divisé en ensembles d'entraînement et de test. Les données sont divisées en deux parties, une pour l'entraînement et une pour les tests. Le paramètre du facteur de division a été

fixé à 70:30, soit 70% pour l'entraînement et 30% pour les tests, respectivement, pour de meilleures performances. Cela signifie que 4922 des données de l'ensemble seront les données d'entraînement et les 2110 autres seront les données de test. Cette méthode évite de réutiliser l'ensemble d'entraînement pour les tests et évite le surajustement en appliquant les classificateurs.

4.3. Mesures de performance du modèle

L'objectif de l'utilisation des mesures de performance est d'explorer les modèles utilisés et de présélectionner le modèle optimal. Tous les modèles candidats énumérés ci-dessus ont été entraînés en utilisant les mesures de performance suivantes : précision, taux d'erreur, rappel et score F1.

Les indicateurs suivants : précision, rappel, exactitude, taux d'erreur et score F1, sont calculés à partir des informations capturées à l'aide de la matrice de confusion présentée dans le tableau 4. Les vrais positifs et les faux positifs sont indiqués par TP et FP, et les faux négatifs et les vrais négatifs sont indiqués par FN et TN respectivement.

Les mesures de performance sont calculées comme suit :

$$\text{Précision} = \frac{\text{TP}}{\text{TP+FP}} \tag{8}$$

$$\text{Rappel} = \frac{\text{TP}}{\text{TP+FN}} \tag{9}$$

$$\text{Exactitude} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}} \tag{10}$$

$$\text{Taux d'erreur} = 1 - \text{Exactitude} \tag{11}$$

$$\text{Score F1} = \frac{2 * \text{Précision} * \text{Rappel}}{\textit{Précision+Rappel}} \tag{12}$$

Tableau 4. « Matrice de confusion ».

		Résultat de la prédiction	
		Positif	Négatif
Valeur réelle	Positif	TP	FN
	Négatif	FP	TN

4.4. Optimisation des hyperparamètres

La méthode de validation croisée holdout est une méthode simple qui consiste généralement à entraîner des modèles à partir de l'ensemble d'entraînement et à les évaluer à partir de l'ensemble de test. Cependant, ces hyperparamètres prédéfinis par défaut ne sont probablement pas adéquats. Pour cette raison, une méthode basée sur la méthode de validation croisée holdout a été développée, appelée validation croisée k-fold, qui est une méthode alternative utilisée lors du réglage des hyperparamètres avec "Grid Search" et "Randomized Search". Elle consiste à diviser l'ensemble de données en k sous-ensembles, et la méthode de validation croisée holdout est répétée k fois. Cette méthode est fréquemment utilisée lorsque les données sont déséquilibrées, car elle garantit que chaque sous-ensemble contient un pourcentage presque égal d'échantillons de chaque classe (résiliation et non-résiliation) que l'ensemble complet. Le résultat de la validation croisée est k valeurs pour toutes les métriques sur le k-fold CV. C'est-à-dire que dans le modèle k-NN, pour de meilleures performances, GridSearchCV est nécessaire pour déterminer le nombre optimal de voisins. Dans le modèle Random Forest, RandomizedSearchCV est utilisé pour optimiser les hyperparamètres max_features, n_estimators, max_depth, bootstrap et criterion. Dans le modèle de régression logistique, GridSearchCV est utilisé pour appliquer différentes valeurs de régularisation afin de réduire l'impact des caractéristiques insignifiantes à zéro ou de simplifier le modèle en relativisant les modèles forts identifiés lors de l'entraînement, afin de déterminer le meilleur modèle. Dans le modèle SVM, GridSearchCV est également appliqué pour déterminer la valeur optimale de la marge autour du vecteur de support.

5. Evaluation et la comparaison des modèles

Les paramètres des matrices de confusion (TP, TN, FP, FN) de chacun des modèles proposés sont présentés dans le tableau 5.

Tableau 5. « Matrices de confusion des modèles proposés ».

Classifieurs	TP	TN	FP	FN
k-NN	1036	345	126	603
Log R	1294	331	140	345
RF	1122	352	120	516
SVM	1441	320	152	197

À partir de la matrice de confusion des modèles utilisés, les mesures de performance suivantes sont calculées : précision, rappel et score F1. Les tableaux 6, 7, 8 et 9 présentent les mesures de performance pour chaque modèle proposé avant le réglage des hyperparamètres.

Tableau 6. « Mesures de performance du modèle SVM ».

	precision	**recall**	**f1-score**
Churn = No	0.90	0.88	0.89
Churn = Yes	0.62	0.68	0.65
moyenne	0.76	0.78	0.77
accuracy			83.46 %
taux d'erreur			16.54 %

Tableau 7. « Mesures de performance du modèle k-NN ».

	precision	**recall**	**f1-score**
Churn = No	0.89	0.63	0.74
Churn = Yes	0.36	0.73	0.49
moyenne	0.62	0.68	0.61
accuracy			65.45 %
taux d'erreur			34.55 %

Tableau 8. « Mesures de performance du modèle Log. R ».

	precision	**recall**	**f1-score**
Churn = No	0.90	0.79	0.84
Churn = Yes	0.49	0.70	0.58
moyenne	0.69	0.74	0.71
accuracy			77.01 %
taux d'erreur			22.99 %

Tableau 9. « Mesures de performance du modèle RF ».

	precision	**recall**	**f1-score**
Churn = No	0.90	0.68	0.78
Churn = Yes	0.41	0.75	0.53
moyenne	0.65	0.71	0.65
accuracy			69.86 %
taux d'erreur			30.14 %

Le tableau 10 résume les mesures de performance de tous les modèles proposés avant le réglage des hyperparamètres.

Tableau 10. « Mesures de performance des modèles proposés avant le réglage des hyperparamètres ».

Model	**accuracy**	**taux d'erreur**	**M. precision**	**M. recall**	**M. f1-score**
SVM	83.46 %	16.54 %	0.76	0.78	0.77
KNN	65.45 %	34.55 %	0.62	0.68	0.61
Log. R	77.01 %	22.99 %	0.69	0.74	0.71
RF	69.86 %	30.14 %	0.65	0.71	0.65

D'après le tableau ci-dessus, on peut constater que les paramètres de performance sont faibles avant le réglage des hyperparamètres. Par conséquent, les hyperparamètres ont été ajustés afin d'améliorer les performances des modèles utilisés. Le tableau 11 résume les mesures de performance de tous les modèles proposés après le réglage des hyperparamètres.

Tableau 11. « Mesures de performance des modèles proposés après le réglage des hyperparamètres ».

Model	**accuracy**	**taux d'erreur**	**M. precision**	**M. recall**	**M. f1-score**
SVM	96.92 %	3.08 %	0.94	0.97	0.95
KNN	79.05 %	20.95 %	0.72	0.78	0.73
Log. R	91.37 %	8.63 %	0.86	0.91	0.88
RF	88.82 %	11.18 %	0.83	0.90	0.85

Sur la base des tableaux ci-dessus, on peut observer une augmentation significative des mesures de performance après l'ajustement des hyperparamètres. Ainsi, selon les mesures de

performance, on peut voir que le modèle SVM surpasse les autres modèles avec une exactitude de 96,92 %, une précision de 0,94, un rappel de 0,97, un score F1 de 0,95 et un taux d'erreur de 3,08 %.

6. Synthèse

Prédire la résiliation est un enjeu majeur pour les entreprises. En effet, identifier les clients insatisfaits des services fournis permet aux entreprises de travailler sur leurs tarifs, leurs offres et leurs promotions, ainsi que sur les préférences des clients, afin de réduire la possibilité de résiliation des clients. Il existe de nombreuses techniques pour prédire la résiliation des clients. Cette étude s'est concentrée sur la comparaison de différents modèles de classification d'apprentissage automatique afin de trouver le modèle offrant les meilleures performances pour la prédiction de résiliation en fonction des mesures de performance. Les résultats ont montré que les modèles proposés ont obtenu une meilleure exactitude en ajustant les hyperparamètres des différents modèles via une validation croisée. Ainsi, le classifieur SVM a surpassé les autres modèles. Cette étude pourrait être améliorée en utilisant une analyse de données volumineuses. En effet, l'analyse des réseaux sociaux est fortement recommandée pour déterminer le taux de satisfaction des clients à l'égard des services d'une organisation de télécommunications, puis ces données peuvent être exploitées pour minimiser le taux de résiliation des clients. De plus, l'utilisation d'autres ensembles de données peut accroître le niveau de confiance des résultats.

Partie II. Amélioration de la fidélisation des clients grâce au deep learning et aux techniques de données déséquilibrées

1. Contexte

Le taux de désabonnement des clients, ou rotation des clients, est un problème courant dans les affaires qui se réfère à la perte de clients sur une période donnée. C'est un enjeu critique pour les entreprises, car cela peut entraîner des pertes financières et une baisse des revenus. Le désabonnement des clients peut résulter de divers facteurs, tels qu'un mauvais service client, des problèmes de qualité des produits ou la disponibilité de meilleures alternatives sur le marché.

Ces dernières années, le désabonnement des clients est devenu une préoccupation de plus en plus importante pour les entreprises en raison de la concurrence croissante dans de nombreux secteurs. Avec la montée du commerce électronique et des achats en ligne, les clients ont plus d'options que jamais, ce qui facilite leur passage à un concurrent s'ils ne sont pas satisfaits d'un produit ou d'un service.

Pour résoudre ce problème, les entreprises doivent prédire avec précision quels clients sont les plus susceptibles de se désabonner et prendre des mesures proactives pour les fidéliser. La prédiction du désabonnement des clients est donc une tâche essentielle pour les entreprises qui cherchent à améliorer la fidélisation des clients et réduire les pertes financières.

Les approches traditionnelles de prédiction du désabonnement des clients ont fait appel à des méthodes statistiques telles que la régression logistique ou les arbres de décision. Cependant, ces méthodes présentent des limites, telles que le besoin d'ingénierie manuelle des caractéristiques, ce qui peut être chronophage et nécessiter une expertise dans le domaine. De plus, ces méthodes peuvent ne pas être en mesure de capturer les relations complexes entre les comportements des clients et le désabonnement.

Récemment, les approches d'apprentissage profond, telles que les réseaux neuronaux profonds (DNN), ont gagné en popularité dans la modélisation prédictive en raison de leur capacité à apprendre et à extraire automatiquement des caractéristiques à partir de données brutes. Les DNN ont atteint des performances de pointe dans de nombreux domaines, tels que le traitement du langage naturel (NLP), la vision par ordinateur et la modélisation prédictive. Dans la prédiction du désabonnement des clients, les DNN peuvent aider les entreprises à

identifier des motifs et des relations dans des ensembles de données volumineux et complexes qui ne seraient pas évidents avec les méthodes statistiques traditionnelles.

Cependant, l'efficacité des DNN dans la prédiction du désabonnement des clients n'a pas été largement explorée, et il existe un manque dans la littérature concernant leur application à ce problème. De plus, les données déséquilibrées, où le nombre d'observations dans une classe l'emporte nettement sur l'autre, sont un problème courant dans la prédiction du désabonnement des clients. Les données déséquilibrées peuvent entraîner des modèles biaisés qui ont de mauvaises performances dans la prédiction de la classe minoritaire. Pour résoudre ce problème, des algorithmes de suréchantillonnage et de sous-échantillonnage peuvent être appliqués pour équilibrer la distribution des classes.

2. Méthodologie

Pour améliorer la prédiction de désabonnement des clients en utilisant les réseaux neuronaux profonds et les techniques de données déséquilibrées, les étapes suivantes (figure 8) ont été suivies dans cette étude :

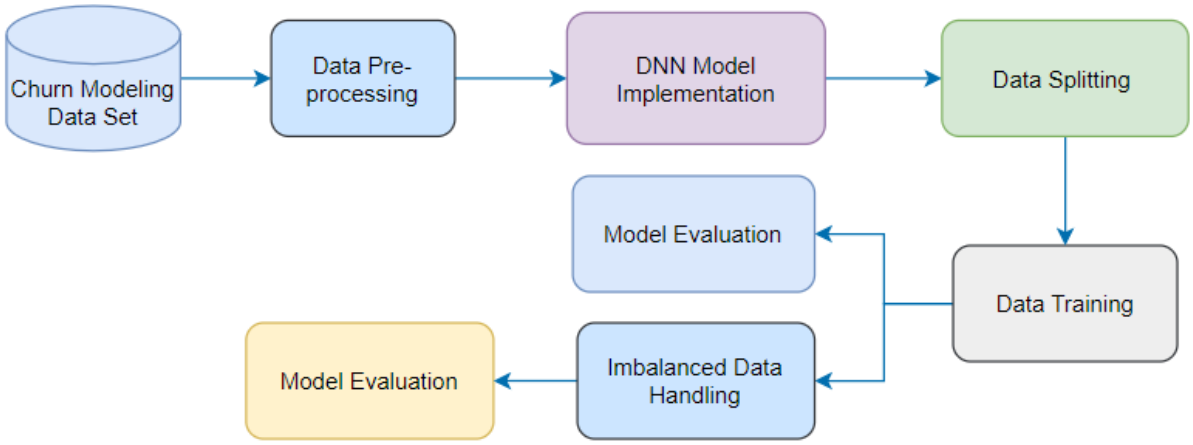


Figure 8. « Les étapes de la méthodologie adoptée pour la prédiction du désabonnement des clients en utilisant les réseaux neuronaux profonds et les techniques de données déséquilibrées ».

2.1. Prétraitement des données

Le dataset de modélisation de désabonnement utilisé se compose de 10 000 lignes et 11 colonnes contenant des détails sur les clients d'une banque (y compris : l'âge, le score de crédit, le genre, l'ancienneté, la géographie, le solde, le membre actif, la possession de carte de crédit, le nombre de produits et le salaire estimé) et la variable cible indiquant si le client a quitté la banque (fermé son compte) ou reste client (tableau 12).

Tableau 12. « L'ensemble de données utilisé pour la prédiction du désabonnement des clients ».

	Credit-Score	Geogra-phy	Gender	Age	Tenure	Balance	Num-Of-Products	Has-CrCard	IsActive-Member	Estimated Salary	Exit-ed
0	619	France	Female	42	2	0.00	1	1	1	101348.88	1
1	608	Spain	Female	41	1	83807.86	1	0	1	112542.58	0
2	502	France	Female	42	8	159660.80	3	1	0	113931.57	1
3	699	France	Female	39	1	0.00	2	0	0	93826.63	0
4	850	Spain	Female	43	2	125510.82	1	1	1	79084.10	0
…	…	…	…	…	…	…	…	…	…	…	…
9995	771	France	Male	39	5	0.00	2	1	0	96270.64	0
9996	516	France	Male	35	10	57369.61	1	1	1	101699.77	0
9997	709	France	Female	36	7	0.00	1	0	1	42085.58	1
9998	772	Germany	Male	42	3	75075.31	2	1	0	92888.52	1
9999	792	France	Female	28	4	130142.79	1	1	0	38190.78	0

Ensuite, les distributions des différentes variables en fonction de la variable cible ont été visualisées afin de comprendre l'ensemble de données et d'observer les variables qui influencent la variable cible, comme le montrent les figures 9 et 10.

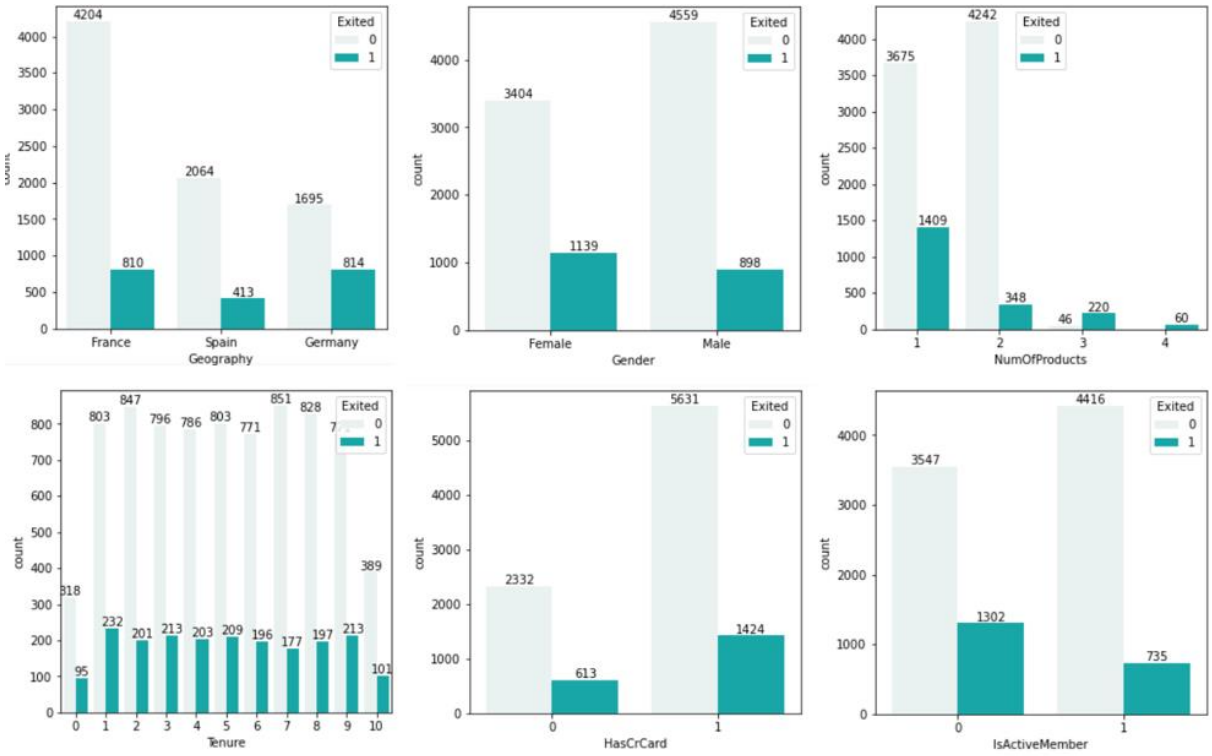


Figure 9. « Observation de l'ensemble des variables discrètes par rapport à la variable cible ''Exited''.

».

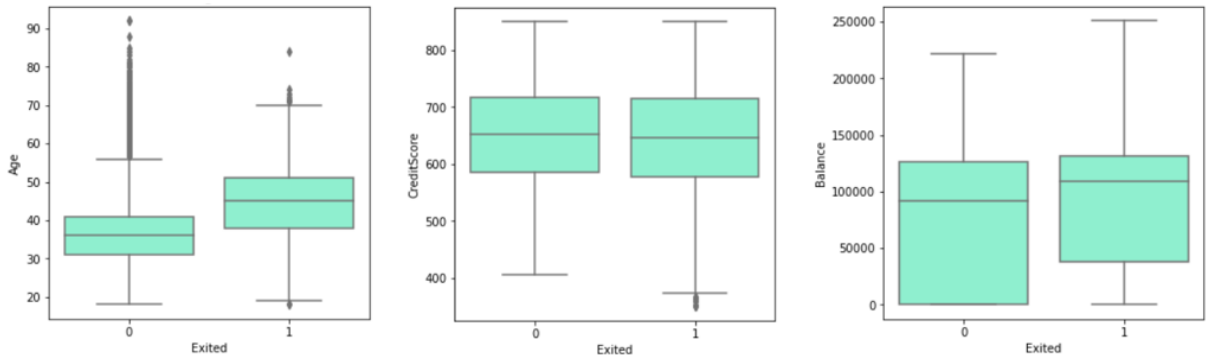
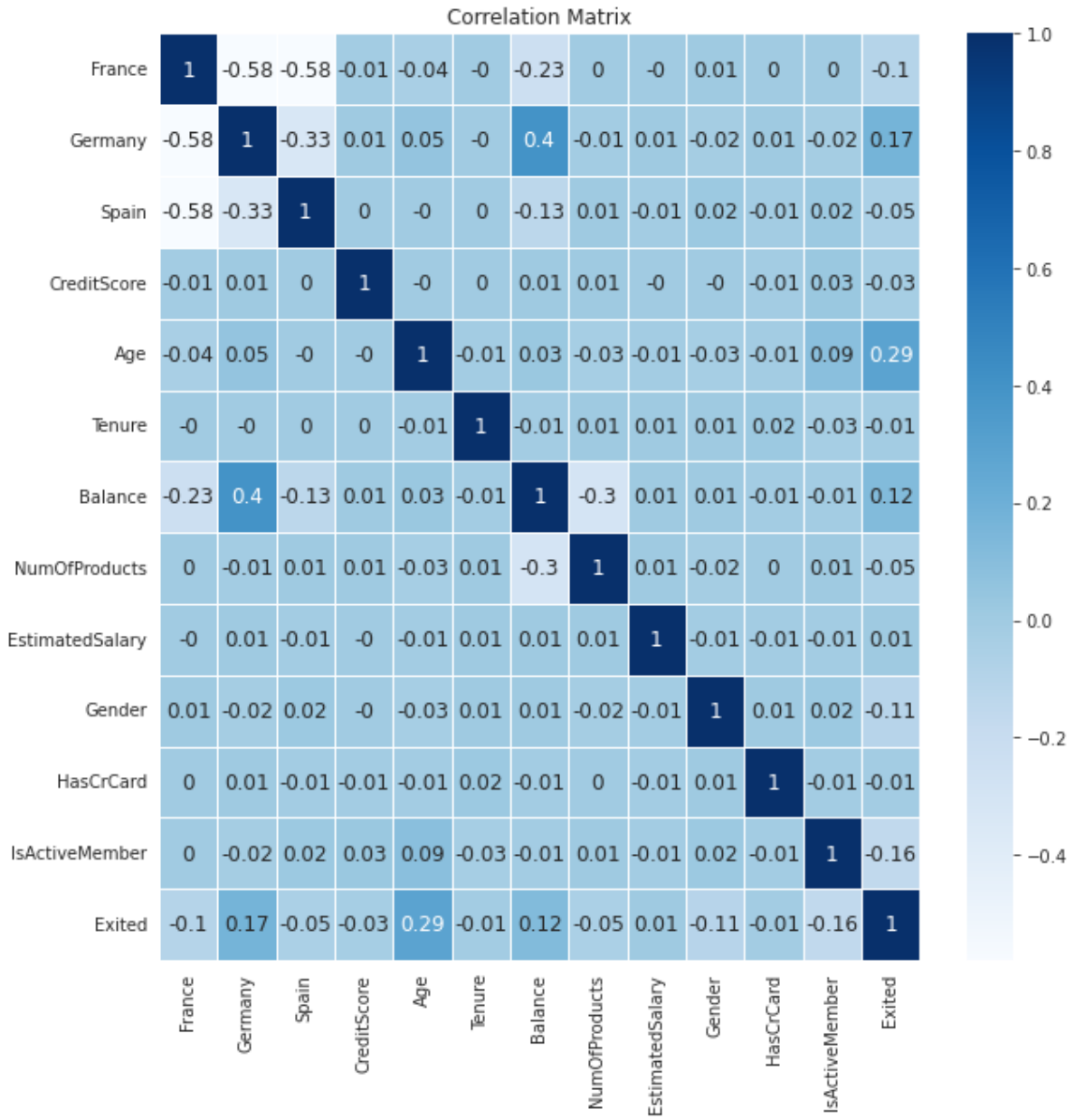


Figure 10. « La distribution des variables continues par rapport à la variable cible ».

Plusieurs étapes de prétraitement ont été effectuées sur les données avant la modélisation. Les caractéristiques ont été standardisées pour avoir une moyenne de zéro et une variance unitaire, les étiquettes cibles ont été encodées en entiers entre 0 et n-1 classes, et le codage one-hot a été utilisé pour convertir les caractéristiques catégorielles en un tableau numérique. La corrélation entre les caractéristiques a également été examinée pour déterminer les variables redondantes ou fortement corrélées, comme le montre le tableau 13.

Tableau 13. « La matrice de corrélation ».



2.2. Modèle prédictif

Pour le modèle prédictif, un réseau neuronal profond avec une architecture de modèle séquentiel a été utilisé. Le modèle DNN était composé de 11 couches, comprenant 8 couches denses entièrement connectées, 3 couches de désactivation pour la régularisation et une couche de sortie avec une fonction d'activation sigmoïde avec 1 unité, car la variable cible était binaire (soit 0 soit 1). La figure 11 présente l'architecture du modèle.

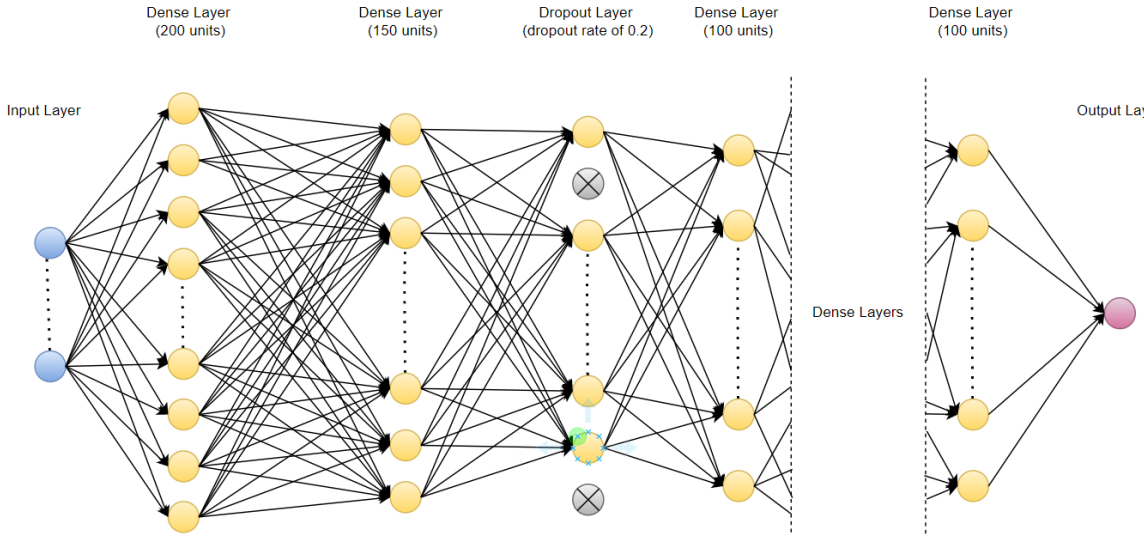


Figure 11. « Architecture du modèle Deep Neural Network ».

Les couches du modèle sont les suivantes :

- Couche d'entrée : vecteur d'entrée de 12 dimensions.
- Couche dense 1 : 200 neurones avec une fonction d'activation ReLU.
- Couche dense 2 : 150 neurones avec une fonction d'activation ReLU.
- Couche de désactivation 1 : applique une désactivation de 0,2.
- Couche dense 3 : 100 neurones avec une fonction d'activation ReLU.
- Couche dense 4 : 100 neurones avec une fonction d'activation ReLU.
- Couche de désactivation 2 : applique une désactivation de 0,2.
- Couche dense 5 : 100 neurones avec une fonction d'activation ReLU.
- Couche dense 6 : 100 neurones avec une fonction d'activation ReLU.
- Couche dense 7 : 100 neurones avec une fonction d'activation ReLU.
- Couche de désactivation 3 : applique une désactivation de 0,2.
- Couche dense 8 : 100 neurones avec une fonction d'activation ReLU.
- Couche de sortie : 1 neurone avec une fonction d'activation sigmoïde.

Le modèle a été compilé en appliquant l'optimiseur Adam (estimation adaptative du moment) et la fonction de perte de l'entropie croisée binaire, avec l'erreur quadratique moyenne et l'exactitude utilisées comme mesures pour évaluer les performances du modèle.

2.3. Entraînement

Le DNN pour la prédiction de la réduction du nombre de clients a été entraîné en utilisant la fonction "fit" de la bibliothèque Keras. L'ensemble de données a été divisé en ensembles d'entraînement et de test à l'aide de la fonction "train_test_split" de la bibliothèque sklearn.model_selection, avec une taille de l'ensemble de test de 10 % et l'état aléatoire réglé sur 42. Le modèle a ensuite été entraîné sur l'ensemble d'entraînement en utilisant une taille de

lot de 10 et 50 époques. Les données de validation (l'ensemble de test) ont également été incluses dans le processus d'entraînement pour évaluer les performances de généralisation du modèle.

2.4. Évaluation

Les performances du DNN dans la prédiction de la réduction du nombre de clients ont été évaluées en utilisant la fonction "evaluate" de la bibliothèque Keras sur l'ensemble de test. Cette fonction a fourni les métriques de l'erreur quadratique moyenne (MSE), de l'exactitude et de la perte du modèle.

La métrique MSE mesure la différence quadratique moyenne entre les valeurs prédites et réelles de la variable cible (un MSE plus faible indique un meilleur ajustement du modèle aux données). Le MSE est écrit comme suit (13) :

$$\text{MSE} = \frac{1}{n}\sum|yi - \hat{y}i|^2 \tag{13}$$

Où :

n = nombre total d'observations

y_i = valeur réelle de la variable cible pour la i-ème observation

$\hat{y}_i$ = valeur prédite de la variable cible pour la i-ème observation

La métrique d'exactitude (Accuracy) mesure le pourcentage de prédictions correctes faites par le modèle (une exactitude plus élevée indique un meilleur ajustement du modèle aux données). L'exactitude est écrite sous la forme suivante (14) :

$$\text{Accuracy} = \frac{number\ of\ correct\ predictions}{total\ number\ of\ predictions} * 100 \tag{14}$$

La métrique de perte (Loss) mesure la différence entre les valeurs prédites et réelles de la variable cible (une perte plus faible indique un meilleur ajustement du modèle aux données).

$$Loss = -\sum(yi * log(\hat{y}i) + (1 - yi) * log(1 - \hat{y}i)) \tag{15}$$

Où :

n = nombre total d'observations

y_i = valeur réelle de la variable cible pour la i-ème observation (0 ou 1)

$\hat{y}_i$ = probabilité prédite que la variable cible soit égale à 1 pour la i-ème observation

La perte est calculée en fonction de l'entropie croisée entre les distributions réelles et prédites de la variable cible.

L'application des techniques de données déséquilibrées au modèle consisterait d'abord à équilibrer la distribution des classes de l'ensemble de données en utilisant la technique souhaitée. L'ensemble de données équilibré résultant serait ensuite utilisé pour entraîner et évaluer le modèle en utilisant les mêmes étapes décrites ci-dessus.

2.5. Traitement des données déséquilibrées

En raison du déséquilibre des classes dans les données, comme le montre la figure 12, plusieurs techniques de sur-échantillonnage et de sous-échantillonnage ont été appliquées aux données d'entraînement pour résoudre le problème du déséquilibre des classes dans l'ensemble de données.

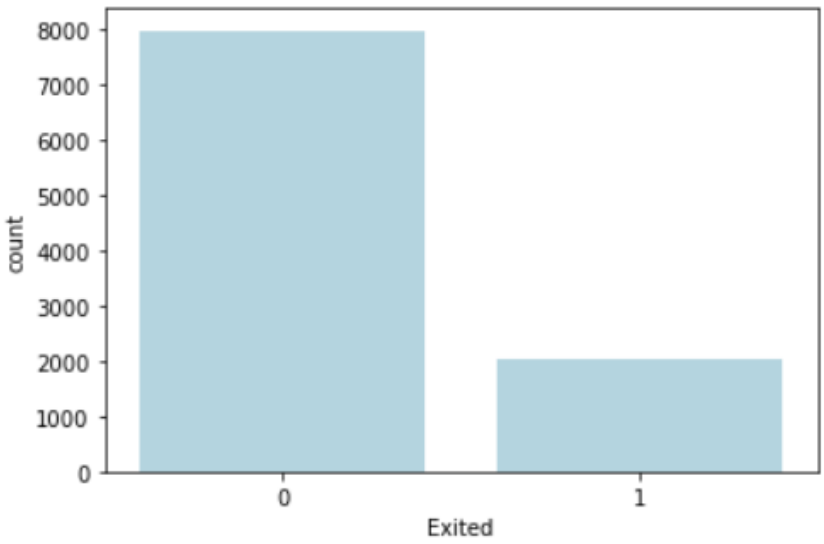


Figure 12. « Classes déséquilibrées de la variable cible ‘‘Exited’’ ».

Les techniques de sur-échantillonnage utilisées sont les suivantes : la technique SMOTE (Synthetic Minority Oversampling Technique), RandomOverSampler, BorderlineSMOTE, Adaptive Synthetic, KMeansSMOTE et Support Vector Machine SMOTE.

SMOTE (Synthetic Minority Oversampling Technique) : il s'agit d'une technique de sur-échantillonnage populaire qui génère des échantillons synthétiques de la classe minoritaire en interpolant entre les échantillons de la classe minoritaire existants. Les échantillons synthétiques

sont générés en choisissant deux échantillons de la classe minoritaire au hasard et en créant un nouvel échantillon à un point aléatoire le long de la ligne reliant les deux échantillons.

RandomOverSampler : il s'agit d'une technique de sur-échantillonnage simple qui sélectionne au hasard des échantillons de la classe minoritaire et les réplique jusqu'à ce que la distribution des classes soit équilibrée.

BorderlineSMOTE : il s'agit d'une technique de sur-échantillonnage qui génère des échantillons synthétiques de la classe minoritaire en se basant sur la frontière entre les échantillons de la classe minoritaire et de la classe majoritaire. Elle crée des échantillons synthétiques en sélectionnant au hasard un échantillon de la classe minoritaire et un échantillon de la classe majoritaire qui est le plus proche de l'échantillon de la classe minoritaire, puis en générant un nouvel échantillon à un point aléatoire le long de la ligne reliant les deux échantillons.

Adaptive Synthetic (ADASYN) : il s'agit d'une technique de sur-échantillonnage qui génère des échantillons synthétiques de la classe minoritaire en se basant sur la densité des échantillons dans l'espace des caractéristiques. Elle crée des échantillons synthétiques en sélectionnant un échantillon de la classe minoritaire et en trouvant les k plus proches voisins de cet échantillon dans l'espace des caractéristiques. Un nouvel échantillon est ensuite généré à un point aléatoire le long de la ligne reliant l'échantillon sélectionné et l'un de ses voisins, la probabilité de choisir un voisin étant basée sur la densité des échantillons dans le voisinage.

KMeansSMOTE : il s'agit d'une technique de sur-échantillonnage qui utilise l'algorithme de regroupement K-means pour regrouper les échantillons de la classe minoritaire et générer des échantillons synthétiques au sein des clusters. Elle commence par regrouper les échantillons de la classe minoritaire en k clusters à l'aide de K-means, puis génère des échantillons synthétiques au sein de chaque cluster en sélectionnant au hasard deux échantillons et en créant un nouvel échantillon à un point aléatoire le long de la ligne reliant les deux échantillons.

Support Vector Machine SMOTE (SVMSMOTE) : il s'agit d'une technique de sur-échantillonnage qui utilise l'algorithme de machine à vecteurs de support (SVM) pour générer des échantillons synthétiques de la classe minoritaire. Elle commence par entraîner un classifieur SVM sur les échantillons de la classe minoritaire, puis génère des échantillons synthétiques en échantillonnant à partir de la frontière de décision définie par le SVM.

Les techniques de sous-échantillonnage appliquées sont les suivantes : ClusterCentroids, AllKNN, NeighbourhoodCleaningRule et RandomUnderSampler.

ClusterCentroids : cette technique consiste à regrouper les observations de la classe majoritaire en k clusters à l'aide d'un algorithme d'apprentissage non supervisé tel que k-means. Les centroïdes de ces clusters sont ensuite sélectionnés comme échantillons représentatifs de la classe majoritaire, et le reste des observations de la classe majoritaire est supprimé. Cela permet d'obtenir un ensemble de données équilibré avec un nombre réduit d'observations de la classe majoritaire.

AllKNN : cette technique consiste à sélectionner un sous-ensemble des observations de la classe majoritaire qui sont les plus proches des observations de la classe minoritaire, en fonction d'une mesure de distance telle que la distance euclidienne. Les observations de la classe majoritaire qui sont plus éloignées des observations de la classe minoritaire sont supprimées.

NeighbourhoodCleaningRule : cette technique consiste à sélectionner un sous-ensemble des observations de la classe majoritaire qui sont les plus proches des observations de la classe minoritaire, en fonction d'une mesure de distance telle que la distance euclidienne. Les observations de la classe majoritaire qui sont plus éloignées des observations de la classe minoritaire sont supprimées. De plus, les observations de la classe majoritaire qui sont les plus proches des observations de la classe minoritaire sont également supprimées si elles ne sont pas similaires aux observations de la classe minoritaire, en fonction d'une mesure de similarité telle que le coefficient de Jaccard.

RandomUnderSampler : cette technique consiste à sélectionner au hasard un sous-ensemble des observations de la classe majoritaire pour équilibrer la distribution des classes. La taille du sous-ensemble est déterminée par l'équilibre souhaité entre les classes.

2.6. Comparaison des modèles

L'efficacité du modèle DNN combiné aux méthodes d'échantillonnage des données a été évaluée en comparant les métriques de performance du modèle DNN avant et après l'application des techniques de sur-échantillonnage et de sous-échantillonnage. De plus, les performances du modèle DNN ont été comparées à celles d'un modèle d'apprentissage automatique couramment utilisé, la régression logistique, qui est un modèle linéaire simple pouvant servir de référence pour la comparaison avec des modèles plus complexes comme le DNN. Une bonne performance

dans la prédiction de la résiliation des clients a été démontrée dans de nombreux cas par la régression logistique. En comparant les performances du modèle DNN proposé avec un benchmark bien établi, une évaluation plus robuste de son efficacité et de son utilité pratique a été fournie.

De plus, les différentes techniques d'échantillonnage des données utilisées en combinaison avec le modèle DNN ont été comparées pour sélectionner la meilleure technique qui donne la plus haute précision. L'exactitude, la précision et le rappel du modèle DNN ont été évalués en utilisant plusieurs techniques d'échantillonnage des données. L'ensemble de données équilibré résultant a ensuite été utilisé pour entraîner et évaluer le modèle. La comparaison du modèle DNN avec différentes techniques d'échantillonnage des données permet de sélectionner la technique la plus efficace pour améliorer l'exactitude du modèle.

3. Résultats et discussion

Dans cette étude, un réseau neuronal profond (DNN) a été mis en œuvre et évalué sur l'ensemble de données ‘’Churn Modeling’’ en utilisant diverses techniques d'équilibrage des données déséquilibrées pour résoudre le déséquilibre de classe dans les données. Les résultats des prédictions sont résumés dans les tableaux 14 et 15.

Tableau 14. « Évaluation du modèle DNN sans techniques d'échantillonnage des données ».

Model	Accuracy	MSE	Loss
DNN	0.88	0.11	0.37

Tableau 15. « Évaluation du modèle DNN avec différentes techniques d'échantillonnage des données ».

Over/Under -sampling method	Accuracy	MSE	Loss
SMOTE	0.88	0.08	0.31
RandomOverSampler	0.90	0.07	0.30
BorderlineSMOTE	0.88	0.08	0.38
ADASYN	0.89	0.08	0.30
KMeansSMOTE	0.88	0.08	0.34
SVMSMOTE	0.89	0.08	0.34
ClusterCentroids	0.91	0.07	0.28
AllKNN	0.90	0.07	0.31

NeighbourhoodCleaningRule	0.94	0.04	0.35
RandomUnderSampler	0.95	0.04	0.18

Il a été constaté que le DNN combiné à la technique RandomUnderSampler a surpassé les autres techniques d'échantillonnage, avec une exactitude de 0,95, une erreur quadratique moyenne (MSE) de 0,04 et une perte de 0,18. Le DNN seul a atteint une exactitude de 0,88, une MSE de 0,11 et une perte de 0,37. Les techniques de sur-échantillonnage (SMOTE, RandomOverSampler, BorderlineSMOTE, ADASYN, KMeansSMOTE, SVMSMOTE) ont produit des résultats mitigés, certaines atteignant une plus grande exactitude et une plus faible MSE et perte par rapport au DNN seul, tandis que d'autres avaient une exactitude plus faible et une MSE et une perte plus élevées.

De plus, les performances du DNN avec et sans techniques d'équilibrage des données déséquilibrées ont été comparées à un modèle de référence utilisant la régression logistique. Le DNN avec RandomUnderSampler a nettement surpassé le modèle de régression logistique, avec une exactitude de 0,95 contre 0,78 pour le modèle de régression logistique.

En résumé, le DNN combiné aux techniques d'équilibrage des données déséquilibrées améliorait efficacement l'exactitude de la prédiction de la résiliation des clients. L'utilisation de la technique RandomUnderSampler semble en particulier être efficace pour améliorer les performances du DNN sur cet ensemble de données.

Les résultats de cette étude ont des implications importantes pour les entreprises qui comptent sur la fidélisation des clients. Les résultats démontrent que l'utilisation de DNN et de techniques d'équilibrage des données déséquilibrées peut améliorer significativement l'exactitude de la prédiction de la résiliation des clients, ce qui peut orienter le développement de modèles de prédiction de la résiliation des clients plus précis et efficaces.

Une limitation de cette étude est que seul un nombre limité de techniques de sur-échantillonnage et de sous-échantillonnage ont été testées, et des recherches supplémentaires sont nécessaires pour déterminer l'efficacité de ces techniques sur d'autres ensembles de données. Une autre limitation est que l'étude n'a pas étudié l'interprétabilité des modèles DNN, ce qui est une considération importante pour les applications pratiques de ces modèles.

4. Synthèse

Cette étude a démontré l'efficacité de l'utilisation de réseaux neuronaux profonds (DNN) et de techniques d'équilibrage des données déséquilibrées pour prédire le roulement des clients dans l'industrie bancaire. L'utilisation de l'ensemble de données Churn Modeling et de diverses techniques de sur-échantillonnage et de sous-échantillonnage a permis de résoudre le déséquilibre de classe dans les données, ce qui a entraîné une amélioration de l'exactitude de la prédiction de la résiliation des clients.

Le modèle DNN combiné à la technique RandomUnderSampler a obtenu la meilleure exactitude parmi toutes les techniques d'équilibrage des données déséquilibrées, avec une exactitude de 0,95, une MSE de 0,04 et une perte de 0,18. Cela suggère que la technique RandomUnderSampler peut être particulièrement efficace pour résoudre le déséquilibre de classe dans cet ensemble de données et améliorer les performances du DNN.

Les futures directions de recherche pourraient inclure l'exploration de techniques supplémentaires d'équilibrage des données déséquilibrées, telles que des techniques hybrides combinant sur-échantillonnage et sous-échantillonnage, ainsi que la synthèse de données, l'augmentation de données et l'apprentissage sensible aux coûts. De plus, l'étude de différentes architectures de réseaux neuronaux tels que les réseaux neuronaux convolutionnels ou récurrents pourrait conduire à de nouvelles améliorations de l'exactitude.

Cette étude a apporté une contribution précieuse au domaine de la prédiction de la résiliation des clients dans l'industrie bancaire. En démontrant l'efficacité des DNN et des techniques d'équilibrage des données déséquilibrées, les entreprises peuvent améliorer leurs stratégies de fidélisation des clients et, en fin de compte, améliorer leurs résultats financiers.

Partie III. Le machine learning pour les systèmes de recommandation dans le e-commerce

1. Contexte

Aujourd'hui, avec le développement des technologies numériques et la facilité d'accès à Internet, nous sommes de plus en plus exposés à une multitude d'informations. Ce développement apporte beaucoup de diversité aux utilisateurs, mais cette multitude de sources d'informations et cette surcharge d'informations peuvent poser un problème. Afin de remédier à ce problème, divers outils ont été développés pour filtrer les informations avant de les transmettre aux utilisateurs. Les systèmes de recommandation sont des outils qui permettent un tel filtrage. L'objectif principal de ces systèmes est de faciliter la prise de décision pour les utilisateurs en leur proposant des informations selon leurs préférences.

Généralement, il existe deux grandes catégories de systèmes de recommandation. L'une est réalisée avec un filtrage basé sur le contenu et l'autre est réalisée avec un filtrage collaboratif. L'approche basée sur le contenu est mise en œuvre en tenant compte des caractéristiques des produits recommandés et en créant des groupes de produits en utilisant une mesure de similarité sur leur contenu. Le contenu implique une relation entre un mot et un produit. L'un des principaux inconvénients de cette approche est que les caractéristiques sont généralement obtenues à l'aide d'informations externes qui ne sont pas toujours disponibles. D'autre part, le filtrage collaboratif utilise généralement un voisinage d'utilisateurs similaires et recommande des produits en se basant sur l'historique des autres utilisateurs du même voisinage. Lorsque les systèmes émettent leurs recommandations, ils peuvent le faire de manière globale ou locale. "Globalement" signifie que tous les utilisateurs recevront les mêmes recommandations. Alors que "localement" signifie que les articles recommandés ne seront pas les mêmes pour tous les utilisateurs.

Deux approches existent dans la littérature lors du développement d'un système de recommandation utilisant le filtrage collaboratif. La première approche est basée sur la mémoire. Cela signifie que le système émettra ses recommandations en se basant sur un voisinage contenu dans la mémoire. Tandis que la deuxième approche est plutôt basée sur des modèles. Ainsi, il est nécessaire de créer d'abord des modèles qui ressemblent aux comportements des utilisateurs, puis d'utiliser ces modèles pour émettre des recommandations. En pratique, il a été démontré que l'approche basée sur la mémoire offre de meilleures performances en termes de précision,

tandis que l'approche basée sur des modèles est plus efficace à grande échelle avec de grands ensembles de données.

Il existe principalement deux types d'informations utilisées par les systèmes de recommandation, les informations implicites et les informations explicites. Les informations explicites se traduisent par des informations fournies par l'utilisateur. Il est possible de fournir ces informations en utilisant des évaluations, des listes de souhaits et des commentaires. Contrairement aux informations explicites, les informations implicites utilisent uniquement des informations pour lesquelles l'utilisateur n'a pas été sollicité. Par exemple, lorsque l'utilisateur achète un produit, cela crée un historique d'achats pour cet utilisateur. Il est possible pour les systèmes de recommandation basés sur les informations implicites d'utiliser ces informations pour créer différents profils d'utilisateurs.

2. Aperçu des systèmes de recommandation

Il existe différents types d'approches pour réaliser un système de recommandation. Dans la présente section, les approches examinées s'appliquent aux systèmes de recommandation d'informations explicites et implicites. En général, un système de recommandation est un moyen de filtrer les données des utilisateurs "U", un ensemble de produits "P" et une fonction d'utilité "h" qui indique le niveau d'intérêt d'un utilisateur pour un produit donné. Par conséquent, les systèmes de recommandation peuvent être considérés comme une fonction à maximiser, qui peut être écrite sous la forme suivante (16).

$$\forall u \in U, i'_u = arg_{i \in I} \max h(u, i) \tag{16}$$

Où "I" se réfère à un groupe de produits pour "U", et la fonction d'utilité "h" est particulière au type d'approche employé. De même, la fonction d'utilité "h" dépend du type d'information utilisé, qu'elle soit implicite ou explicite. Le groupe de produits "I" joue un rôle crucial dans la formulation d'un système de recommandation, car il représente l'ensemble des articles qui peuvent être recommandés aux utilisateurs. La taille et la composition de ce groupe peuvent avoir un impact significatif sur la précision et la pertinence des recommandations générées par le système. En outre, le choix de la fonction d'utilité "h" est spécifique à l'approche employée et est influencé par le type d'informations utilisées, qu'elles soient implicites ou explicites. Par exemple, la fonction d'utilité pour une approche de filtrage collaboratif est souvent basée sur la similarité entre les utilisateurs ou les éléments, tandis que la fonction d'utilité pour une approche

basée sur le contenu est basée sur les attributs ou les caractéristiques des éléments. La fonction d'utilité doit être choisie avec soin pour garantir que les recommandations générées par le système correspondent aux préférences et aux besoins de l'utilisateur.

2.1. Filtrage collaboratif

Le filtrage collaboratif est le processus qui consiste à utiliser la collaboration des utilisateurs pour formuler une recommandation. Le fonctionnement du filtrage collaboratif est décrit ci-dessous :

Tout d'abord, il recueille des informations telles que des avis ou des comportements d'achat auprès des utilisateurs afin de déterminer leurs centres d'intérêt.

Ensuite, il compare les informations collectées à celles d'autres utilisateurs et repère les similitudes les plus fortes.

Enfin, le système suggère des produits aux utilisateurs cibles similaires.

Le filtrage collaboratif peut être réalisé soit en gardant en mémoire toutes les informations sur les utilisateurs similaires, soit en créant un modèle de préférence à partir de ces informations. Dans le cas d'un système où les informations sont conservées en mémoire, les deux mesures de similarité les plus courantes sont la mesure du cosinus (17) et la corrélation de Pearson (18). La corrélation de Pearson est utile lorsque le système de recommandation utilise des informations explicites, tandis que la mesure du cosinus est utilisée lorsqu'il existe des informations implicites telles que l'historique des achats. Lorsque le filtrage collaboratif est utilisé pour créer un modèle, d'autres techniques telles que la segmentation sont utilisées pour découvrir des informations latentes qui pourraient expliquer l'achat d'un produit ou le résultat d'un devis.

$$\cos(\vec{x}, \vec{y}) = \frac{\vec{x} \,.\, \vec{y}}{||\vec{x}|| * ||\vec{y}||} \tag{17}$$

$$PC(x, y) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{n * stdev(x) * stdev(y)} \tag{18}$$

2.2. Réseau de confiance

Les réseaux de confiance reposent sur des principes similaires à ceux du filtrage collaboratif, mais poussent la similarité entre les utilisateurs un peu plus loin. La recherche sur les réseaux de confiance est motivée par l'hypothèse selon laquelle il existe d'autres facteurs que

les mesures de similarité traditionnelles utilisées dans le filtrage collaboratif, qui peuvent amener un utilisateur à acheter un produit. Cette hypothèse repose sur l'idée que, dans la vie de tous les jours, les gens demandent parfois conseil à des personnes dignes de confiance et discutent d'un produit avec leurs amis. Contrairement au filtrage collaboratif où la similarité est basée sur les profils des individus, l'intérêt principal des réseaux de confiance est la relation de confiance qui lie les utilisateurs entre eux.

Le principal avantage de l'ajout d'un réseau de confiance à un système utilisant le filtrage collaboratif est qu'il permet d'assouplir le problème du démarrage à froid. Avec le réseau de confiance, même si l'utilisateur est nouveau dans le système, il peut toujours recevoir des recommandations s'il est connecté à d'autres individus. Évidemment, cela suppose que l'individu connaisse d'autres utilisateurs qui font partie du système. En outre, lorsqu'un produit n'est pas trouvé dans le réseau immédiat de l'utilisateur, la recherche peut se poursuivre avec un deuxième niveau de confiance, et ainsi de suite. Cependant, même si cette méthode permet d'obtenir une couverture très élevée, la précision diminue au fur et à mesure que la recherche s'éloigne de la source.

Il existe deux grandes catégories de confiance : la confiance interpersonnelle spécifique au contexte et la confiance impersonnelle. Dans le premier cas, l'utilisateur fait confiance à son entourage uniquement dans un contexte spécifique et peut faire confiance à un groupe complètement différent dans un autre contexte. Par exemple, un utilisateur peut faire confiance à son entourage pour aller voir un film, mais il peut aussi faire confiance à un agent immobilier pour acheter une maison. La seconde indique une confiance systématique qui décrit la confiance d'un utilisateur dans son ensemble au sein d'un groupe. Par exemple, deux personnes se font confiance parce qu'elles font partie du même groupe.

2.3. Modèle latent

Contrairement aux approches qui utilisent un réseau de confiance ou le filtrage collaboratif, les systèmes de recommandation qui utilisent des modèles latents ne gardent pas les informations de l'utilisateur en mémoire pendant la recommandation. Au contraire, ils utilisent ces informations pour créer des modèles à l'aide de techniques d'exploration de données. Pour créer un modèle, le système entraîne d'abord son modèle avec des paramètres aléatoires, puis affine les paramètres à l'aide d'une ou plusieurs fonctions de coût. Contrairement à l'approche

en mémoire, l'approche par modèle latent n'essaiera pas nécessairement de trouver des liens directs entre les utilisateurs et les produits. Elle tente plutôt de découvrir des liens latents qui inciteront l'utilisateur à acheter un produit. En suivant cette idée, il est alors possible de saisir des faits qui pourraient motiver certains utilisateurs à préférer certains produits sans que cela soit évident pour un système traditionnel. Cependant, il est difficile avec cette approche d'expliquer le raisonnement qui sous-tend une recommandation, car les lignes directrices n'ont pas d'étiquettes correctement identifiées. En d'autres termes, un facteur latent peut être un mélange de différents facteurs réguliers.

Le développement de modèles latents dans le domaine des systèmes de recommandation fait souvent appel à la technique de factorisation matricielle. Cette technique permet d'associer les utilisateurs et les produits dans un nouvel espace où les relations entre les utilisateurs et les produits peuvent être définies à l'aide du produit de points. La matrice de l'espace initial est généralement constituée d'utilisateurs pour les lignes et de produits pour les colonnes. Ainsi, pour calculer un score estimé pour un produit, il faut prendre le vecteur de l'utilisateur et le vecteur du produit dans l'espace latent, puis calculer le produit point entre ces deux vecteurs. Cependant, comme les ensembles de données pour les systèmes de recommandation sont souvent dispersés, il est important de veiller à ne pas ajouter trop de bruit au nouvel espace.

2.4. Règles d'association

Les règles d'association sont généralement utilisées pour découvrir des schémas répétitifs dans des ensembles de données qui peuvent être très volumineux. Comme les réseaux de confiance, les règles d'association sont souvent combinées à une approche de filtrage collaboratif pour rendre cette dernière encore plus précise. En général, les règles d'association suivent un modèle probabiliste de cause à effet. En d'autres termes, le système analyse une multitude d'événements et regroupe ensuite les instances qui précèdent un fait.

Les systèmes de recommandation basés sur des règles d'association utilisent généralement l'historique des transactions pour extraire les règles qui permettront aux systèmes de faire des associations. Toutefois, ces méthodes aboutissent souvent à un très grand nombre de règles, ce qui rend inefficace la connaissance obtenue à partir de ces règles. En outre, les méthodes basées sur les règles d'association nécessitent des paramètres explicites tels que le nombre de règles et le niveau de confiance associé à ces règles. Dans des travaux antérieurs, un système de

recommandation a été créé sur la base de règles d'association, qui sont produites à l'aide d'un modèle qui effectue un contrôle de qualité sur ces règles dès le moment de leur création. Fondamentalement, il existe deux types d'associations : les associations entre produits et les associations entre utilisateurs. Par exemple, une association entre produits peut être que les utilisateurs qui aiment le produit A et le produit B aiment également le produit C. Quant à l'association entre utilisateurs, elle peut être que les articles aimés par l'utilisateur A et l'utilisateur B sont également aimés par l'utilisateur C. Cette pratique nous permet de trouver des inférences entre les règles et l'analyse du chevauchement entre les utilisateurs nous permet de faire des recommandations aux utilisateurs qui n'ont pas beaucoup de similitudes directes entre eux.

Bien que l'utilisation d'un modèle pour créer des règles d'association augmente la qualité des règles d'association, le problème reste que le nombre de règles produites est toujours élevé. Par conséquent, le bruit causé par un très grand nombre de règles réduit considérablement la confiance dans les recommandations. Une règle d'association peut être définie comme une table de vérité résultant de la combinaison de deux caractéristiques ou plus. Les règles d'association sont un ensemble de déclarations "si-alors" conçues pour fournir une probabilité de relations entre les caractéristiques des données, dans de grands ensembles de données de divers types de bases de données. L'extraction de règles d'association offre une variété de cas d'utilisation et est largement utilisée pour aider à découvrir les corrélations entre les ventes dans les ensembles de données transactionnelles. Les règles d'association sont basées sur les trois paramètres suivants :

- Support : mesure du pourcentage d'éléments, qu'il s'agisse d'un seul élément ou d'une combinaison d'éléments, qui apparaissent dans l'ensemble de données. La valeur de soutien est une mesure de la fréquence à laquelle l'ensemble d'éléments apparaît dans une association.
- Confiance : mesure de la précision de la probabilité d'occurrence d'un élément par rapport à un autre. La valeur de confiance est comprise entre zéro et un. Plus la valeur de confiance est proche de un, plus les deux éléments sont susceptibles d'apparaître ensemble.
- Lift : c'est le rapport entre le soutien observé et le soutien qui serait attendu si X et Y étaient indépendants. Le lift indique la valeur de la règle basée sur des événements aléatoires X et Y.

Les règles d'association sont utiles pour analyser et prédire le comportement des clients. Elles sont particulièrement importantes pour l'analyse des clients, l'analyse des paniers d'achat, le regroupement des produits et des clients, la conception de catalogues et la gestion des magasins. Les algorithmes qui utilisent les règles d'association comprennent AIS (Artificial Immune Systems), SETM (Self-Exciting Temporal Point Process Model) et Apriori.

- Algorithme Apriori

L'algorithme Apriori utilise des ensembles d'éléments fréquents pour élaborer des règles d'association. L'algorithme est conçu pour fonctionner sur des ensembles de données transactionnelles. Sur la base des règles d'association, il peut évaluer la force ou la faiblesse du lien entre deux objets. Grâce à une recherche de type "breadth-first" et à un arbre de hachage, l'algorithme peut calculer avec succès les associations d'ensembles d'éléments. Le processus est itératif afin de trouver des ensembles fréquents à partir d'un grand ensemble de données.

- Algorithme FP-Growth

L'algorithme FP-Growth (Frequent Pattern Growth) est une version améliorée de l'algorithme Apriori. Il s'agit d'une méthode alternative pour trouver des ensembles d'éléments fréquents sans utiliser de générations de candidats, en employant une stratégie de division et de conquête. Cette approche efficace et évolutive repose sur l'utilisation d'une structure d'arbre préfixe étendue pour stocker des informations compressées et cruciales sur les motifs fréquents, appelée arbre de motifs fréquents (FP-tree), qui préserve les informations d'association des ensembles d'éléments.

3. Méthodologie

Pour résoudre le problème de la prise de décision des clients et augmenter les ventes, il est essentiel de mettre en œuvre un système de recommandation de produits précis et efficace. Dans le présent document, un système de recommandation basé sur des règles d'association, via l'algorithme FP-Growth, est adopté parce qu'il offre une grande précision tout en étant facile à mettre en œuvre et à expliquer. Ainsi, les résultats de ce modèle sont compréhensibles : toutes les règles spécifiques à l'activité peuvent être lues dans la table d'association. La figure 13 montre les étapes de la méthodologie adoptée.

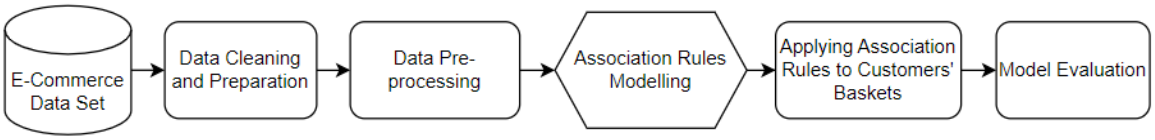


Figure 13. « Processus du système de recommandation proposé ».

3.1. Visualisation des données

L'ensemble de données utilisé est "Online Retail" du Machine Learning Repository. Il s'agit d'un ensemble de données transactionnelles contenant 406 829 transactions pour une société de commerce électronique enregistrée basée au Royaume-Uni qui vend des cadeaux uniques. L'ensemble de données comprend des informations sur l'identification du client, le numéro de facture, le code de stock, la description du produit, la quantité, la date de facturation, le prix unitaire, le prix total de la facture et le pays. Le tableau 16 montre les 10 premières lignes du dataset utilisé.

Tableau 16. « Le dataset transactionnel utilisé pour la recommandation ».

```
(406829, 9)
```

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country	TotalPrice
0	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	12/1/2010 8:26	2.55	17850.0	United Kingdom	15.30
1	536365	71053	WHITE METAL LANTERN	6	12/1/2010 8:26	3.39	17850.0	United Kingdom	20.34
2	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	12/1/2010 8:26	2.75	17850.0	United Kingdom	22.00
3	536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	12/1/2010 8:26	3.39	17850.0	United Kingdom	20.34
4	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	12/1/2010 8:26	3.39	17850.0	United Kingdom	20.34
5	536365	22752	SET 7 BABUSHKA NESTING BOXES	2	12/1/2010 8:26	7.65	17850.0	United Kingdom	15.30
6	536365	21730	GLASS STAR FROSTED T-LIGHT HOLDER	6	12/1/2010 8:26	4.25	17850.0	United Kingdom	25.50
7	536366	22633	HAND WARMER UNION JACK	6	12/1/2010 8:28	1.85	17850.0	United Kingdom	11.10
8	536366	22632	HAND WARMER RED POLKA DOT	6	12/1/2010 8:28	1.85	17850.0	United Kingdom	11.10
9	536367	84879	ASSORTED COLOUR BIRD ORNAMENT	32	12/1/2010 8:34	1.69	13047.0	United Kingdom	54.08

3.2. Prétraitement des données

Tout d'abord, nous n'avons conservé que les produits que le client a effectivement mis dans son panier en supprimant les produits qui correspondent aux cadeaux de l'entreprise aux clients. Ensuite, nous avons regroupé tous les produits qu'un client a achetés dans un nouvel ensemble de données. Chaque ligne correspond à une transaction comprenant le numéro de facture, l'identifiant du client et tous les produits achetés (tableau 17).

Tableau 17. « L'ensemble des données relatives à tous les produits achetés par chaque client ».

InvoiceNo	CustomerID	StockCode
536365	17850.0	[84406B, 85123A, 21730, 22752, 84029G, 84029E,...
536366	17850.0	[22632, 22633]
536367	13047.0	[21777, 84969, 22748, 22749, 84879, 22623, 226...
536368	13047.0	[22914, 22912, 22913, 22960]
536369	13047.0	[21756]

3.3. Modélisation des règles d'association

Pour la détermination des règles d'association, nous avons utilisé le modèle FP-Growth qui permet, à partir d'un historique de transactions, de déterminer l'ensemble des règles d'association les plus fréquentes dans l'ensemble de données. Pour ce faire, il a besoin comme paramètre d'entrée de l'ensemble des transactions composé des paniers de produits que les clients ont déjà achetés. Pour établir la table d'associations, il est nécessaire de définir 2 hyperparamètres :

- minSupRatio : Il représente le support minimum pour qu'un ensemble d'éléments soit identifié comme fréquent.
- minConf : Il représente la confiance minimale pour générer une règle d'association.

Le système a généré 4 970 règles d'association (tableau 18).

Tableau 18. « Ensemble de données sur les règles d'association générées par le modèle ».

	basket	next_product	p
199	{22916, 22921, 22920, 22917, 22919}	{22918}	0.992537
424	{22916, 22921, 22919, 22917}	{22918}	0.986014
306	{22920, 22921, 22918, 22917}	{22916}	0.985714
310	{22920, 22921, 22916, 22917}	{22918}	0.985714
96	{22920, 22921, 22919, 22917}	{22918}	0.985401
231	{22920, 22916, 22921, 22919}	{22918}	0.985401
202	{22918, 22921, 22920, 22917, 22919}	{22916}	0.985185
201	{22916, 22921, 22918, 22920, 22919}	{22917}	0.985185
366	{22921, 22916, 22917}	{22918}	0.979866
458	{22921, 22919, 22917}	{22918}	0.979730

Une fois les règles d'association établies à l'aide du modèle FP-Growth, elles ont été appliquées aux paniers de produits des clients. Cette opération a permis de recommander des produits aux clients sur la base de leur historique d'achat. Les prix estimés ont ensuite été calculés à partir des produits recommandés, comme le montre le tableau 19.

Tableau 19. « Ensemble de données de règles d'association appliquées aux paniers des clients ».

InvoiceNo	CustomerID	Customer basket	Recommended Product	Product description	Probability	Price estimation
536365	**17850.0**	[85123A, 84406B, 84029G, 84029E, 22752, 21730,...	0	Null	0.000000	0.000000
536366	**17850.0**	[22633, 22632]	22865	HAND WARMER OWL DESIGN	0.470982	0.871317
536367	**13047.0**	[22310, 84969, 48187, 21754, 22623, 22749, 217...	22750	FELTCRAFT PRINCESS LOLA DOLL	0.593516	2.225686
536368	**13047.0**	[22912, 22960, 22913, 22914]	22961	JAM MAKING SET PRINTED	0.322280	0.467306
536369	**13047.0**	[21756]	21754	HOME BUILDING BLOCK WORD	0.576132	3.427984
536370	**12583.0**	[22631, 22629, 22726, 22661, 21913, 21724, 218...	22630	DOLLY GIRL LUNCH BOX	0.595506	1.161236
536371	**13748.0**	[22086]	22910	PAPER CHAIN KIT VINTAGE CHRISTMAS	0.458586	1.169394
536372	**17850.0**	[22633, 22632]	22865	HAND WARMER OWL DESIGN	0.470982	0.871317
536373	**17850.0**	[85123A, 84406B, 20679, 21871, 37370, 82482, 2...	15056BL	EDWARDIAN PARASOL BLACK	0.509158	3.029487

4. Évaluation du modèle et résultats

L'évaluation d'un système de recommandation est une étape cruciale pour déterminer son efficacité à améliorer les ventes et la satisfaction des clients. Dans cette étude, la performance du système de recommandation a été évaluée en calculant la probabilité moyenne du prochain produit que le client achètera ($P_{average}$). En outre, les revenus attendus des produits prédits et suggérés ont également été calculés et présentés dans le tableau 20 ci-dessous. Cette évaluation a permis d'estimer l'augmentation potentielle des revenus qui pourrait résulter de la mise en œuvre du système de recommandation proposé.

Tableau 20. « Évaluation du système de recommandation ».

Nombre de suggestions pour chaque client	$P_{average}$ (en %)	Chiffre d'affaires généré (en $)
1	69.3374553	36,891.891891

5. Synthèse

À l'ère de l'information, la surcharge de données est devenue un obstacle à la prise des bonnes décisions au bon moment. En particulier pour les cyberconsommateurs, déterminer avec précision et rapidité les informations qui les intéressent dans ce chaos permanent est un défi majeur pour toute organisation de commerce électronique. C'est pourquoi les systèmes de recommandation sont appelés à résoudre ce problème et à aider les utilisateurs à prendre les bonnes décisions. Dans cette optique, l'objectif de notre étude était d'implémenter un système de recommandation basé sur des règles d'association via l'algorithme FP-Growth. La limite de ce travail est que certaines des caractéristiques d'évaluation d'un système de recommandation, telles que la diversité et l'explicabilité, sont difficiles à définir. L'évaluation en situation réelle, y compris les tests A/B et les tests par échantillonnage, est une technique puissante pour évaluer un nouveau système de recommandation. En effet, dans le prolongement de cette étude, nous mettrons en place des systèmes de recommandation basés sur différents modèles d'apprentissage automatique afin d'effectuer des tests A/B et de définir le modèle le plus performant.

Partie IV. Se défendre contre le vol numérique : une approche machine learning pour prédire la fraude au paiement électronique

1. Contexte

Le e-commerce est devenu omniprésent au quotidien, englobant l'achat et la vente de produits ou de services en ligne. Tout, des livres aux produits alimentaires, en passant par les vêtements et même les voitures, peut être acheté en ligne. La signification du commerce électronique réside dans sa capacité à gagner du temps et de l'argent. En utilisant des plateformes de commerce électronique, il n'est pas nécessaire de se rendre dans un magasin pour acheter ce dont nous avons besoin physiquement. Au lieu de cela, nous pouvons le chercher sur nos smartphones ou ordinateurs portables et passer commande immédiatement, ce qui rend les achats plus rapides et plus pratiques.

Selon les prévisions, 95 % de tous les achats se feront en ligne d'ici 2040 (Statista, 2022). Avec les appareils mobiles et l'Internet des objets (IoT) facilitant de plus en plus l'achat à la demande, les consommateurs trouvent la possibilité de faire leurs achats en ligne de manière transparente. Cette tendance devrait contribuer à l'augmentation du volume des ventes de commerce électronique dans le monde, qui devrait atteindre 4,9 milliards de dollars en 2021 et près de 7,3 milliards de dollars d'ici 2025, comme illustré dans la figure suivante (figure 14).

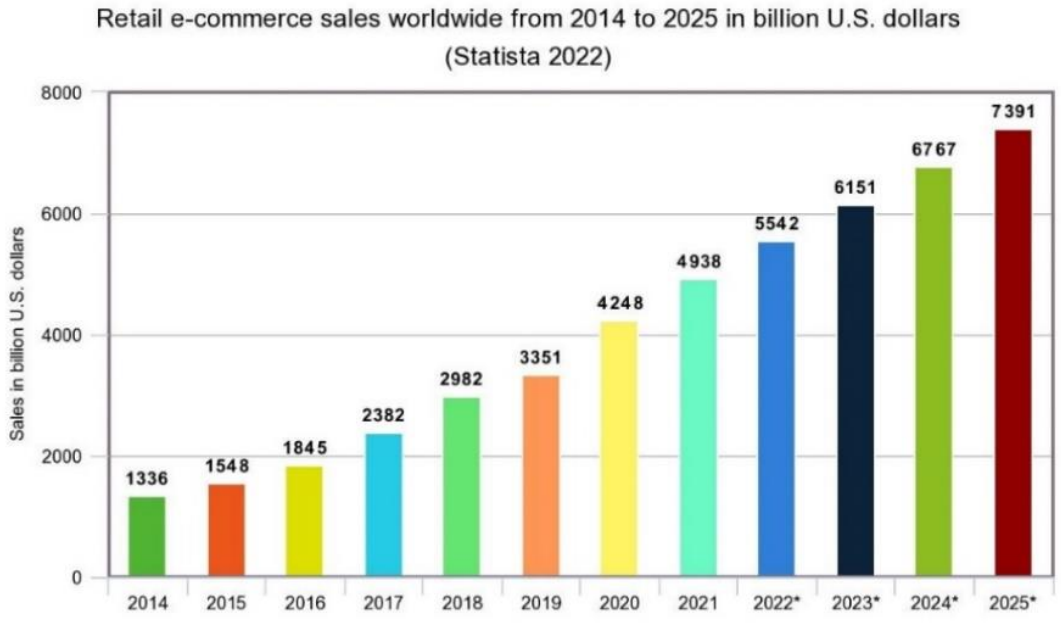


Figure 14. « Ventes au détail du commerce électronique dans le monde ».

Les cartes de crédit sont le moyen de paiement le plus largement utilisé pour effectuer des achats via des systèmes de paiement électronique dans le commerce électronique. Les systèmes de paiement électronique impliquent le transfert électronique ou numérique de fonds, représenté par l'échange de données. Les cartes de crédit sont devenues la norme pour les paiements sur Internet, et des millions de transactions sont traitées chaque jour. L'architecture des systèmes de paiement électronique est illustrée dans la figure 15 ci-dessous.

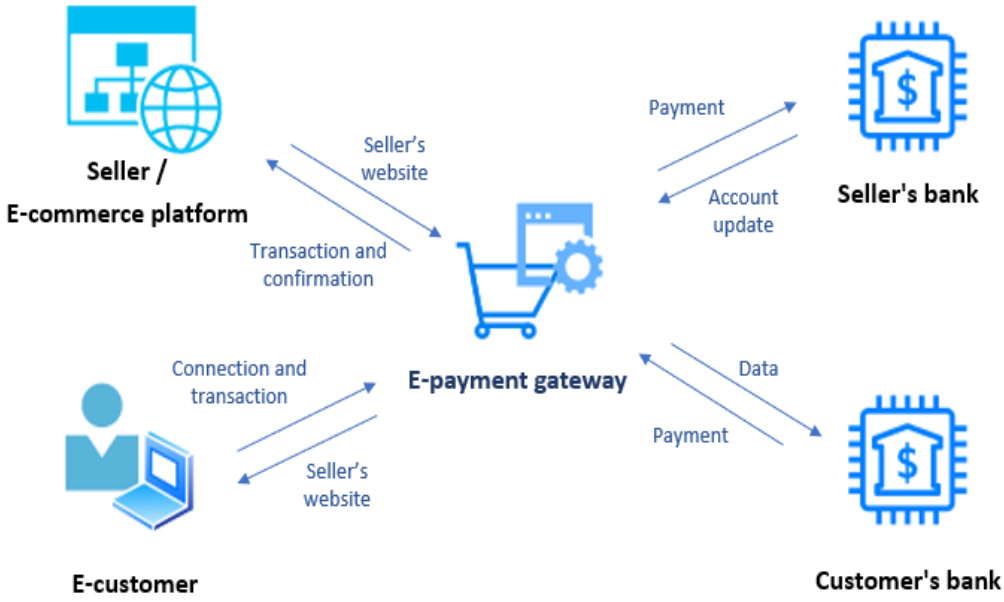


Figure 15. « Architecture du système de paiement électronique ».

Une passerelle de paiement électronique, ou processeur de cartes de crédit, établit le lien entre le site web du vendeur et le panier d'achat, la banque qui reçoit la transaction (celle du vendeur) et la banque qui émet la transaction (celle du client). La passerelle de paiement électronique gère chaque échange entre ces parties. En traitant les parties essentielles du traitement des cartes de crédit, notamment l'autorisation et le règlement des paiements, la passerelle de paiement électronique est un composant crucial des transactions de paiement électronique.

Les systèmes de paiement électronique ont permis de gagner du temps, permettant ainsi de faire des achats d'un simple clic. Ainsi, les transactions sont traitées et finalisées en une fraction du temps. De plus, ils ont permis une efficacité d'utilisation ; avec les systèmes de paiement électronique, le client évite de faire de longues queues aux distributeurs automatiques de billets ou aux guichets bancaires pour retirer de l'argent ou de visiter physiquement les

magasins. De plus, les transactions en espèces sont associées à des risques tels que le vol, la perte ou d'autres incidents similaires. Cependant, les systèmes de paiement électronique sont équipés de protocoles de sécurité qui garantissent la sécurité des fonds.

La croissance rapide du commerce électronique a considérablement augmenté l'utilisation des systèmes de paiement électronique, en particulier des cartes de crédit, dans les transactions en ligne. Cependant, cette augmentation des paiements électroniques a également rendu les cartes de crédit attrayantes pour les fraudeurs. Les données sur la fraude par carte de crédit indiquent qu'elle est devenue plus répandue et diversifiée avec l'essor des achats en ligne. Les systèmes de détection de fraude traditionnels peinent à suivre les méthodes en constante évolution utilisées par les fraudeurs, ce qui rend impératif de trouver des approches plus efficaces. Par conséquent, le paysage mondial de la cybercriminalité connaît une escalade préoccupante chaque jour.

La fraude par carte de crédit peut être largement classée en deux types : la fraude hors ligne et la fraude en ligne. La fraude hors ligne implique généralement des actions physiques, telles que le vol d'un sac à main ou d'un portefeuille contenant des éléments de valeur tels que des cartes de paiement et des documents d'identification, et l'utilisation des informations volées à des fins lucratives. D'autre part, la fraude en ligne se produit lorsque les fraudeurs exploitent des plateformes en ligne ou créent des sites web trompeurs pour collecter des données personnelles sensibles et effectuer des transactions non autorisées sur les comptes des clients. Les fraudeurs utilisent diverses techniques pour collecter ou voler des données personnelles, notamment le phishing, le piratage, la fraude d'identité, l'épaule surfeuse, les logiciels espions, le fouillage des poubelles, et plus encore.

Les solutions de détection de fraude traditionnelles se sont révélées insuffisantes pour détecter la fraude par carte de crédit dans les environnements de commerce électronique diversifiés. Par conséquent, des techniques d'apprentissage automatique et d'intelligence artificielle ont été mises en œuvre pour améliorer la précision de la détection de fraude. Les méthodes d'apprentissage automatique permettent aux experts en détection de fraude de lutter contre les activités frauduleuses.

Le but de cette étude est de mettre en œuvre un système d'apprentissage automatique pour détecter les paiements électroniques frauduleux basé sur le modèle AdaBoost. Les

performances du modèle AdaBoost ont été comparées à celles d'autres modèles d'apprentissage automatique pour évaluer son efficacité dans la détection des paiements électroniques frauduleux. Une évaluation complète de ses performances a été réalisée en comparant les résultats obtenus à partir du modèle AdaBoost avec ceux de différents modèles, tels que CatBoost et XGBoost. Cette comparaison a permis d'obtenir des informations sur les forces et les faiblesses relatives de chaque modèle dans la détection précise et la prévention de la fraude en ligne. En tenant compte des mesures de performance et en les comparant entre différents modèles, les organisations peuvent prendre des décisions éclairées concernant l'adoption de l'approche d'apprentissage automatique la plus adaptée à leurs besoins spécifiques de détection de fraude.

2. Fondements de l'apprentissage automatique pour la détection de la fraude

2.1. Systèmes de détection de fraude basés sur l'apprentissage automatique

La fraude est en constante évolution et changement. Elle a considérablement augmenté avec la croissance du commerce électronique. Les fraudeurs exploitent toutes les vulnérabilités détectées dans les plateformes de banque en ligne ou de commerce électronique. En effet, la fraude est devenue une préoccupation importante pour les commerçants du commerce électronique et de la banque en ligne dans le monde entier. La lutte contre la fraude nécessite la mise en place d'un système anti-fraude efficace, qui comprend des mesures pour détecter et évaluer les transactions de paiement en ligne pour d'éventuelles fraudes. Ces systèmes se concentrent sur des paramètres spécifiques de chaque transaction pour identifier une activité suspecte. Lorsque la transaction ne répond pas aux exigences définies, elle est examinée et des anomalies par rapport à la norme sont signalées.

L'apprentissage automatique est un cas d'utilisation majeur pour la détection et la prévention des paiements frauduleux. Il peut détecter la fraude parmi les transactions légitimes.

Le principe sous-jacent à l'utilisation de l'apprentissage automatique pour la détection de la fraude est que les transactions frauduleuses incluent des informations qui ne correspondent pas aux transactions légitimes. Avec l'apprentissage automatique, il est possible de détecter plus efficacement et rapidement de grands flux de fraudes par rapport aux êtres humains. Les algorithmes d'apprentissage automatique peuvent traiter de grandes quantités d'informations de manière plus précise et rapide.

La principale faiblesse des systèmes de détection de fraude traditionnels basés sur des systèmes statiques et basés sur des règles est qu'ils dépendent fortement de l'intervention humaine. Ainsi, le travail des analystes de haut niveau nécessite beaucoup de temps et de ressources. De plus, ces experts peuvent ne pas avoir les compétences, les connaissances, la rapidité et l'expérience des algorithmes de détection de fraude qui utilisent des approches d'apprentissage automatique.

2.2. Fonctionnement des systèmes de détection de fraude basés sur l'apprentissage automatique

Un modèle de détection de fraude par apprentissage automatique doit d'abord collecter, analyser, segmenter et extraire des données pertinentes pour identifier la fraude. Contrairement aux humains, qui ont besoin de plus de temps et d'efforts pour comprendre de grandes quantités d'informations, plus les données à analyser pour le modèle d'apprentissage automatique sont importantes, meilleurs sont les résultats. Les données sont le moyen d'accélérer la courbe d'apprentissage d'un système d'apprentissage automatique.

La prochaine étape consiste en l'extraction de caractéristiques. Elle consiste à ajouter des caractéristiques contenant des informations sur le client (localisation, identité, commandes, réseau et mode de paiement choisi, etc.). Ces caractéristiques peuvent varier en fonction de la complexité du système.

Ensuite, le modèle d'apprentissage automatique reçoit des ensembles d'entraînement qui lui apprennent à identifier et à prédire les transactions frauduleuses. Plus les données fournies pour un ensemble d'entraînement sont nombreuses, meilleures seront les performances du modèle d'apprentissage automatique.

2.3. Algorithmes d'apprentissage automatique pour la prédiction de paiements frauduleux

Plusieurs classificateurs d'apprentissage automatique peuvent être utilisés pour la prédiction de paiements frauduleux dans le commerce électronique. Certains des plus couramment utilisés sont :

- La régression logistique : cet algorithme est utile pour les problèmes de classification binaire comme la détection de fraude. Il peut aider à identifier les transactions frauduleuses en fonction des caractéristiques fournies en entrée.

- Les arbres de décision : cet algorithme peut être utilisé pour créer un arbre de décision basé sur les caractéristiques fournies en entrée, et l'arbre peut aider à classer les transactions comme frauduleuses ou légitimes.
- Les forêts aléatoires : cet algorithme peut être utilisé pour créer plusieurs arbres de décision, puis combiner leurs prédictions pour obtenir un résultat plus précis. Il est particulièrement utile lorsqu'il y a de nombreuses caractéristiques d'entrée.
- Les machines à vecteurs de support (SVM) : cet algorithme peut être utilisé pour séparer les données en différentes classes à l'aide d'un hyperplan. Il peut aider à identifier les transactions frauduleuses en créant une frontière entre les transactions frauduleuses et légitimes.
- Les réseaux neuronaux : cet algorithme peut entraîner un modèle sur un grand ensemble de données de transactions et identifier les schémas associés à la fraude. Il peut être particulièrement efficace lorsqu'il s'agit de données complexes et non linéaires.

En plus des algorithmes ci-dessus, des algorithmes d'apprentissage d'ensemble tels que Boosting peuvent également être utilisés pour la prédiction des paiements frauduleux du commerce électronique. Les algorithmes de boosting, tels que AdaBoost et Gradient Boosting, combinent plusieurs modèles plus faibles pour créer un modèle plus fort et plus précis. Pour ce faire, l'algorithme entraîne plusieurs modèles sur des sous-ensembles de données et attribue des pondérations à chaque modèle en fonction de leurs performances. Ensuite, les modèles sont combinés pour créer un modèle final qui peut faire des prédictions plus précises. Cette technique est particulièrement utile pour gérer des ensembles de données complexes et de grande dimension, tels que ceux que l'on trouve dans les transactions de commerce électronique.

Dans cette étude, des algorithmes de boosting non supervisés, à savoir : CatBoost, AdaBoost et XGBoost sont implémentés.

AdaBoost, abréviation de Adaptive Boosting, est un algorithme d'apprentissage d'ensemble couramment utilisé dans l'apprentissage automatique. Cela fonctionne en tirant parti des instances mal classées du modèle précédent et en générant un nouvel échantillon de données pondéré où les instances mal classées reçoivent plus de poids. Cela signifie que l'algorithme s'adapte aux erreurs commises par le modèle précédent et tente de l'améliorer. Le nouveau

modèle généré est ensuite ajouté à l'ensemble, et le processus est répété jusqu'à ce que le niveau de précision souhaité soit atteint.

XGBoost, ou Extreme Gradient Boosting, est un algorithme qui vise à améliorer la technique de gradient boosting en augmentant les ressources de calcul. Il s'agit d'une version avancée de l'amplification de gradient qui a deux caractéristiques principales. Premièrement, il améliore le surajustement grâce à la régularisation, empêchant le modèle de s'adapter trop étroitement aux données d'apprentissage. Deuxièmement, il améliore la vitesse d'exécution en optimisant l'ordonnancement à l'aide d'une exécution parallèle. L'algorithme utilise également la profondeur maximale de l'arbre de décision comme paramètre pour élaguer l'arbre, ce qui permet de réduire considérablement le temps d'exécution.

CatBoost est un algorithme d'apprentissage automatique conçu pour gérer efficacement les données catégorielles. Il est capable de traiter des variables catégorielles sans avoir besoin de les convertir sous forme numérique, ce qui peut être un processus chronophage dans certains cas. Cet algorithme est particulièrement utile lorsqu'il s'agit d'ensembles de données contenant un grand nombre de variables catégorielles. CatBoost utilise un algorithme avancé d'amplification de gradient qui crée une séquence d'arbres de décision, chaque arbre suivant apprenant des erreurs du précédent. Cette approche peut aider à améliorer la précision du modèle et à réduire le risque de surajustement. De plus, CatBoost peut gérer les valeurs manquantes dans l'ensemble de données et choisir automatiquement la meilleure valeur pour les remplir.

3. Méthodologie

Cette étude vise à prédire si une transaction de paiement électronique est frauduleuse ou légitime. Pour ce faire, le classificateur d'apprentissage automatique AdaBoost a été implémenté et comparé à d'autres algorithmes de renforcement de l'apprentissage automatique, à savoir CatBoost et XGBoost, simplifiant la détection des paiements frauduleux en le transformant en un problème de classification binaire. Les méthodes de boosting ont été choisies car elles impliquent une collection d'algorithmes basés sur des ensembles de classificateurs binaires, pour optimiser leurs performances. Pour atteindre notre objectif, nous avons suivi les étapes décrites dans la figure 16.

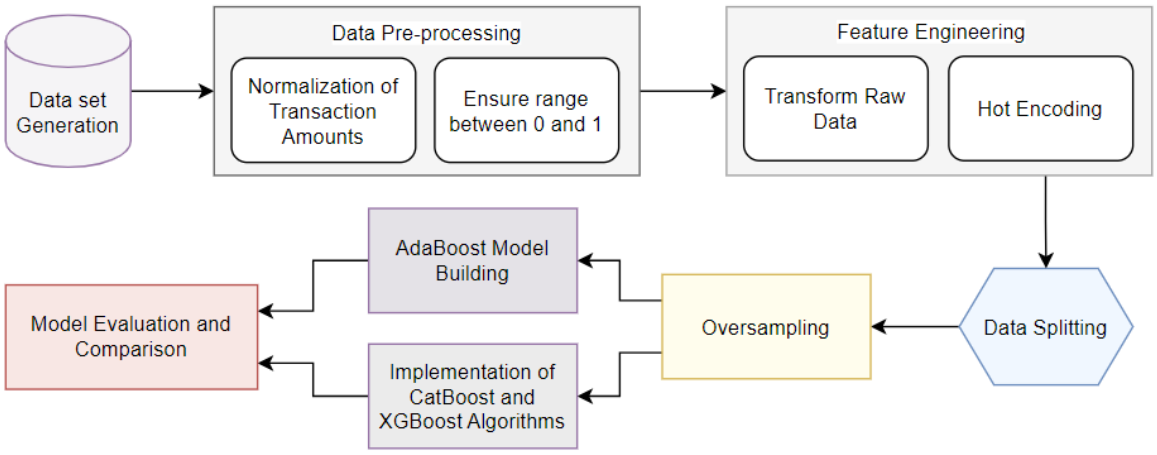


Figure 16. « Flowchart de l'étude de prédiction de paiements frauduleux ».

3.1. Base de données

L'ensemble de données utilisé dans cette étude a été généré par un simulateur de paiement bancaire basé sur un échantillon de données transactionnelles agrégées fournies par une banque. Le simulateur a été conçu pour produire des données artificielles qui pourraient être utilisées pour la recherche sur la détection des fraudes. Il a duré 180 étapes, correspondant à six mois de transactions, pour générer une distribution qui pourrait être considérée comme fiable à des fins de test. L'ensemble de données résultant comprend 594 643 transactions, dont 587 443 sont des paiements légitimes et 7 200 sont des transactions frauduleuses, comme le montre le tableau 21.

Tableau 21. « Description du dataset relatif à la détection des payements frauduleux ».

	step	customer	age	gender	zipcodeOri	merchant	zipMerchant	category	amount	fraud
0	0	'C1093826151'	'4'	'M'	'28007'	'M348934600'	'28007'	'es_transportation'	4.55	0
1	0	'C352968107'	'2'	'M'	'28007'	'M348934600'	'28007'	'es_transportation'	39.68	0
2	0	'C2054744914'	'4'	'F'	'28007'	'M1823072687'	'28007'	'es_transportation'	26.89	0
3	0	'C1760612790'	'3'	'M'	'28007'	'M348934600'	'28007'	'es_transportation'	17.25	0
4	0	'C757503768'	'5'	'M'	'28007'	'M348934600'	'28007'	'es_transportation'	35.72	0
...	...	...	...	...	...	...	...	...	...	...
594638	179	'C1753498738'	'3'	'F'	'28007'	'M1823072687'	'28007'	'es_transportation'	20.53	0
594639	179	'C650108285'	'4'	'F'	'28007'	'M1823072687'	'28007'	'es_transportation'	50.73	0
594640	179	'C123623130'	'2'	'F'	'28007'	'M349281107'	'28007'	'es_fashion'	22.44	0
594641	179	'C1499363341'	'5'	'M'	'28007'	'M1823072687'	'28007'	'es_transportation'	14.46	0
594642	179	'C616528518'	'4'	'F'	'28007'	'M1823072687'	'28007'	'es_transportation'	26.93	0

594643 rows × 10 columns

3.2. Prétraitement des données

L'objectif de la création d'un modèle de classification des fraudes est d'attribuer une probabilité élevée à chaque transaction entrante, permettant l'identification des transactions potentiellement frauduleuses et empêchant les tentatives illégales. Ainsi, des techniques de prétraitement des données ont été utilisées pour assurer l'exactitude et la fiabilité du modèle.

Les transactions avec des valeurs plus élevées sont plus vulnérables à la fraude, et comparer les valeurs des transactions entre différents clients n'est pas approprié puisque les montants des transactions varient entre les clients. Pour résoudre ce problème, les valeurs de transaction de chaque client ont été normalisées en divisant le montant de la transaction par le montant maximal de la transaction pour ce même client. Ce processus de normalisation a permis au modèle d'identifier un comportement inhabituel pour chaque transaction individuelle. Avec 4 112 clients et un nombre variable de transactions pour chaque client dans l'ensemble de données, la normalisation des montants des transactions de cette manière garantissait que chaque montant de transaction était compris entre 0 et 1, comme illustré à la figure 17.

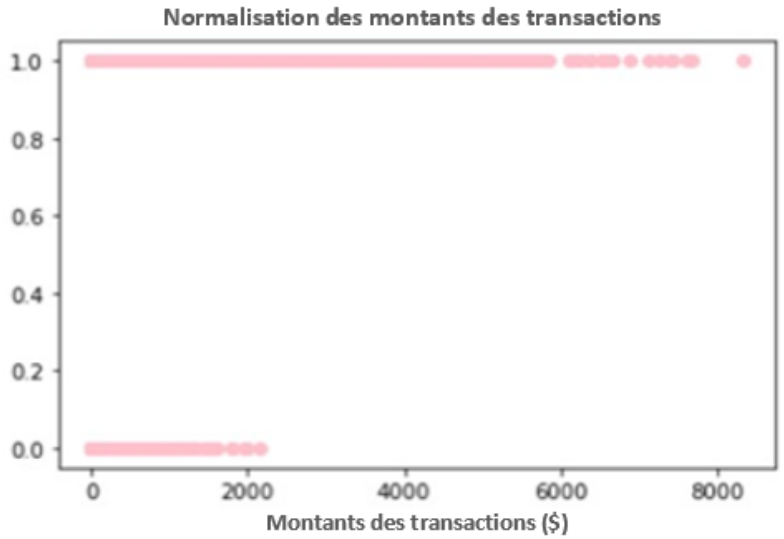


Figure 17. « Normalisation des montants des transactions ».

3.3. Ingénierie des fonctionnalités

L'ingénierie des fonctionnalités est une étape cruciale dans le pré-traitement d'un problème de modélisation. Il vise à transformer les données brutes d'origine en dimensions plus précisément représentées par un modèle prédictif, améliorant ainsi considérablement les mesures d'évaluation. Un autre aspect important de l'ingénierie des fonctionnalités consiste à garantir qu'aucune des fonctionnalités utilisées pour former le modèle ne sera à nouveau

disponible lorsqu'un nouveau cas se présentera. Cela permet d'éviter les biais et d'améliorer la capacité du modèle à généraliser à de nouveaux cas.

3.4. Encodage à chaud (one-hot)

L'encodage à chaud a été utilisé comme étape de prétraitement pour les caractéristiques catégorielles dans les modèles d'apprentissage automatique. Cette méthode est largement utilisée en apprentissage automatique pour encoder des caractéristiques catégorielles. L'encodage à chaud crée une nouvelle caractéristique binaire pour chaque catégorie possible et attribue une valeur de 1 à la caractéristique de chaque échantillon qui correspond à sa catégorie d'origine.

3.5. Fractionnement des données

Extraire environ 20 % de l'ensemble de données pour la production afin de simuler les données réelles et sera utilisé pour tester le modèle. Cet ensemble de données est différent de l'ensemble de données de test, qui est également utilisé pour tester le modèle.

3.6. Suréchantillonnage

Le problème des pondérations extrêmement déséquilibrées entre les cas frauduleux et légitimes dans l'ensemble de données actuel, avec seulement 1.2 % du total des cas étant une fraude et seulement 7 000 observations étiquetées comme telles tandis que les 580 000 observations restantes sont étiquetées comme des transactions propres, a été résolu en utilisant le La technique de suréchantillonnage synthétique des minorités (SMOTE), qui est un algorithme de suréchantillonnage qui génère des échantillons synthétiques pour la classe minoritaire. Étant donné sa conception, cette technique a surmonté le surajustement.

3.7. Évaluation du modèle

Divers paramètres, notamment la précision, le rappel, l'exactitude et le score F, ont été utilisés pour évaluer les performances et la capacité prédictive du modèle AdaBoost mis en œuvre pour prédire les transactions frauduleuses sur l'ensemble de test. Ces indicateurs ont été dérivés des informations recueillies à partir de la matrice de confusion de chaque modèle, où TP et FP représentent les vrais positifs et les faux positifs. En revanche, FN et TN représentent des faux négatifs et des vrais négatifs. En comparant les mesures de performance du modèle AdaBoost à celles d'autres modèles d'apprentissage automatique, une évaluation complète de

ses capacités prédictives et de son efficacité globale dans la détection des fraudes pourrait être réalisée. Cette analyse a facilité une compréhension approfondie de la performance du modèle AdaBoost par rapport aux approches alternatives d'apprentissage automatique, fournissant un aperçu de ses forces et de ses domaines d'amélioration.

4. Résultats et discussion

4.1. Résultats

Les résultats obtenus à partir de l'évaluation des performances des modèles XGBoost, CatBoost et AdaBoost dans la prédiction des transactions frauduleuses sont présentées dans les tableaux ci-dessous.

Le tableau 22 présente la matrice de confusion des modèles, affichant les vrais positifs (TP), les faux positifs (FP), les vrais négatifs (TN) et les faux négatifs (FN) pour chaque modèle. Ces valeurs servent de base au calcul de diverses mesures de performance.

Tableau 22. « Matrice de confusion des modèles utilisés pour la détection de fraude ».

Model	**TP**	**FP**	**TN**	**FN**
AdaBoost	1003	2335	91288	17
CatBoost	1020	5534	88089	0
XGBoost	990	2262	91361	30

Les mesures de performance des modèles, comme indiqué dans les tableaux 23, 24 et 25, donnent un aperçu de leur efficacité à prédire les transactions frauduleuses.

Tableau 23. « Évaluation du modèle XGBoost ».

Class	**Precision**	**Recall**	**F-score**
Fraudulent	0.30	0.97	0.46
Legitimate	1	0.98	0.99
Macro avg	0.65	0.97	0.73
Accuracy			0.98
Execution time			52.34s

Tableau 24. « Évaluation du modèle CatBoost ».

Class	Precision	Recall	F-score
Fraudulent	0.16	1	0.27
Legitimate	1	0.94	0.97
Macro avg	0.58	0.97	0.62
Accuracy			0.94
Execution time			0.02s

Tableau 25. « Évaluation du modèle AdaBoost ».

Class	Precision	Recall	F-score
Fraudulent	0.30	0.98	0.99
Legitimate	1	0.98	0.99
Macro avg	0.65	0.98	0.72
Accuracy			0.98
Execution time			5.90s

4.2. Discussion

Une analyse complète des résultats est présentée sur la base des mesures d'évaluation obtenues à partir des performances des modèles sur l'ensemble de test. Les mesures de performance comprennent la précision, le rappel, l'exactitude et le score F1, qui donnent un aperçu de l'efficacité des modèles pour prédire les transactions frauduleuses.

Le tableau 23 présente les métriques de performance pour le modèle XGBoost. Il a obtenu un rappel de 0.97, indiquant qu'il a correctement identifié 97 % des transactions frauduleuses. Le score de précision de 0.30 suggère que sur toutes les transactions classées comme frauduleuses par XGBoost, seules 30% étaient véritablement frauduleuses. Le modèle a obtenu une précision de 0.98, indiquant un taux de prédiction correct global élevé. Le score F1 de 0.46, la moyenne harmonique de la précision et du rappel, fournit une mesure équilibrée des performances du modèle.

Dans le tableau 24, le modèle CatBoost affiche un rappel de 1, indiquant qu'il a correctement identifié toutes les transactions frauduleuses dans l'ensemble de test. Cependant, le score de précision de 0.16 indique que le modèle a mal classé de nombreuses transactions légitimes comme frauduleuses. Cela a conduit à une exactitude globale inférieure de 0.94. Le score F1 de 0.27 met en évidence le compromis entre rappel et précision dans ce modèle.

Comme le montre le tableau 25, le modèle AdaBoost a obtenu un rappel de 0.98, identifiant correctement 98 % des transactions frauduleuses. Le score de précision de 0.30 suggère qu'il avait un taux de précision similaire à XGBoost. Le modèle a obtenu une exactitude de 0.98, correspondant aux performances de XGBoost. Le score F1 de 0.99 indique une performance bien équilibrée entre précision et rappel.

En comparant les trois modèles, CatBoost a démontré le temps d'exécution le plus court et a détecté avec succès toutes les transactions frauduleuses. Cependant, il avait un nombre plus élevé de faux positifs, ce qui réduisait l'exactitude. XGBoost a atteint des scores de précision et de rappel élevés, mais a eu un temps d'exécution plus long. D'autre part, AdaBoost a obtenu de bons résultats en termes d'exactitude, de précision et de rappel tout en maintenant un temps d'exécution raisonnable.

Ces résultats soulignent les compromis entre les différentes métriques d'évaluation et l'importance de prendre en compte les exigences et contraintes spécifiques du système de détection des fraudes. En fonction du cas d'utilisation spécifique et des considérations opérationnelles, les organisations peuvent choisir le modèle qui correspond le mieux à leurs priorités, qu'il s'agisse de maximiser le rappel, la précision, l'exactitude ou un équilibre entre ces métriques.

5. Synthèse

La détection des transactions frauduleuses par carte de crédit en ligne est de la plus haute importance pour atténuer les pertes financières substantielles des organisations. Dans cette étude, l'utilisation de méthodes simples d'apprentissage automatique de type boosting a été proposée pour cette tâche. Plus précisément, le modèle AdaBoost a été mis en œuvre et ses performances en matière de détection des transactions frauduleuses ont été évaluées.

L'efficacité de l'identification et de la prévention de la fraude en ligne a été évaluée en comparant les mesures de performance du modèle AdaBoost avec d'autres modèles d'apprentissage automatique. Le modèle AdaBoost a affiché des taux de rappel, de précision et d'exactitude élevés, ce qui témoigne de sa forte capacité de prédiction. Ces résultats suggèrent que l'utilisation des algorithmes de boosting peut être une approche prometteuse pour améliorer la détection des transactions frauduleuses dans les plateformes de commerce électronique et les systèmes de cartes de crédit.

Cependant, il est important de reconnaître les limites de cette étude. Tout d'abord, l'évaluation a été réalisée à l'aide de données de transaction synthétiques, qui peuvent ne pas refléter pleinement les complexités et les nuances des modèles de fraude du monde réel. Les travaux futurs devraient comprendre l'évaluation de ces modèles sur des ensembles de données réelles diversifiées et à grande échelle afin de valider leur performance et leur généralisabilité.

En outre, la fraude aux paiements électroniques est dynamique, les fraudeurs adaptant continuellement leurs techniques. Il est essentiel que les recherches futures explorent le développement de systèmes de détection de la fraude adaptatifs et proactifs, capables de s'adapter rapidement aux schémas de fraude émergents et de fournir une protection en temps utile.

En mettant en œuvre des systèmes efficaces de détection des fraudes, les organisations peuvent protéger leurs intérêts financiers et garantir un environnement plus sûr et plus sécurisé pour l'utilisation des cartes de crédit et les transactions de commerce électronique. Les progrès continus des techniques d'apprentissage automatique, associés aux efforts continus de recherche et de développement, peuvent encore améliorer la précision et l'efficacité des systèmes de détection des fraudes, protégeant ainsi les entreprises et les consommateurs contre le vol numérique.

Partie V. Analyse des sentiments à travers les commentaires sur les produits pour la recommandation dans le e-commerce : une approche machine learning

1. Contexte

Le retour d'information des autres clients est essentiel lorsqu'il s'agit de prendre une décision, en particulier lorsque les choix sont multiples et les ressources importantes. C'est pourquoi les cyberconsommateurs se réfèrent fréquemment aux expériences antérieures des acheteurs. Ces informations n'étant pas structurées, la collecte des sentiments du public sur une variété de sujets a conduit à l'évolution des domaines de l'exploration d'opinion et de l'analyse des sentiments. Lorsqu'un consommateur souhaite prendre une décision concernant l'achat d'un produit ou d'un service, un grand nombre d'avis d'utilisateurs sont disponibles, mais les lire et les analyser tous est une tâche fastidieuse. De même, lorsqu'une organisation cherche à obtenir un retour d'information du public ou à commercialiser ses produits, ou encore à trouver de nouveaux contrats, à prévoir les tendances des ventes ou même à gérer sa réputation, elle est confrontée à un énorme volume d'avis accumulés de la part de ses clients. Les technologies d'analyse des sentiments permettent d'analyser une vaste gamme de données disponibles et d'extraire les sentiments des clients, ce qui peut en fin de compte aider le client et l'organisation à atteindre leurs objectifs. L'analyse des sentiments est un domaine de l'informatique qui étudie les opinions des gens exprimées dans un texte, la recherche se concentrant sur le traitement du texte afin d'identifier les données d'opinion.

L'analyse des sentiments, ou l'exploration des opinions ou l'intelligence artificielle des émotions, est une technique de traitement du langage naturel (NLP) utilisée pour évaluer si les données sont positives, négatives ou neutres. L'analyse des sentiments est couramment appliquée aux données textuelles dans le but d'aider les entreprises à suivre l'évolution de leur marque et de leurs produits à partir des commentaires de leurs clients et à comprendre les besoins de ces derniers. En effet, connaître les préférences des clients, c'est-à-dire les produits qu'ils recommandent, est un défi majeur pour les entreprises de commerce électronique.

Pour l'analyse du sentiment des critiques de produits en ligne, cette étude propose des techniques d'apprentissage automatique supervisé : Random Forest, Régression logistique, k-Nearest Neighbor, et Catboost Classifier. Le prétraitement des données, l'extraction des caractéristiques et les classifications de polarité ou de sentiment sont les trois principales étapes de l'étude proposée sur l'analyse des sentiments.

2. Méthodologie

Dans cette étude, le but est de prédire si le client recommandera ou non le produit en se basant sur le texte des avis, c'est-à-dire s'il l'apprécie suffisamment pour le recommander. Il s'agit donc d'un problème de classification binaire (la variable cible peut prendre deux valeurs 0 ou 1). Pour ce faire, les étapes illustrées dans la figure ci-dessous (figure 18) sont suivies :

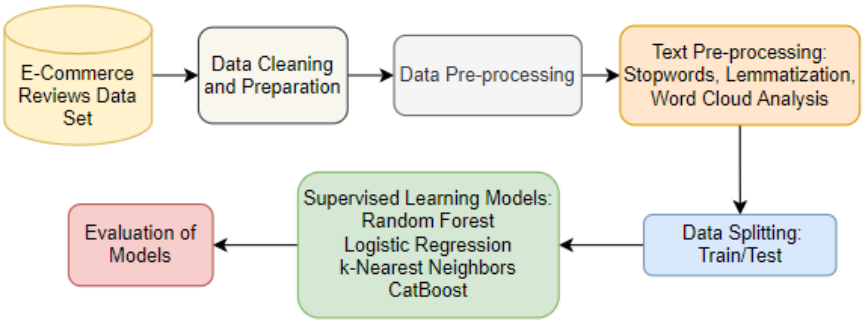


Figure 18. « Le pipeline de l'étude de prédiction des recommandations des clients ».

2.1. Description de l'ensemble de données

L'ensemble de données "E-Commerce Customer Reviews" utilisé dans nos expériences est composé des variables suivantes : "Clothing ID", "Age", "Title", "Rating", "Review Text", "Recommended IND", "Positive Feedback Count", "Division Name", "Class Name" et "Department Name" (tableau 26).

Tableau 26. « Dataset sur les commentaires et les évaluations des clients ».

	Unnamed: 0	Clothing ID	Age	Title	Review Text	Rating	Recommended IND	Positive Feedback Count	Division Name	Department Name	Class Name
0	0	767	33	NaN	Absolutely wonderful - silky and sexy and comf...	4	1	0	Initmates	Intimate	Intimates
1	1	1080	34	NaN	Love this dress! it's sooo pretty. i happene...	5	1	4	General	Dresses	Dresses
2	2	1077	60	Some major design flaws	I had such high hopes for this dress and reall...	3	0	0	General	Dresses	Dresses
3	3	1049	50	My favorite buy!	I love, love, love this jumpsuit. it's fun, fl...	5	1	0	General Petite	Bottoms	Pants
4	4	847	47	Flattering shirt	This shirt is very flattering to all due to th...	5	1	6	General	Tops	Blouses
...	...	...	...	...	...	...	...	...	...	...	...
23481	23481	1104	34	Great dress for many occasions	I was very happy to snag this dress at such a ...	5	1	0	General Petite	Dresses	Dresses
23482	23482	862	48	Wish it was made of cotton	It reminds me of maternity clothes. soft, stre...	3	1	0	General Petite	Tops	Knits
23483	23483	1104	31	Cute, but see through	This fit well, but the top was very see throug...	3	0	1	General Petite	Dresses	Dresses
23484	23484	1084	28	Very cute dress, perfect for summer parties an...	I bought this dress for a wedding i have this ...	3	1	2	General	Dresses	Dresses
23485	23485	1104	52	Please make more like this one!	This dress in a lovely platinum is feminine an...	5	1	22	General Petite	Dresses	Dresses

23486 rows × 11 columns

2.2. Prétraitement et exploration des données

Dans cette étape, nous avons supprimé les données manquantes de chaque variable. Nous avons regroupé la variable "titre du produit" avec le texte du commentaire. Nous avons ensuite montré la distribution de la variable cible "Recommend_IND" (figure 19). Nous pouvons constater que nos données ne sont pas équilibrées car la majorité des clients recommandent les produits achetés (plus de 80 %). Nous avons Créé ensuite une nouvelle variable qui représente la longueur de chaque commentaire "Text_Length". Nous avons analysé ensuite la relation entre la variable longueur du texte et la variable cible (figure 20). Ensuite, on a visualisé la corrélation entre la variable cible et la variable "Rating" (figures 21, 22). Les bibliothèques Python courantes utilisées pour effectuer ces tâches de prétraitement comprennent "NLTK" (Natural Language Toolkit) et "RE" (Regular Expression).

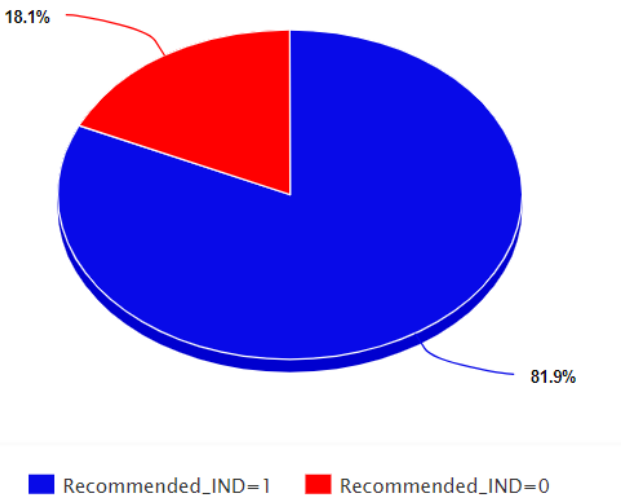


Figure 19. « Proportions des produits recommandés et non recommandés ».

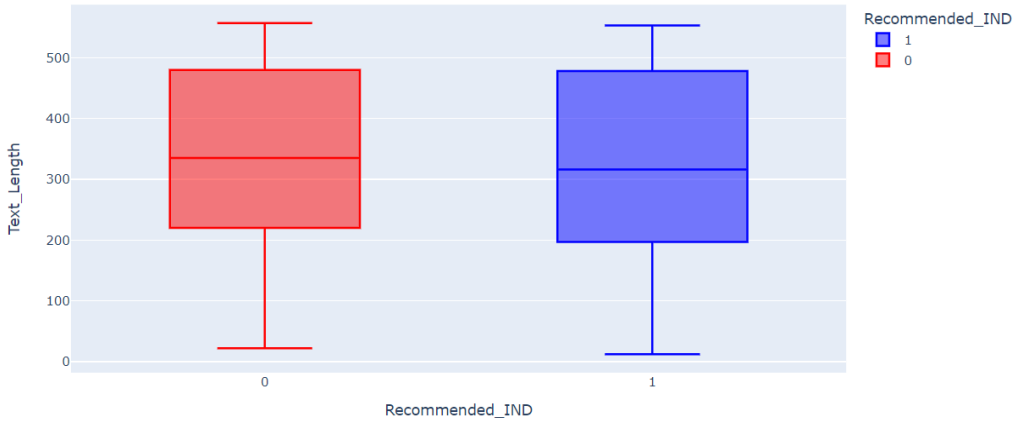


Figure 20. « Variable cible "Recommend_IND" en fonction de la longueur du texte ».

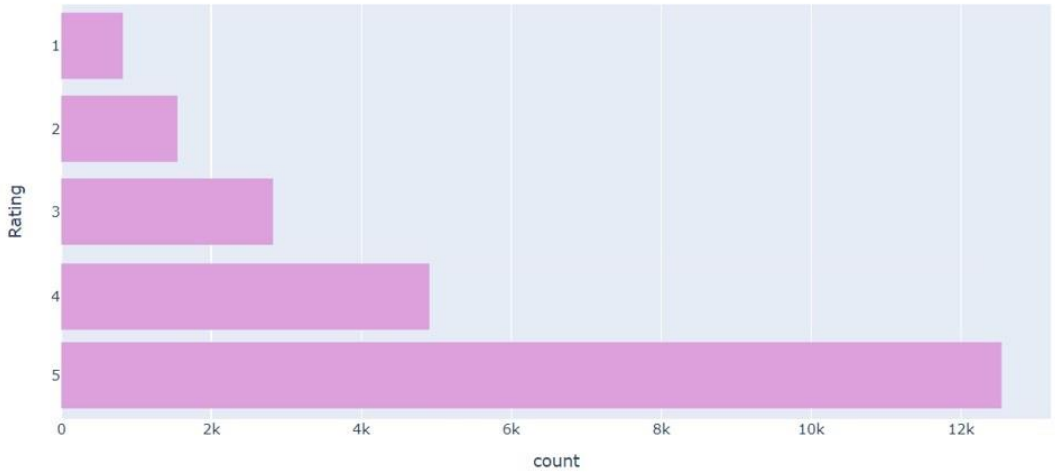


Figure 21. « Distribution des évaluations des produits ».

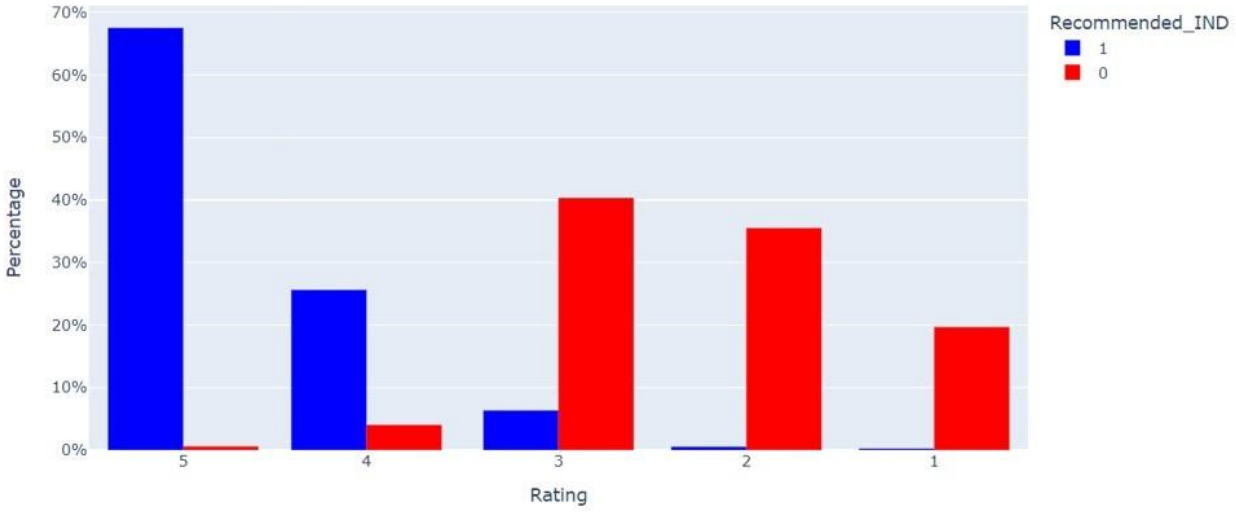


Figure 22. « La variable cible "Recommend_IND" par rapport à l'évaluation du produit ».

2.3. Analyse de la polarité

La polarité est une valeur flottante qui se situe dans l'intervalle [-1,1], où 1 signifie un retour d'information positif et -1 un retour d'information négatif. La distribution du score de polarité dans les commentaires des clients est illustrée à la figure 23. La majorité des commentaires se situe du côté positif du graphique [0,1].

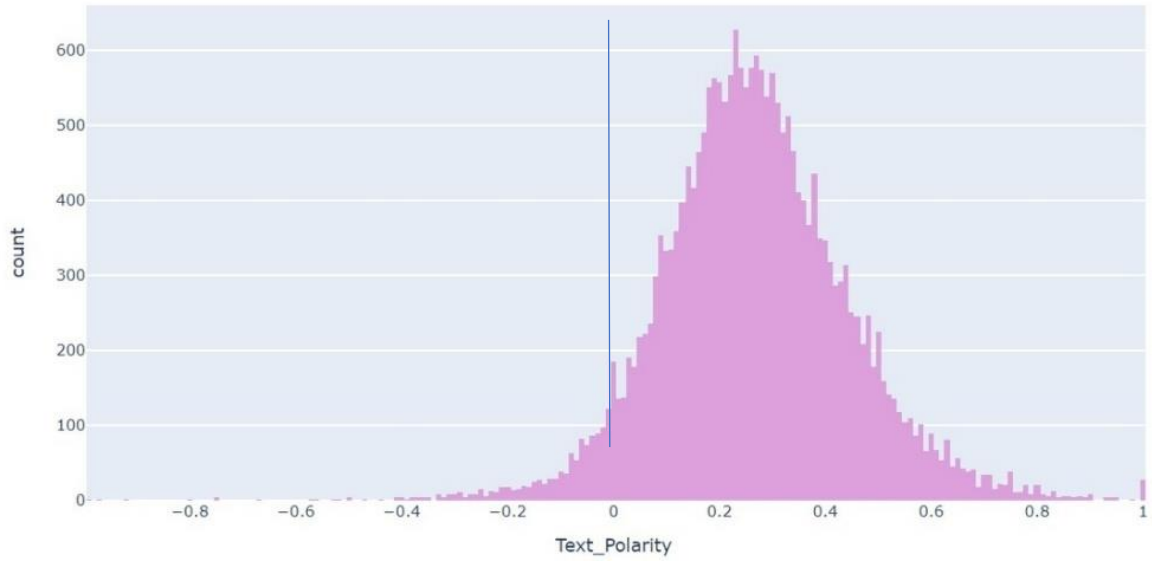


Figure 23. « La polarité du texte des commentaires des clients ».

2.4. Prétraitement du texte

Dans cette phase, les caractères spéciaux ont été supprimés (!"#$%&\'()*+,-./:;<=>?@[\\N]^_`{|}~), et tout le texte des commentaires est converti en minuscules. En fonction des valeurs de la variable Text_Polarity, deux sentiments différents peuvent être dégagés. Un sentiment positif est présent si Text_Polarity est supérieur à zéro. Inversement, un sentiment négatif est présent si Text_Polarity est inférieur à zéro (tableau 27).

Tableau 27. « Dataset sur la polarité et le sentiment dans les commentaires des clients ».

	Rating	Class_Name	Recommended_IND	Text	Text_Length	Text_Polarity	Sentiment
0	4	Intimates	1	absolutely wonderful silky and sexy and comf...	54	0.633	Positive
1	5	Dresses	1	love this dress its sooo pretty i happened ...	304	0.340	Positive
2	3	Dresses	0	some major design flaws i had such high hopes ...	524	0.073	Positive
3	5	Pants	1	my favorite buy i love love love this jumpsuit...	141	0.561	Positive
4	5	Blouses	1	flattering shirt this shirt is very flattering...	209	0.513	Positive
5	2	Dresses	0	not for the very petite i love tracy reese dre...	512	0.181	Positive
6	5	Knits	1	cagrcoal shimmer fun i aded this in my basket ...	517	0.158	Positive
7	4	Knits	1	shimmer surprisingly goes with lots i ordered ...	519	0.230	Positive
8	5	Dresses	1	flattering i love this dress i usually get an ...	177	0.003	Positive
9	5	Dresses	1	such a fun dress im 55 and 125 lbs i ordered t...	378	0.202	Positive
10	3	Dresses	0	dress looks like its made of cheap material dr...	381	-0.047	Negative
11	5	Dresses	1	this dress is perfection so pretty and flatte...	52	0.250	Positive

La figure 24 représente le pourcentage des sentiments positifs et négatifs par rapport à la variable cible "Recommend_IND".

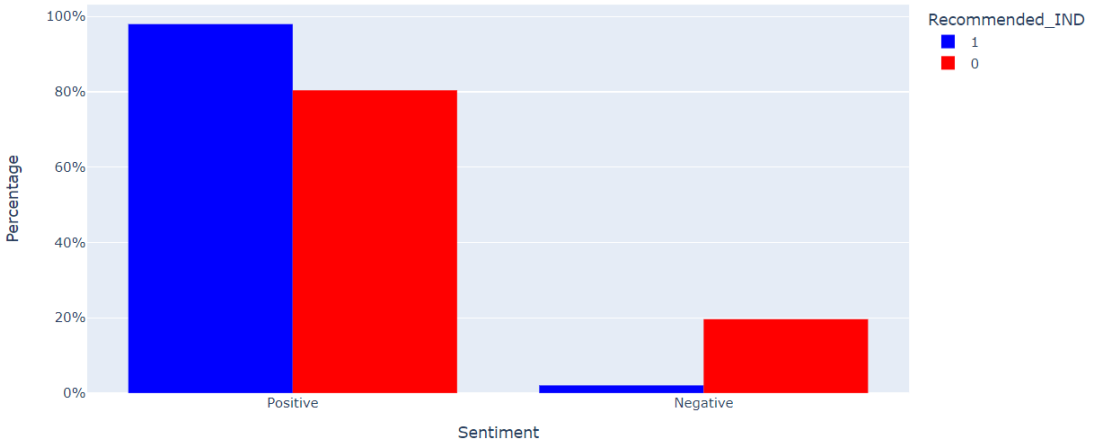


Figure 24. « Les sentiments par rapport à la variable cible "Recommend_IND" en pourcentage ».

2.5. Mots vides, stemming et lemmatisation

Cette étape consiste à supprimer les expressions régulières et les mots vides, ainsi qu'à tronconner et à lemmatiser le texte des critiques. Le stemming consiste à supprimer les derniers caractères d'un mot, ce qui entraîne souvent des erreurs d'orthographe et de signification. En tenant compte du contexte, la lemmatisation réduit un terme à sa forme logique de base, ou lemme.

Exemple :

- Original : changing, arrived
- Stemming : chang, arriv
- Lemmatization : change, arrive

2.6. Analyse du nuage de mots

Enfin, les mots communs ont été supprimés qui n'affectent pas la prédiction des avis négatifs ou positifs. La fréquence des mots est représentée graphiquement dans les nuages de mots. La taille du mot dans le graphique créé augmente proportionnellement à la fréquence d'apparition du mot clé dans les avis analysés. Le nuage de mots positif est représenté à la figure 25 et le nuage de mots négatif à la figure 26. L'ensemble de données final contenant la version actualisée des critiques et la variable cible est illustré au tableau 28.



Figure 25. « Nuage de mots positifs relatif au commentaires des clients ».



Figure 26. « Nuage de mots négatifs relatif au commentaires des clients ».

Tableau 28. « Dataset contenant le texte des commentaires et la variable cible ».

	Updated_Review_Text	Recommended_IND
0	absolutely wonderful silky sexy comfortable	1
1	love sooo pretty happen find store glad bc nev...	1
2	major design flaws high hope really wanted wor...	0
3	favorite buy love love love jumpsuit fun flirt...	1
4	flatter flatter due adjustable front tie perfe...	1
5	petite love tracy reese petite 5 foot tall usu...	0
6	cagrcoal shimmer fun aded basket hte last mint...	1
7	shimmer surprisingly go lot carbon store pick ...	1
8	flatter love usually get xs run little snug bu...	1
9	fun 55 125 lb petite make sure length wasnt lo...	1

2.7. Modèles prédictifs

L'ensemble d'échantillons a été divisé en un ensemble de formation de 70 % et un ensemble de test de 30 % des données. Ensuite, les modèles d'apprentissage supervisé ont été mis en œuvre, à savoir : Random Forest, Logistic Regression, k-Nearest Neighbor et Catboost Classifier, comme le montre la figure 27.

```
Classification_models = {
    "K-Nearest Neighbors": KNeighborsClassifier(weights='distance', metric='euclidean',n_neighbors = 3),

    "Logistic Regression": LogisticRegression(C=1.0, class_weight=None, dual=False, fit_intercept=True,
                   intercept_scaling=1, l1_ratio=None, max_iter=100,
                   multi_class='auto', n_jobs=None, penalty='l2',
                   random_state=0, solver='lbfgs', tol=0.0001, verbose=0,
                   warm_start=False),

    "Random Forest": RandomForestClassifier(max_depth=50, min_samples_split=25, criterion = 'entropy', random_state=606),

    "CatBoost Classifier": CatBoostClassifier(depth = 3, max_bin = 75, boosting_type = 'Plain', bootstrap_type = 'MVS',
                    verbose = False),
}
```

Figure 27. « Implémentation des modèles de classification (script) ».

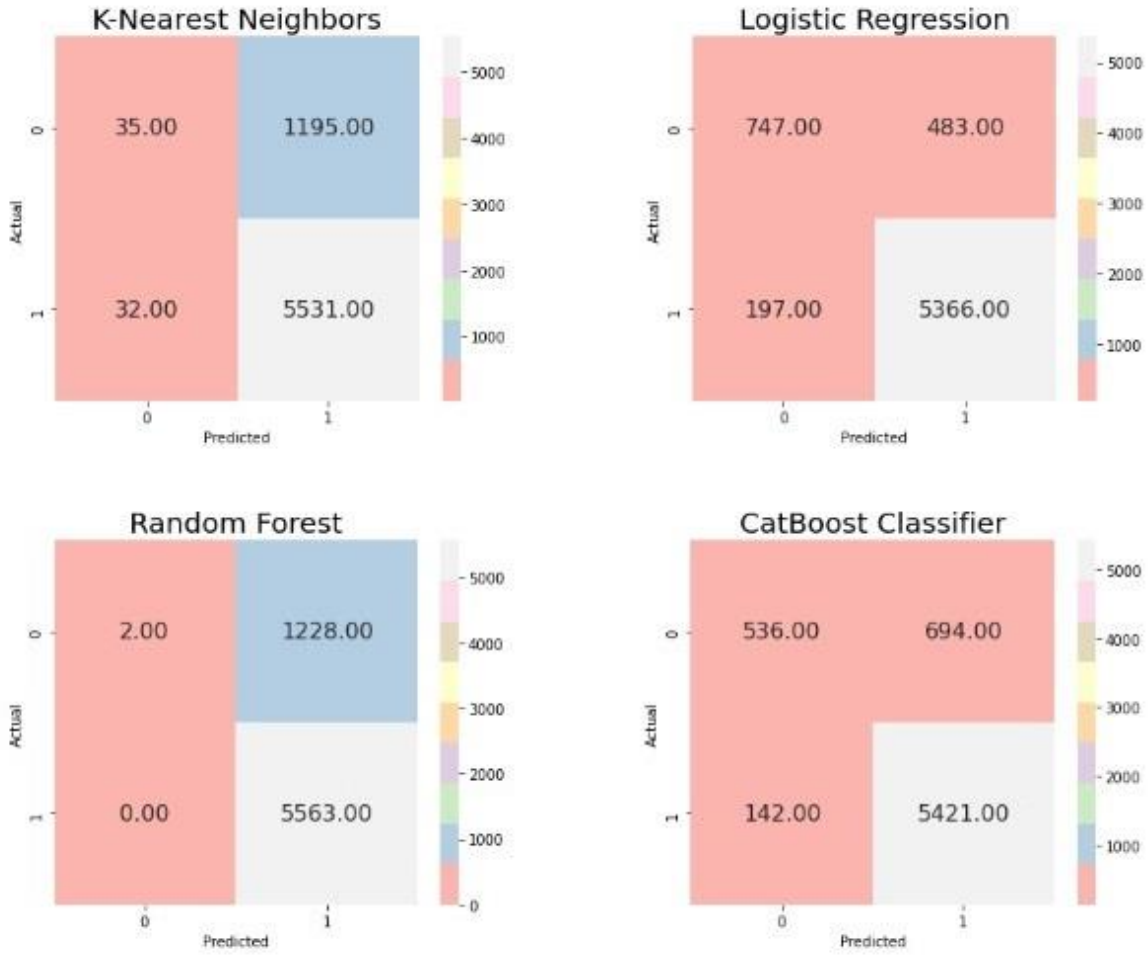
3. Résultats expérimentaux et discussion

3.1. Matrice de confusion

Les résultats des prédictions dans une tâche de classification sont résumés dans une matrice de confusion. Les prédictions correctes et incorrectes sont mises en évidence et divisées en classes. Le résultat des prédictions est comparé aux valeurs réelles. La représentation des matrices de confusion est illustrée dans le tableau 29. Les vrais positifs et les faux positifs sont notés TP et FP, tandis que les faux négatifs et les vrais négatifs sont notés FN et TN. Où :

- TP : est le nombre de clients qui ont recommandé un produit (classe 1), et que le modèle prédictif a correctement prédit.
- TN : est le nombre de clients qui n'ont pas recommandé un produit (classe 0), et que le modèle prédictif a correctement prédit.
- FP : est le nombre de clients qui n'ont pas recommandé de produit (classe 0), mais que l'algorithme prédictif a identifié comme étant de classe 1.
- FN : est le nombre de clients qui ont recommandé un produit (classe 1), mais que le modèle prédictif a identifié comme étant la classe 0.

Tableau 29. « Matrices de confusion des modèles utilisés pour la prédiction des sentiments des commentaires des clients ».



3.2. Indicateurs de performance

Dans le but d'évaluer la performance des modèles appliqués ou la prédiction de la recommandation de produits par les clients sur l'ensemble de test, nous avons utilisé différentes métriques, telles que la précision, le rappel, l'exactitude et le F1-score. Ils mesurent la capacité des modèles prédictifs à prédire correctement la recommandation du client (0 ou 1). Les quatre indicateurs mentionnés précédemment sont calculés à partir des informations capturées à l'aide de la matrice de confusion. Où :

- Le rappel est le ratio de vrais churners ou de vrais positifs (TP), et se calcule comme suit :

$$\text{Rappel} = \frac{\text{TP}}{\text{TP+FN}} \quad (9)$$

- La précision est le ratio des churners corrects prédits, sa formule est la suivante :

$$\text{Précision} = \frac{\text{TP}}{\text{TP+FP}} \quad (8)$$

- L'exactitude est le ratio du nombre de toutes les prédictions correctes, sa formule est la suivante :

$$\text{Accuracy} = \frac{number\ of\ correct\ predictions}{total\ number\ of\ predictions} * 100 \quad (14)$$

- Le score F est la moyenne harmonique de la précision et du rappel et s'écrit comme suit :

$$\text{Score F1} = \frac{2 * \text{Précision} * \text{Rappel}}{Précision+Rappel} \quad (12)$$

Le tableau suivant (tableau 30) montre les valeurs d'exactitude, de rappel, de précision et de score f1 des quatre modèles utilisés.

Tableau 30. « Valeurs d'exactitude, de précision, de rappel et de score f1 des modèles utilisés ».

	Accuracy	Precision	Recall	F1-Score
K-Nearest Neighbors	0.819	0.672	0.511	0.477
Logistic Regression	0.900	0.854	0.786	0.814
Random Forest	0.819	0.910	0.501	0.452
CatBoost Classifier	0.877	0.839	0.705	0.745

3.3. Courbe AUC-ROC

Pour illustrer la capacité de diagnostic des classificateurs binaires, une courbe ROC (Receiver Operator Characteristic) sert de représentation graphique. Il peut être utile de résumer les performances de chaque classificateur en une seule mesure lors de la comparaison de plusieurs classificateurs. Le calcul de la SSC, souvent appelée aire sous la courbe ROC, est l'une de ces stratégies. L'AUC fonctionne bien dans les scénarios du monde réel en tant qu'indicateur général de la précision de la prédiction. La courbe AUC-ROC pour les modèles utilisés est représentée dans la figure ci-dessous (figure 28).

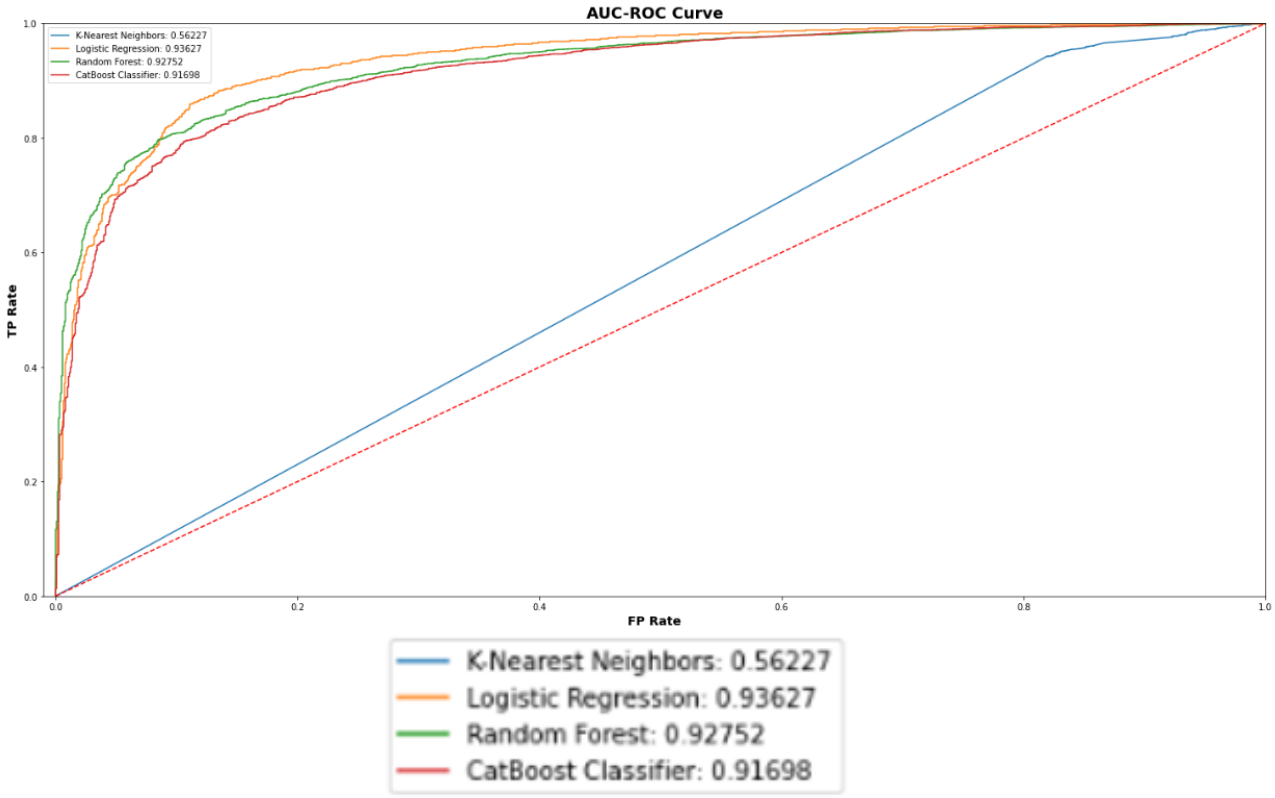


Figure 28. « Courbe AUC-ROC des modèles de prédiction des sentiments des commentaires des clients ».

3.4. Discussion

Le modèle des k-voisins les plus proches a obtenu une précision de 81.9 %, mais de faibles valeurs de précision, de rappel et de score f1. Il a également obtenu le score AUC le plus faible (56.22 %), ce qui a réduit ses performances en matière de prédiction.

Le modèle Random Forest a obtenu la même précision que le modèle k-Nearest Neighbors, mais l'a surpassé en termes de précision (91 %) et de score AUC (92.75 %).

Le modèle Catboost a obtenu des valeurs satisfaisantes pour l'exactitude (A= 87.7 %), le rappel (R= 70.5 %), la précision (P= 83.9 %), le score f1 (F1= 74.5 %) et le score AUC (AUC= 91.69 %). Il a surpassé les deux modèles précédents, à savoir k-Nearest Neighbors et Random Forest.

Cependant, le modèle de régression logistique a surpassé tous les modèles précédents en termes d'indicateurs de performance (A= 90 %, R= 78.6 %, F1= 81.4 %), ainsi qu'un score AUC de 93.62 %, ce qui le place en tête de liste, suivi par le modèle CatBoost.

4. Synthèse

L'utilisation des commentaires des clients dans le commerce électronique afin de recueillir des informations pour la gestion des relations avec la clientèle est un concept intéressant. Les modèles proposés ci-dessus ont atteint une précision comprise entre 80 et 90 %. Le modèle de régression logistique a surpassé les autres modèles en termes de paramètres de la matrice de confusion et de score AUC.

Certains travaux futurs peuvent être intégrés pour améliorer le modèle et le rendre plus performant dans des situations réelles. Un développement ultérieur de ce modèle pourrait inclure la communication avec l'utilisateur pour obtenir plus de données sur le client ou un produit particulier, afin de construire un outil puissant qui aidera non seulement les consommateurs à prendre une bonne décision d'achat, mais sera également un outil très intéressant pour la gestion des relations avec la clientèle. En outre, comme nous avons utilisé un ensemble de données à grande échelle, nous pouvons mettre en œuvre le modèle le plus performant sur des sites de marché locaux afin d'atteindre des niveaux plus élevés de précision et de facilité de gestion. Par conséquent, nous poursuivrons cette recherche jusqu'à ce que nous puissions généraliser ce modèle à tous les types d'avis textuels.

Partie VI. Prédire les ventes futures en utilisant les algorithmes machine learning

1. Contexte

La satisfaction des clients est un facteur primordial dans le monde des affaires. Les organisations cherchent toujours à répondre à leurs demandes le plus souvent possible tout en cherchant constamment à réaliser davantage de profits grâce à un investissement soigneusement calculé pour augmenter les ventes. D'autre part, la prédiction des ventes futures permet aux organisations d'augmenter leurs profits en prenant les bonnes décisions. La prévision des ventes est le processus de prédire les ventes futures sur la base des ventes passées. Elle est essentielle pour les entreprises qui souhaitent se développer sur de nouveaux marchés, proposer de nouveaux produits ou services ou connaître une expansion significative. La raison principale de la prévision est d'adapter les ressources marketing et commerciales à la capacité de production. La prévision des ventes garantit que la bonne combinaison de tarification, de promotions et de marketing est en place pour répondre à la demande, qu'il y a suffisamment de vendeurs pour obtenir le volume de commandes attendu et que les ressources essentielles sont disponibles du côté de la demande. Par conséquent, de nombreuses entreprises consacrent des ressources financières et humaines importantes à cette tâche, ce qui nécessite un investissement considérable. Il est essentiel pour les organisations de production et les entreprises commerciales de disposer d'un système de prévision des ventes efficace et précis afin de ne pas risquer des pertes dues à des prédictions de modèles fausses ou inexactes. D'autre part, une meilleure prévision des ventes permet une utilisation optimisée des ressources, une planification, un transport et un contrôle des stocks efficaces. Cela augmente la satisfaction des clients et réduit les coûts de production.

Le machine learning s'est révélé très efficace dans plusieurs tâches dans le monde des affaires en ligne, notamment la prédiction de la désertion, l'analyse des sentiments, les recommandations et la prédiction des ventes. À l'aide d'algorithmes d'apprentissage automatique, les entreprises peuvent analyser de vastes quantités de données sur les clients pour identifier des schémas et des tendances pouvant être utilisés pour prédire le comportement des clients. Ces informations peuvent être utilisées pour développer des stratégies de fidélisation des clients et de réduction de la désertion, ainsi que pour formuler des recommandations de produits personnalisées. L'apprentissage automatique peut également aider les entreprises à

prédire avec précision les volumes de ventes, ce qui leur permet d'optimiser l'utilisation des ressources et le contrôle des stocks.

2. Méthodologie

L'objectif de cette étude est de comparer quatre algorithmes d'apprentissage automatique pour la prédiction des ventes, à savoir la régression linéaire, le régresseur de forêt aléatoire, XGBoost et LSTM. Les raisons de choisir ces algorithmes sont les suivantes :

La régression linéaire est simple à comprendre et à expliquer, et peut être régularisée pour éviter le surajustement.

La forêt aléatoire fournit des prédictions solides et facilement compréhensibles, traite efficacement de grands ensembles de données et atteint un niveau de précision supérieur à l'algorithme de l'arbre de décision pour prédire les résultats.

XGBoost utilise la puissance du traitement parallèle, est très flexible et peut gérer les données manquantes grâce à ses fonctions intégrées. De plus, il est plus rapide que le Gradient Boosting.

Enfin, les cellules LSTM sont utilisées dans les réseaux de neurones récurrents qui apprennent à prédire l'avenir à partir de séquences de longueurs variables. Les réseaux de neurones récurrents fonctionnent avec tout type de données séquentielles et, contrairement à ARIMA et Prophet, ne sont pas limités aux séries chronologiques.

Dans le but de développer et d'évaluer les modèles de prédiction utilisés, nous avons adopté la méthodologie présentée dans la figure 29. De plus, nous avons développé une gamme de fonctions répertoriées dans le tableau 31.

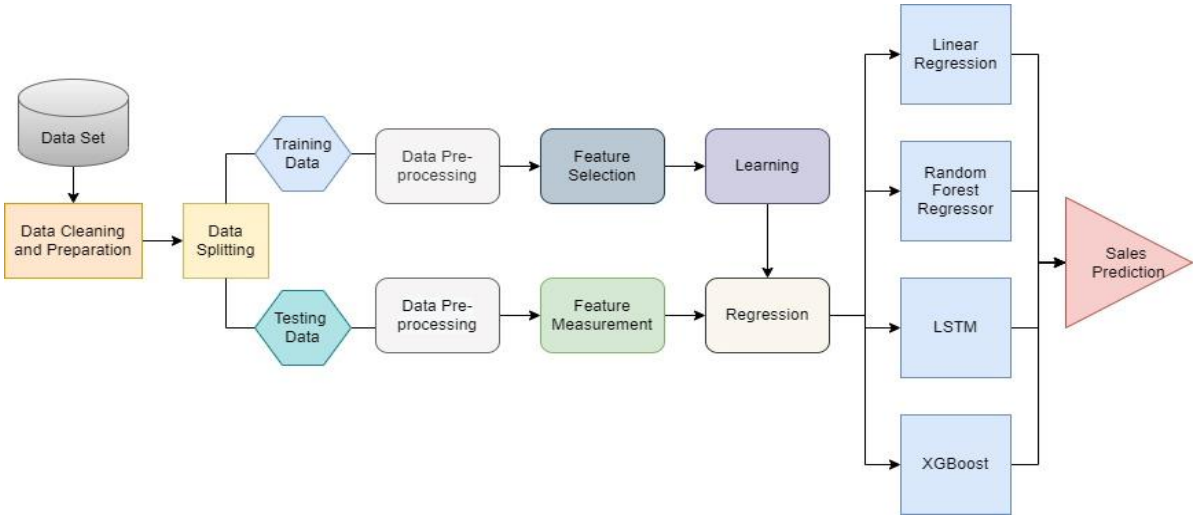


Figure 29. « Architecture de la méthodologie adoptée pour la prévision des ventes ».

Tableau 31. « Les fonctions utilisées pour la prédiction des ventes ».

Fonction	**Description**
Train test split	les données sont divisées de telle manière que les 12 derniers mois font partie de l'ensemble de test, et le reste des données sert à former notre modèle
Scaling the data	les données sont mises à l'échelle de sorte que toutes les variables soient entre -1 et 1 grâce à un scalage min-max
Inverting the scale	l'objectif de cette fonction est d'inverser la mise à l'échelle de la fonction précédente
Create prediction data frame	le data frame comprend les ventes réelles collectées dans notre ensemble de test ainsi que les résultats prédits par notre modèle pour mesurer les performances de chaque modèle
Score the models	évaluer les modèles en fonction de l'erreur quadratique moyenne, de l'erreur absolue moyenne et du coefficient de détermination

2.1. Description et visualisation des données

Le jeu de données utilisé est le "Store Item Demand Forecasting Challenge", qui contient les données de vente d'une chaîne de magasins. Le jeu de données comprend quatre caractéristiques, à savoir "date", "store", "item" et "sales". Il contient un total de 913 000 lignes,

chaque ligne représentant les données de vente pour une date, un magasin et un article spécifiques.

Tout d'abord, les données sont chargées et converties dans une forme qui sera ultérieurement utilisée pour chacun des modèles de prévision des ventes. Dans sa structure initiale, une journée de ventes dans l'un des dix magasins est représentée par chaque ligne de données, comme le montre le tableau 32 (a).

Tableau 32. (a) « Data frame des ventes quotidiennes par magasin », (b) « Data frame des ventes mensuelles totales ».

	date	store	item	sales
0	2013-01-01	1	1	13
1	2013-01-02	1	1	11
2	2013-01-03	1	1	14
3	2013-01-04	1	1	13
4	2013-01-05	1	1	10

(a)

	date	sales
0	2013-01-01	454904
1	2013-02-01	459417
2	2013-03-01	617382
3	2013-04-01	682274
4	2013-05-01	763242

(b)

L'objectif est de prévoir les ventes mensuelles, donc toutes les journées et tous les magasins sont d'abord regroupés en ventes mensuelles totales dans un nouveau data frame, comme le montre le tableau 32 (b), où chaque ligne correspond aux ventes totales pour un mois spécifique dans tous les magasins.

2.2. Analyse exploratoire des données

La figure ci-dessous (figure 30) représente graphiquement la distribution des ventes par jour et la distribution des ventes totales par magasin.

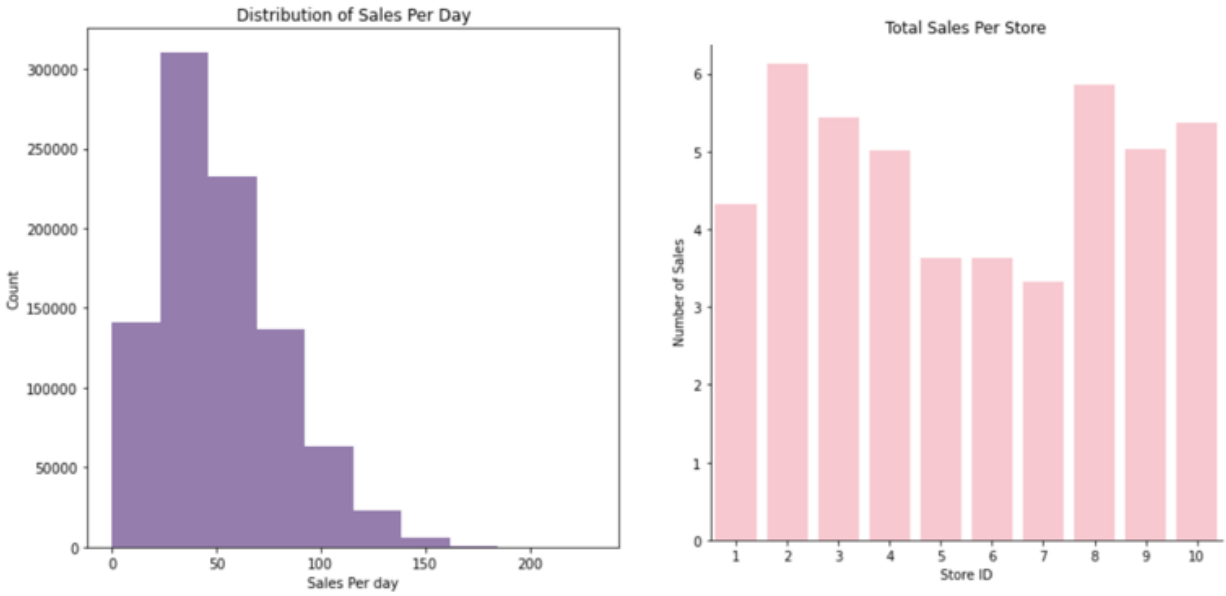


Figure 30. « Distribution des ventes par jour et des ventes totales par magasin ».

2.3. Détermination de la stationnarité

Le graphe des ventes mensuelles totales en fonction du temps dans la figure 31 montre que les ventes mensuelles moyennes augmentent avec le temps, ce qui indique que les données ne sont pas stationnaires. Pour rendre les données stationnaires, la différence entre les ventes de chaque mois a été calculée et ajoutée au data frame en tant que nouvelle caractéristique (sales_diff), comme le montre le tableau 33.

La figure 32 est une représentation graphique de l'apparence des données après la transformation de différenciation.

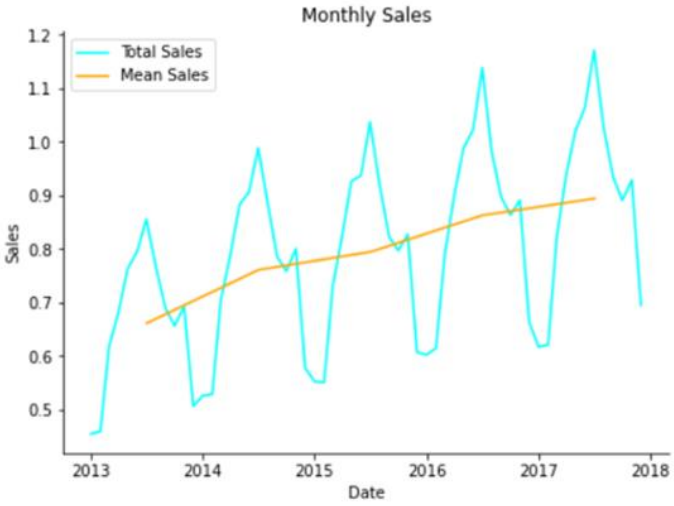


Figure 31. « Graphique des ventes mensuelles ».

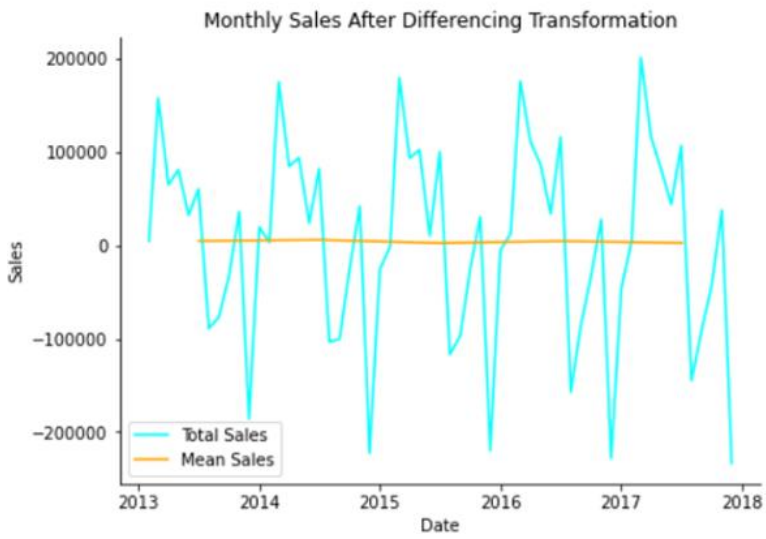


Figure 32. « Graphique des ventes mensuelles après la transformation de différenciation ».

2.4. Observation des retards

Les graphiques d'autocorrélation (figure 33) et d'autocorrélation partielle (figure 34) ci-dessous fournissent une période de rétroaction cohérente pour les modèles de régression.

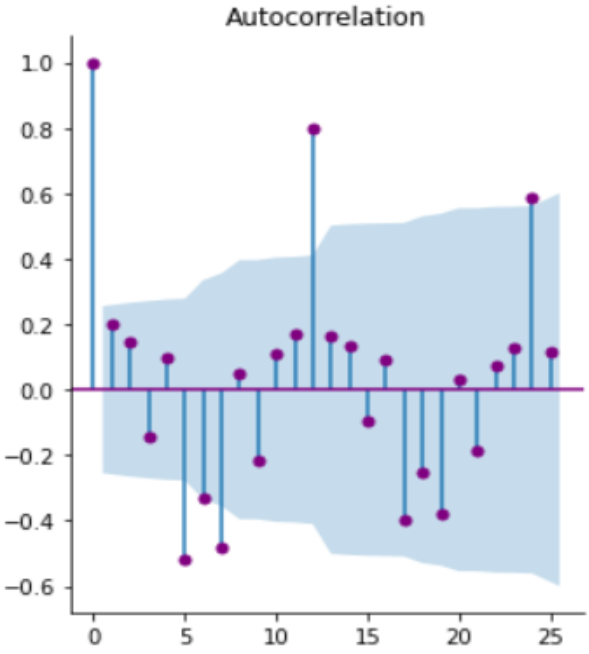


Figure 33. « Graphique d'autocorrélation ».

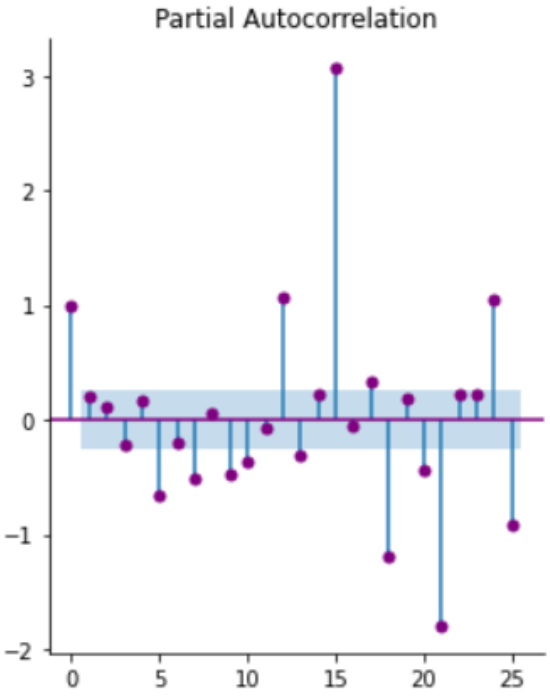


Figure 34. « Graphique d'autocorrélation partielle ».

2.5. Modélisation régressive de l'ensemble de données

Sur la base des graphiques précédents (figures 33 et 34), une période de rétroaction de 12 mois a été choisie pour être incluse dans notre ensemble de caractéristiques. Ensuite, un data frame a été généré, composé de 13 colonnes, avec 12 colonnes représentant les mois et une colonne pour la variable dépendante "sales_diff" (tableau 33).

Tableau 33. « Data frame des ventes mensuelles totales ».

	date	sales	sales_diff	lag_1	lag_2	lag_3	lag_4	lag_5	lag_6	lag_7	lag_8	lag_9	lag_10	lag_11	lag_12
0	2014-02-01	529117	3130.0	19380.0	-186036.0	36056.0	-33320.0	-76854.0	-89161.0	60325.0	32355.0	80968.0	64892.0	157965.0	4513.0
1	2014-03-01	704301	175184.0	3130.0	19380.0	-186036.0	36056.0	-33320.0	-76854.0	-89161.0	60325.0	32355.0	80968.0	64892.0	157965.0
2	2014-04-01	788914	84613.0	175184.0	3130.0	19380.0	-186036.0	36056.0	-33320.0	-76854.0	-89161.0	60325.0	32355.0	80968.0	64892.0
3	2014-05-01	882877	93963.0	84613.0	175184.0	3130.0	19380.0	-186036.0	36056.0	-33320.0	-76854.0	-89161.0	60325.0	32355.0	80968.0

2.6. Résultats de la prévision des ventes

Les figures suivantes (figures 35, 36, 37, et 38) montrent les résultats de la prévision des ventes pour chaque modèle.

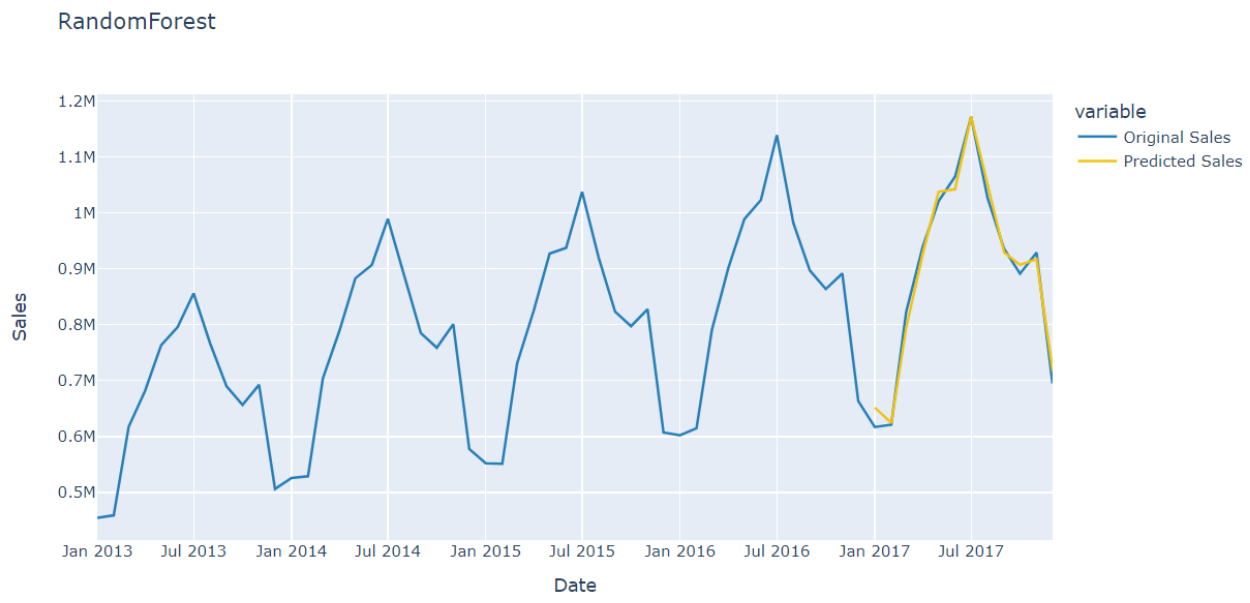


Figure 35. « Graphique de prévision des ventes pour le modèle RF ».

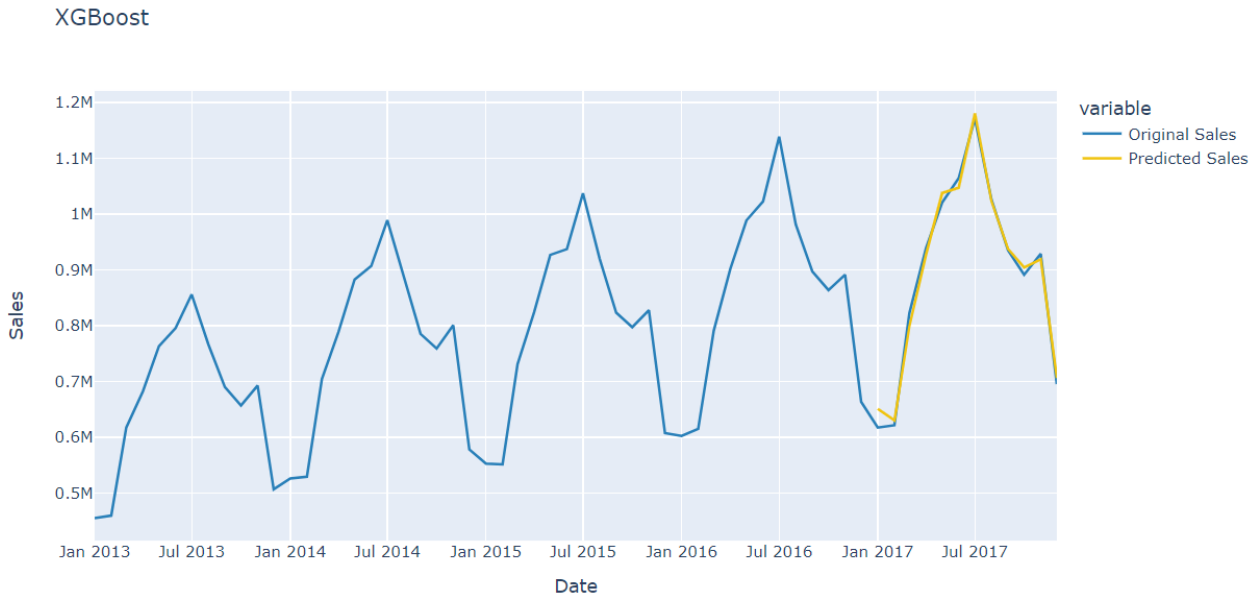


Figure 36. « Graphique de prévision des ventes pour le modèle XGBoost ».

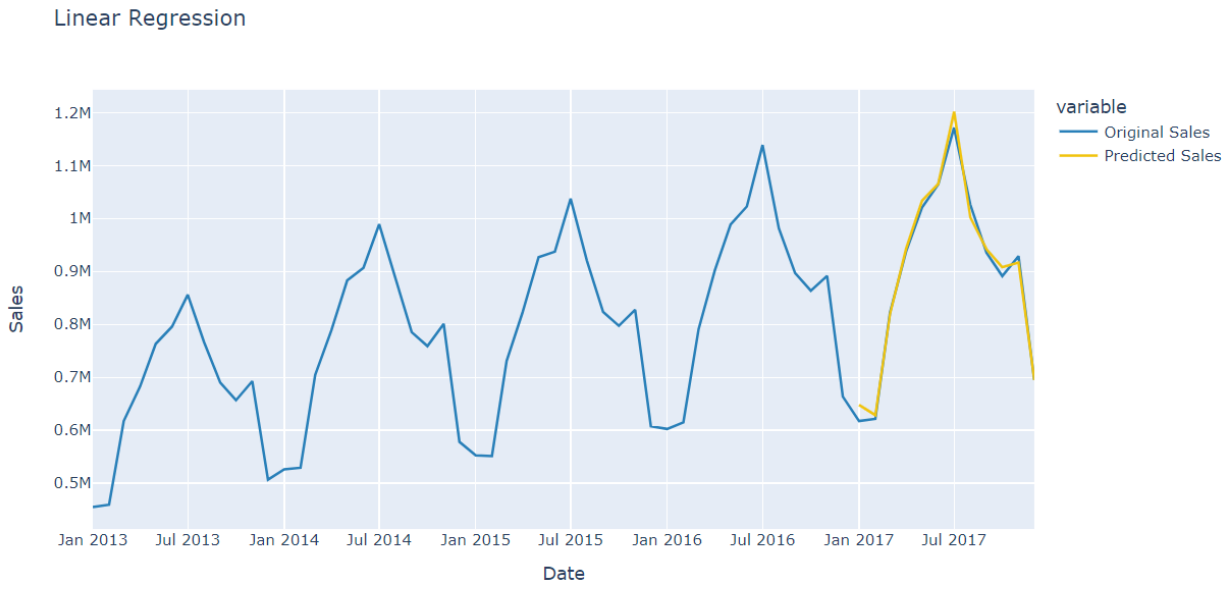


Figure 37. « Graphique de prévision des ventes pour le modèle Linear Regression ».

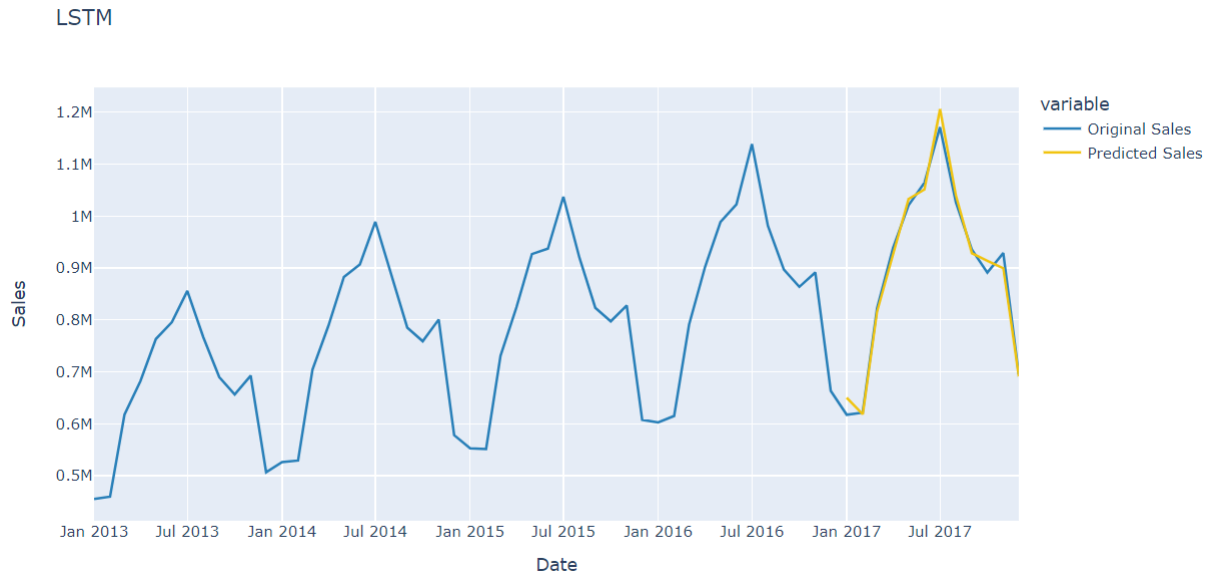


Figure 38. « Graphique de prévision des ventes pour le modèle LSTM ».

3. Comparaison des modèles et résultats

Les indicateurs MAE, RMSE et R^2 ont été utilisés pour comparer les modèles de prévision des ventes, comme indiqué dans le tableau 34.

- Le MAE ou erreur absolue moyenne fournit une mesure de la magnitude moyenne des erreurs dans une plage de prédictions, indépendamment de leur direction. Le MAE est la moyenne sur l'échantillon testé des différences absolues entre la prédiction et l'observation réelle, où toutes les différences individuelles ont le même poids.

- Le RMSE ou erreur quadratique moyenne est une méthode de notation quadratique qui fournit également une mesure de la magnitude moyenne de l'erreur. Il correspond à la racine carrée de la moyenne des différences au carré entre la prédiction et l'observation réelle.
- Les valeurs de MAE et RMSE peuvent varier de 0 à ∞ et sont des scores orientés négativement, ce qui signifie que les valeurs les plus basses sont préférables.
- Le R^2 ou coefficient de détermination sert à évaluer la qualité d'une régression linéaire. Il correspond mathématiquement à la proportion de la variance d'une variable dépendante expliquée par une ou plusieurs variables indépendantes dans le modèle de régression. Il représente le pourcentage de variation de la variable de réponse expliquée par un modèle linéaire. La valeur de R^2 est toujours comprise entre 0 et 1. Si R^2 a une valeur approchante 0, cela signifie que le modèle n'explique aucune de la variabilité des données de réponse autour de sa moyenne. En revanche, si R^2 a une valeur approchante 1, cela signifie que le modèle explique toute la variabilité des données de réponse autour de sa moyenne.

Tableau 34. « Valeurs de RMSE, MAE et R^2 pour chaque modèle ».

Modèle	**RMSE**	**MAE**	**R- Squared**
LSTM	19281.054287	16007.250000	0.976882
Random Forest	19227.637970	16531.583333	0.976955
Linear Regression	16221.040791	15433.000000	0.980716
XGBoost	13701.003360	12342.666667	0.991301

À partir de la comparaison des mesures de performance dans le tableau 32, on peut voir que le modèle XGBoost est celui qui donne les meilleurs résultats, avec un coefficient de détermination de 99.13% et des erreurs relativement faibles par rapport aux autres modèles. L'algorithme de régression linéaire se classe en deuxième position avec un coefficient de détermination de 98.07%. Les algorithmes de forêt aléatoire et de LSTM se classent en troisième position avec des erreurs plus élevées et des coefficients de détermination relativement faibles.

4. Synthèse

Dans cette étude, les algorithmes d'apprentissage automatique supervisés suivants ont été mis en œuvre : régression linéaire, forêt aléatoire, XGBoost et LSTM, pour résoudre un problème de régression qui est la prédiction des ventes futures. Le modèle XGBoost s'est avéré plus performant dans notre cas de prédiction des ventes, suivi de près par le modèle de régression linéaire. Dans cette étude, la forme simplifiée de chaque modèle a été utilisée pour illustrer comment ces modèles peuvent être utilisés pour la prédiction des ventes. Afin de minimiser la complexité, les modèles ont été partiellement ajustés. En revanche, le modèle LSTM, en particulier, peut être amélioré en ajoutant de nombreux nœuds et couches supplémentaires. Cette étude peut être améliorée en intégrant l'analyse des Big Data, qui est de plus en plus essentielle dans le monde des affaires d'aujourd'hui pour une prévision précise des ventes. Des recherches supplémentaires pourraient explorer des configurations de modèles plus complexes et étudier d'autres caractéristiques qui pourraient avoir un impact sur la prédiction des ventes.

Partie VII. Le machine learning au service de la tarification dynamique dans le e-marketing

1. Contexte

Dans le paysage commercial d'aujourd'hui, rapide et hautement concurrentiel, fixer des prix optimaux est un défi critique auquel font face les entreprises opérant dans le domaine du marketing numérique. Les dynamiques du marché en constante évolution, les préférences changeantes des clients et la concurrence intense nécessitent des stratégies de tarification dynamiques qui peuvent s'adapter aux conditions du marché en temps réel. Les modèles de tarification statiques traditionnels sont souvent insuffisants pour capturer la complexité de ces environnements dynamiques, ce qui entraîne des opportunités manquées et des décisions de tarification sous-optimales.

Pour relever ces défis, les chercheurs et les professionnels se sont tournés vers les techniques d'apprentissage automatique pour développer des modèles de tarification sophistiqués capables de gérer les subtilités des marchés dynamiques. L'apprentissage automatique offre la possibilité d'analyser de vastes quantités de données, d'identifier des modèles et des tendances, et de prendre des décisions de tarification éclairées en temps réel. Parmi les différentes approches d'apprentissage automatique, l'apprentissage par renforcement a suscité beaucoup d'attention en raison de sa capacité à optimiser les stratégies de tarification grâce à un apprentissage et une adaptation continue.

L'objectif de cette étude est de présenter une méthodologie permettant de tirer parti de l'apprentissage automatique, plus précisément du modèle Deep Q-Network (DQN), pour la tarification dynamique dans le marketing numérique. En combinant les techniques d'apprentissage par renforcement avec les réseaux neuronaux profonds, le modèle DQN offre un outil puissant pour traiter la complexité des espaces d'état et d'action de grande dimension dans les problèmes de tarification.

Cette étude contribue au domaine de la tarification dynamique en présentant une méthodologie améliorée qui renforce l'efficacité et la précision des décisions de tarification dans le domaine du marketing numérique. Nous évaluons rigoureusement le modèle DQN, en fournissant des informations détaillées sur ses performances et son impact sur des indicateurs clés tels que la génération de revenus, les taux de conversion, la valeur à vie des clients, la part de marché et les marges bénéficiaires. Les résultats soulignent le potentiel des techniques

d'apprentissage automatique pour générer des améliorations significatives dans les stratégies de tarification et les résultats commerciaux.

2. Cadre théorique

La tarification dynamique est une stratégie de tarification qui consiste à ajuster les prix en temps réel en fonction de divers facteurs tels que la demande, les conditions du marché, la concurrence et le comportement des clients. Elle permet aux entreprises d'optimiser leurs décisions de tarification et de maximiser leur rentabilité en réagissant aux changements de l'environnement du marché. Dans le contexte du marketing numérique, où les transactions et les interactions en ligne sont prédominantes, la tarification dynamique joue un rôle crucial dans la capture de valeur et le maintien d'un avantage concurrentiel.

Dans le domaine de la tarification dynamique, les algorithmes d'apprentissage automatique peuvent analyser les données de ventes historiques, le comportement des clients et d'autres facteurs pertinents pour optimiser les stratégies de tarification et améliorer les performances commerciales.

Le modèle Deep Q-Network est une puissante approche d'apprentissage automatique qui combine des réseaux neuronaux profonds avec des techniques d'apprentissage par renforcement. Il convient particulièrement aux problèmes présentant des espaces d'états et d'actions de grande dimension, ce qui en fait une solution adaptée à la tarification dynamique dans le marketing numérique. Le modèle DQN utilise un algorithme d'apprentissage Q-learning, qui est une forme d'apprentissage par renforcement, pour apprendre les actions de tarification optimales en fonction des récompenses observées et des récompenses futures estimées. En mettant à jour de manière itérative les valeurs Q associées à chaque action, le modèle DQN peut prendre des décisions de tarification intelligentes et s'adapter aux changements des conditions du marché.

En intégrant la tarification dynamique, l'apprentissage automatique et le modèle DQN, cette étude vise à exploiter des techniques informatiques avancées pour optimiser les décisions de tarification dans le domaine du marketing numérique. L'utilisation de l'apprentissage automatique permet au modèle d'apprendre à partir des données et d'adapter ses stratégies de tarification, tandis que l'architecture DQN fournit un cadre robuste pour gérer des scénarios de tarification complexes. À travers cette recherche, nous cherchons à améliorer la compréhension

et l'application des stratégies de tarification dynamique et à contribuer à l'avancement des pratiques du marketing numérique.

3. Méthodologie

La méthodologie utilisée dans cette étude implique la mise en œuvre d'un modèle Deep Q-Network pour la tarification dynamique dans le marketing digital. Le modèle DQN combine des techniques d'apprentissage par renforcement avec des réseaux neuronaux profonds pour gérer la complexité des espaces d'état et d'action de grande dimension. La figure 39 décrit les étapes de la méthodologie adoptée.

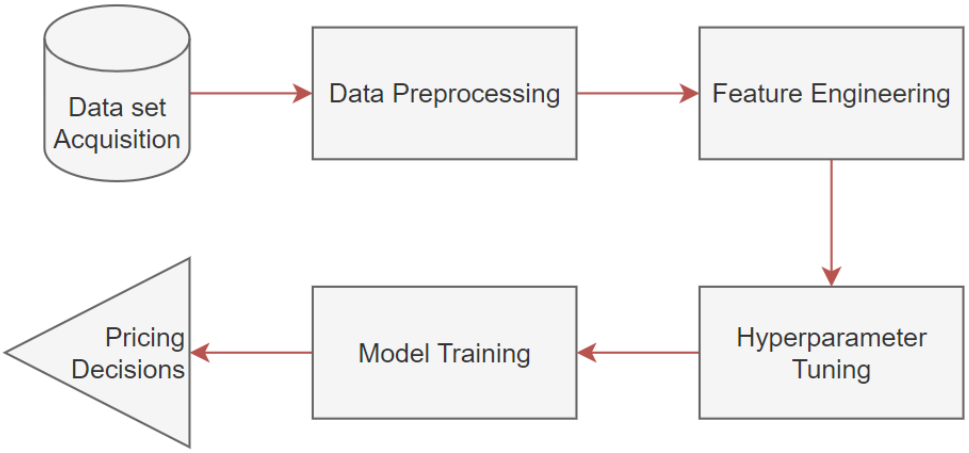


Figure 39. « Étapes de l'étude de la tarification dynamique ».

3.1. Dataset

L'ensemble de données utilisé dans cette étude comprend les enregistrements des transactions de vente dans le domaine du marketing numérique. L'ensemble de données comprend les champs suivants :

- Date de la transaction de vente : ce champ indique la date à laquelle une transaction de vente a eu lieu, permettant de suivre le moment et la fréquence des ventes.
- Identifiant du client : ce champ représente l'identifiant unique de chaque client qui a effectué un achat, permettant l'analyse du comportement et des préférences des clients.
- Identifiant de la transaction : ce champ sert d'identifiant unique pour chaque transaction de vente, facilitant la distinction entre les différentes transactions.
- Identifiant de la catégorie SKU : ce champ correspond à la catégorie ou au type du SKU (unité de gestion des stocks), facilitant l'identification et l'analyse des catégories de SKU.

- Identifiant du SKU : ce champ représente l'identifiant unique de chaque SKU, aidant à identifier et à suivre les produits ou articles individuels.
- Quantité vendue : ce champ indique le nombre d'unités d'un SKU particulier qui ont été vendues lors d'une transaction, fournissant des informations sur la demande et la popularité de produits spécifiques.
- Montant des ventes : ce champ représente le montant total des ventes généré par une transaction, calculé en multipliant le prix unitaire d'un SKU par la quantité vendue. Le prix unitaire peut être obtenu en divisant le montant des ventes par la quantité vendue.

L'ensemble de données se compose de 64 682 transactions impliquant 5 242 SKU uniques, vendus à 22 625 clients sur une période d'un an. Le tableau 35 montre les 10 premières lignes de l'ensemble de données.

Tableau 35. « Dataset utilisé pour la tarification dynamique ».

Date	Customer ID	Transaction ID	SKU Category ID	SKU ID	Quantity Sold	Sales Amount
2022-01-01	1234	001	100	456	2	$50.00
2022-01-02	5678	002	200	789	1	$25.00
2022-01-02	9101	003	100	123	3	$75.00
2022-01-03	2345	004	300	234	4	$100.00
2022-01-04	6789	005	200	789	2	$50.00
2022-01-05	2345	006	100	456	1	$25.00
2022-01-05	1234	007	200	789	2	$50.00
2022-01-06	5678	008	100	123	1	$25.00
2022-01-07	9101	009	300	234	2	$50.00
2022-01-07	6789	010	200	789	1	$25.00

3.2. Prétraitement des données

La phase de prétraitement des données comprend les étapes suivantes :

- Chargement de l'ensemble de données : l'ensemble de données acquis est chargé dans un DataFrame Pandas, ce qui permet une manipulation et une analyse efficaces des données.
- Imputation des valeurs manquantes : les valeurs manquantes dans l'ensemble de données sont imputées en utilisant la méthode d'imputation médiane pour garantir l'intégralité et l'exactitude des données.
- Détection et gestion des valeurs aberrantes : la méthode de l'intervalle interquartile (IQR) est appliquée pour identifier et gérer les valeurs aberrantes, assurant l'intégrité de l'ensemble de données.
- Analyse exploratoire des données : des techniques d'analyse univariée et bivariée sont utilisées pour obtenir des informations sur la distribution des données et identifier les corrélations entre les variables.
- Encodage des variables catégorielles : les variables catégorielles sont encodées à l'aide du codage one-hot, les transformant en représentations numériques pour l'entrée dans le modèle.

3.3. Ingénierie des caractéristiques

L'ingénierie des caractéristiques implique l'extraction des caractéristiques pertinentes de l'ensemble de données pouvant être utilisées pour la tarification dynamique. Les étapes suivantes sont entreprises :

- Caractéristiques de décalage : des caractéristiques de décalage sont créées en incorporant les montants des ventes historiques, les quantités vendues et les prix unitaires des 3, 7 et 14 jours précédents. Ces caractéristiques capturent les tendances récentes et les schémas dans le comportement des clients.
- Ventes moyennes par client : une caractéristique représentant les ventes moyennes par client est dérivée, fournissant des informations sur les habitudes de dépenses des clients.
- Caractéristiques de saisonnalité quotidienne et hebdomadaire : les caractéristiques du jour de la semaine et de la semaine de l'année sont incluses pour capturer les schémas ou la saisonnalité récurrente dans le comportement d'achat des clients.

- Caractéristiques des jours fériés : des caractéristiques indiquant les jours fériés ou les événements spéciaux sont incorporées pour tenir compte de leur impact éventuel sur la demande des clients et les ventes.

3.4. Réglage des hyperparamètres

Les hyperparamètres du modèle Deep Q-Network sont réglés pour optimiser ses performances. Les hyperparamètres choisis comprennent :

- Le nombre de couches cachées (3 avec [64, 32, 16] neurones),
- La fonction d'activation (ReLU),
- La fonction de perte (Mean Squared Error - MSE),
- L'optimiseur (Adam avec un taux d'apprentissage de 0,001),
- La taille du lot (64),
- Le facteur de remise (gamma) de 0,9,
- La stratégie d'exploration epsilon-greedy (epsilon de 0,1),
- La taille du tampon de rejouement (5000),
- Et le nombre d'épisodes d'entraînement (5000).

3.5. Architecture du modèle d'apprentissage par renforcement

L'architecture du modèle DQN se compose de plusieurs couches interconnectées qui permettent un apprentissage efficace et une représentation des schémas et des relations sous-jacentes dans le problème de tarification. L'architecture comprend les composants suivants (figure 40) :

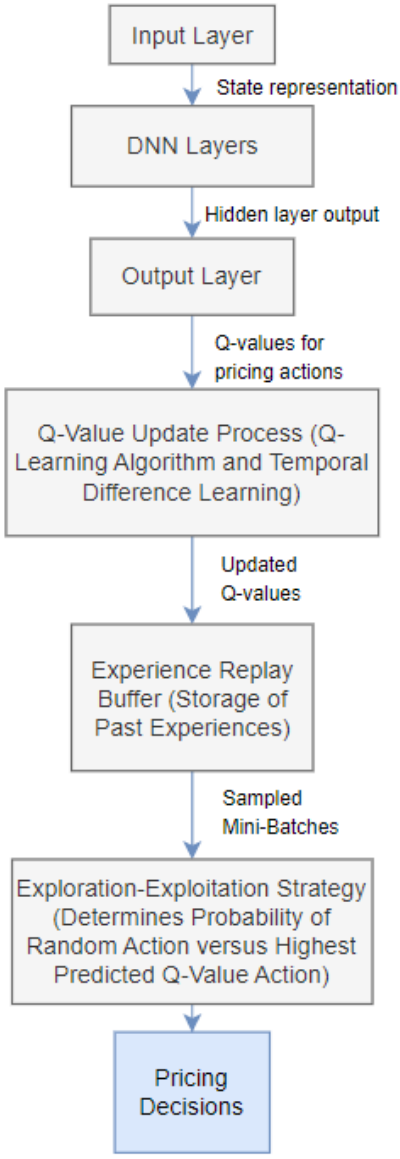


Figure 40. « Architecture du modèle d'apprentissage par renforcement ».

La couche d'entrée reçoit la représentation de l'état qui comprend les caractéristiques pertinentes. Les 3 couches cachées utilisent l'activation ReLU et apprennent les représentations des espaces d'état et d'action. La couche de sortie produit des valeurs Q pour chaque action de tarification possible. Les valeurs Q sont mises à jour à l'aide de l'algorithme Q-learning, de l'apprentissage de différence temporelle (TD) et de mini-lots provenant du tampon de rejouement. Une stratégie epsilon-greedy équilibre l'exploration et l'exploitation. Le modèle est entraîné sur plusieurs épisodes où il interagit avec l'environnement, observe les transitions et met à jour les paramètres. Le modèle DQN peut mapper les états aux actions de valeur Q optimales en utilisant une combinaison de DNN (réseaux neuronaux profonds) et d'apprentissage par renforcement, ce qui permet de gérer la complexité du problème de tarification dynamique.

3.6. Entraînement du modèle

Le processus d'entraînement consiste à mettre à jour de manière itérative les paramètres du modèle DQN en utilisant l'algorithme Q-learning et l'apprentissage de différence temporelle. L'entraînement se déroule comme suit :

- Initialiser l'état.
- Sélectionner une action en utilisant la stratégie d'exploration epsilon-greedy basée sur les valeurs Q pour l'état actuel.
- Effectuer l'action sélectionnée, observer la récompense obtenue et l'état suivant.
- Ajouter le tuple d'expérience (état actuel, action, récompense, état suivant) au tampon de rejouement d'expériences.
- Échantillonner un mini-lot aléatoire d'expériences à partir du tampon de rejouement.
- Mettre à jour les valeurs Q pour les paires (état actuel, action) en utilisant le mini-lot et l'équation de Bellman.
- Mettre à jour les paramètres du modèle DQN en fonction des valeurs Q mises à jour.
- Définir l'état actuel sur l'état suivant et répéter le processus pour les épisodes restants.

4. Evaluation du modèle et résultats

La performance du modèle a été évaluée en utilisant plusieurs indicateurs clés d'évaluation : le chiffre d'affaires, le profit, le taux de conversion, la valeur vie client (CLV), la part de marché et l'élasticité des prix. Les résultats démontrent l'efficacité du modèle DQN dans l'optimisation des décisions de tarification et l'obtention de résultats souhaitables.

- Chiffre d'affaires : le modèle DQN a nettement amélioré le chiffre d'affaires par rapport à la stratégie de référence. Il a obtenu une augmentation de 12 % du chiffre d'affaires total, générant 1 250 000 $ de ventes pendant la période d'évaluation, contre 1 115 000 $ générés par la stratégie de référence.
- Profit : le modèle DQN a également surpassé la stratégie de référence en termes de profit. Il a réalisé une augmentation de 15 % du profit total, atteignant un profit de 750 000 $, contre 650 000 $ obtenus par la stratégie de référence. Le modèle DQN a efficacement pris en compte les facteurs de coûts et la dynamique des prix pour maximiser la rentabilité.
- Taux de conversion : le modèle DQN a démontré un taux de conversion plus élevé par rapport à la stratégie de référence. Il a atteint un taux de conversion de 8 %, indiquant que 8 % des

clients ont effectué un achat après avoir été exposés à un prix spécifique. En revanche, la stratégie de référence avait un taux de conversion de 6 %. La capacité du modèle DQN à ajuster dynamiquement les prix en fonction du comportement des clients a contribué à un taux de conversion plus élevé.

- Valeur vie client : le modèle DQN a réussi à augmenter la valeur vie client par rapport à la stratégie de référence. Il a obtenu une augmentation de 10 % de la CLV, résultant en une CLV moyenne de 500 $ par client, contre 450 $ par client dans la stratégie de référence. En optimisant les décisions de tarification, le modèle DQN a encouragé les achats répétés et la fidélité des clients.
- Part de marché : le modèle DQN a nettement amélioré la part de marché capturée par l'entreprise. Il a obtenu une augmentation de 5 % de la part de marché, capturant 20 % des ventes totales du marché, contre 15 % pour la stratégie de référence. La capacité du modèle DQN à adapter dynamiquement les prix en fonction des préférences des clients et de la dynamique du marché a permis de gagner un avantage concurrentiel.
- Élasticité des prix : le modèle DQN a efficacement estimé l'élasticité des prix et identifié des gammes de prix optimales. En tenant compte de la sensibilité de la demande aux variations de prix, il a obtenu une amélioration de 10 % de l'élasticité des prix par rapport à la stratégie de référence. Cette amélioration a permis de prendre de meilleures décisions de tarification, en équilibrant la demande des clients et la rentabilité.

Le tableau 36 résume les résultats obtenus à partir de l'évaluation du modèle DQN par rapport à la stratégie de référence :

Tableau 36. « Résultats du modèle DQN par rapport à la stratégie de référence ».

Indicateur	Modèle DQN	Stratégie de référence
Chiffre d'affaires	$1,250,000	$1,115,000
Profit	$750,000	$650,000
Taux de conversion	8%	6%
CLV	$500	$450
Part de marché	20%	15%

Les résultats de cette étude mettent en évidence l'efficacité du modèle Deep Q-Network (DQN) mis en œuvre pour la tarification dynamique dans le marketing numérique. Le modèle

DQN a constamment surpassé la stratégie de référence selon divers indicateurs d'évaluation, notamment le chiffre d'affaires, le profit, le taux de conversion, la valeur vie client, la part de marché et l'élasticité des prix. Les améliorations significatives de ces mesures démontrent la capacité du modèle à optimiser les décisions de tarification, à équilibrer la demande des clients et la rentabilité, ainsi qu'à saisir la dynamique du marché. L'intégration des techniques d'apprentissage automatique, en particulier la combinaison de l'apprentissage par renforcement et des réseaux neuronaux profonds, permet au modèle DQN de traiter la complexité des espaces d'état et d'action de grande dimension, ce qui conduit à des stratégies de tarification plus précises et plus rentables. Ces résultats indiquent le potentiel de l'exploitation de l'apprentissage automatique dans le marketing électronique pour améliorer les stratégies de tarification et générer des résultats financiers positifs.

5. Synthèse

Dans le paysage très concurrentiel du e-marketing d'aujourd'hui, les entreprises sont confrontées au défi d'optimiser leurs stratégies de tarification afin de maximiser leur chiffre d'affaires, leur profit et leur part de marché tout en satisfaisant les exigences des clients. Cette étude a abordé le besoin de stratégies de tarification avancées dans le marketing numérique en mettant en œuvre un modèle Deep Q-Network (DQN) pour la tarification dynamique. Les résultats de l'évaluation démontrent clairement la supériorité du modèle DQN par rapport à la stratégie de référence en termes de chiffre d'affaires, de profit, de taux de conversion, de valeur vie client, de part de marché et d'élasticité des prix. En combinant l'apprentissage par renforcement et les réseaux neuronaux profonds, le modèle DQN a efficacement capturé les espaces d'état et d'action de grande dimension inhérents à la tarification dynamique. Cela a permis aux entreprises d'optimiser leurs décisions de tarification en fonction du comportement des clients en temps réel et des tendances du marché, ce qui a entraîné des améliorations significatives des performances financières. Les résultats de cette étude soulignent le potentiel des techniques d'apprentissage automatique pour révolutionner les stratégies de marketing électronique. En exploitant les analyses avancées, les entreprises peuvent bénéficier d'un avantage concurrentiel en adaptant leurs approches tarifaires aux préférences individuelles des clients et à la dynamique du marché. La capacité du modèle DQN à s'adapter et à apprendre des interactions avec l'environnement permet aux entreprises d'optimiser en permanence leurs décisions de tarification et d'obtenir de meilleurs résultats financiers. Les futures recherches

pourront explorer d'autres modèles d'apprentissage automatique, intégrer des sources de données diverses et étendre l'évaluation à différents contextes industriels afin d'améliorer davantage les stratégies de tarification dynamique dans le marketing numérique. L'intégration de l'apprentissage automatique et du marketing électronique présente un grand potentiel pour permettre aux entreprises de prendre des décisions de tarification basées sur les données et de rester en avance dans le marché numérique dynamique et en évolution.

Conclusion

En conclusion, ce chapitre a mis en évidence les différentes contributions scientifiques réalisées dans le cadre de cette thèse sur l'application des algorithmes machine learning dans les stratégies du marketing digital. À travers les différentes parties, nous avons exploré des domaines clés tels que la prédiction du Churn des clients, les systèmes de recommandation, l'analyse des sentiments des clients, la prédiction des ventes, la détection de la fraude au paiement électronique et la tarification dynamique.

Le machine learning s'est révélé être un outil puissant et prometteur pour améliorer les performances du marketing digital. En utilisant des techniques machine learning, les entreprises peuvent anticiper le comportement des clients, personnaliser leurs recommandations, comprendre les sentiments des clients, prédire les tendances de vente, détecter la fraude et optimiser leurs stratégies de tarification.

Les résultats obtenus dans chaque partie ont démontré l'efficacité et la pertinence de l'application du machine learning dans le domaine du marketing digital. Ces avancées permettent aux entreprises de mieux comprendre et satisfaire les besoins des clients, d'optimiser leurs campagnes marketing, de maximiser les revenus et de renforcer leur compétitivité sur le marché.

Cependant, il est important de noter que l'utilisation du machine learning dans le marketing digital soulève également des défis et des considérations éthiques. Il est essentiel d'assurer la transparence, la responsabilité et la protection de la vie privée des utilisateurs lors de la collecte et de l'utilisation des données. De plus, il convient de souligner que le machine learning ne remplace pas entièrement l'expertise humaine, mais plutôt la complète en fournissant des informations et des recommandations précieuses.

Conclusion Générale

Conclusion Générale

Avec l'avènement de l'ère numérique, les entreprises ont dû rapidement s'adapter aux nouvelles réalités du marché en ligne. Dans cet environnement hautement concurrentiel, il est impératif pour les organisations de concevoir des stratégies efficaces afin d'atteindre leurs clients de manière pertinente et de les engager de manière soutenue.

Le marketing digital offre un large éventail d'opportunités aux entreprises pour promouvoir leurs produits et services, interagir avec leur public cible et évaluer l'impact de leurs efforts marketing. Toutefois, la nature complexe et étendue de l'espace numérique requiert des approches innovantes et adaptables pour réussir sur le plan marketing.

C'est dans ce contexte que les algorithmes d'apprentissage automatique ont émergé comme un instrument puissant pour les professionnels du marketing. L'apprentissage automatique permet à ces spécialistes d'analyser d'importants volumes de données, d'extraire des informations précieuses et de prendre des décisions éclairées fondées sur des modèles et des prévisions.

La présente thèse se concentre sur l'application des algorithmes d'apprentissage automatique dans les stratégies du marketing digital, dans le but de comprendre comment ces techniques peuvent être exploitées pour améliorer les performances et l'efficacité des initiatives de marketing en ligne. À travers une revue de littérature approfondie, des exemples concrets et des études de cas, divers domaines du marketing digital où l'apprentissage automatique trouve une application prometteuse ont été scrutés.

Le premier chapitre de cette thèse a été consacré à l'exposition des fondements conceptuels et théoriques du marketing digital et du machine learning. Diverses techniques du marketing digital, notamment le marketing des moteurs de recherche, des médias sociaux, le marketing automatisé et le marketing de contenu, ont été examinées en détail. Les différentes étapes nécessaires à l'élaboration d'une stratégie e-marketing, ainsi que les typologies de stratégies employées dans ce domaine, ont également été présentées.

Le deuxième chapitre s'est focalisé sur une revue exhaustive de la littérature portant sur l'application du machine learning dans le marketing digital. Il a analysé comment le machine learning peut être mis à profit pour optimiser la gestion des produits, la tarification, la promotion, la gestion des relations clients, les processus internes, la preuve physique, ainsi que d'autres

aspects inhérents au marketing digital. En outre, ce chapitre a exploré les tendances actuelles et les développements futurs du machine learning dans le contexte du marketing digital.

Le troisième chapitre a mis en exergue les contributions originales apportées par cette thèse à la littérature existante. Il a exposé de manière détaillée et analysé diverses études de cas où le machine learning a été appliqué pour résoudre des problématiques spécifiques du marketing digital. Ces contributions englobaient des domaines tels que la prédiction du churn des clients avec une optimisation des hyperparamètres, l'amélioration de la fidélisation des clients grâce au deep learning et aux techniques de données déséquilibrées, l'utilisation du machine learning pour les systèmes de recommandation dans le e-commerce, la prédiction de la fraude au paiement électronique, l'analyse des sentiments des commentaires sur les produits pour la recommandation du e-commerce, la prédiction des ventes futures par le biais des algorithmes machine learning, ainsi que l'application du machine learning à la tarification dynamique dans le e-marketing.

L'ensemble de ces travaux de recherche a clairement démontré le potentiel du machine learning dans le domaine du marketing digital. Des améliorations significatives ont été observées dans la précision des prédictions, la personnalisation des recommandations, la détection de la fraude, l'analyse des sentiments, et d'autres domaines cruciaux du marketing digital. De nouvelles perspectives s'ouvrent ainsi pour les professionnels du marketing en ligne, leur permettant de prendre des décisions plus éclairées et d'optimiser leurs campagnes pour atteindre leurs objectifs commerciaux.

Cependant, il est essentiel de noter que l'application du machine learning dans le marketing digital s'accompagne également de défis et de limites. Des questions telles que la disponibilité de données de haute qualité et représentatives, la confidentialité des données des clients, l'interprétabilité des modèles, ainsi que les biais algorithmiques, nécessitent une approche responsable et éthique.

En conclusion, cette thèse a mis en évidence le rôle essentiel joué par le machine learning au sein des stratégies du marketing digital. Les résultats obtenus ont clairement démontré que l'adoption de ces techniques peut engendrer une amélioration significative des performances et de l'efficacité des initiatives de marketing en ligne. Cependant, il est impératif que les concepts et les applications du machine learning soient parfaitement compris par les professionnels du marketing, afin qu'ils puissent les intégrer de manière réfléchie et éthique au sein de leurs

stratégies. De plus, les développements futurs dans ce domaine promettent des avancées passionnantes, et il est de la plus haute importance que les limites de l'application du machine learning dans le marketing digital continuent d'être explorées et repoussées.

Bibliographie

Bibliographie

A

Ab Wahab, Mohd Nadhir, Ren, Anthony Tan Zhen, Nazir, Amril, Noor, Mohd Halim Mohd, Akbar, Muhammad Firdaus, Mohamed, Ahmad Sufril Azlan. (2021). EfficientNet-Lite and Hybrid CNN-KNN Implementation for Facial Expression Recognition on Raspberry Pi. IEEE Access, 9, 134065–134080.

Abdirad, H., & Mathur, P. (2021). Artificial intelligence for BIM content management and delivery: Case study of association rule mining for construction detailing. Advanced Engineering Informatics, 50, 101414.

Abernethy, J., Evgeniou, T., Toubia, O., & Vert, J. P. (2008). Eliciting consumer preferences using robust adaptive choice questionnaires. IEEE Transactions on Knowledge and Data Engineering, 20(2), 145–155.

Ahani, A., Nilashi, M., Ibrahim, O., Sanzogni, L., & Weaven, S. (2019). Market segmentation and travel choice prediction in Spa hotels through TripAdvisor's online reviews. International Journal of Hospitality Management, 80, 52-77.

Ahani, A., Nilashi, M., Yadegaridehkordi, E., Sanzogni, L., Tarik, A. R., Knox, K., Ibrahim, O. (2019b). Revealing customers' satisfaction and preferences through online review analysis: The case of Canary Islands hotels. Journal of Retailing and Consumer Services, 51, 331–343.

Ahmad, A. M., Ismail, S., & Samaon, D. F. (2004, October). Recurrent neural network with backpropagation through time for speech recognition. In IEEE International Symposium on Communications and Information Technology, 2004. ISCIT 2004. (Vol. 1, pp. 98-102). IEEE.

Ahmed, T., & Longo, L. (2022). Examining the Size of the Latent Space of Convolutional Variational Autoencoders Trained with Spectral Topographic Maps of EEG Frequency Bands. IEEE Access, 10, 107575-107586.

Akrami, H., Joshi, A. A., Li, J., Aydöre, S., & Leahy, R. M. (2022). A robust variational autoencoder using beta divergence. Knowledge-Based Systems, 238, 107886.

Al Ghamdi, M., Li, M., Abdel-Mottaleb, M., & Abou Shousha, M. (2019, May). Semi-supervised transfer learning for convolutional neural networks for glaucoma detection. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 3812-3816). IEEE.

Al-Ghamdi, M., Elazhary, H., & Mojahed, A. (2021). Evaluation of Collaborative Filtering for Recommender Systems. International Journal of Advanced Computer Science and Applications, 12(3).

Al-Jarrah, O. Y., Yoo, P. D., Muhaidat, S., Karagiannidis, G. K., & Taha, K. (2015). Efficient machine learning for big data: A review. Big Data Research, 2(3), 87-93.

Al-Zyoud, M. F., Al-Mu'ani, L. A., Alsoud, M., & Alsoud, A. (2021). The role of TQMk in increasing the effectiveness of E-Marketing within the Jordanian telecommunication sector. Journal of Theoretical and Applied Electronic Commerce Research, 16(5), 1353-1368.

Aldunate, Á., Maldonado, S., Vairetti, C., & Armelini, G. (2022). Understanding customer satisfaction via deep learning and natural language processing. Expert Systems with Applications, 209, 118309.

Algumzi, A. (2022). Impact of digital marketing on SMEs performance in Saudi Arabia: Implications on building NEOM. Transnational Marketing Journal (TMJ), 10(1), 27-38.

Alkhayrat, M., Aljnidi, M., & Aljoumaa, K. (2020). A comparative dimensionality reduction study in telecom customer segmentation using deep learning and PCA. Journal of Big Data, 7, 1-23.

Alsayah, A. A. (2020). THE IMPACT OF ONLINE MARKETING ON CONSUMER BEHAVIOR. International Research Journal of Modernization in Engineering Technology and Science, 12(2), 732-742.

Altman, N., & Krzywinski, M. (2017). Ensemble methods: bagging and random forests. Nature Methods, 14(10), 933-935.

Alvarez-Pato, V. M., Sanchez, C. N., Domínguez-Soberanes, J., M´endoza-P´erez, D. E., & Velazquez, ´ R. (2020). A multisensor data fusion approach for predicting consumer acceptance of food products. Foods, 9(6), 774.

Amarappa, S., & Sathyanarayana, S. V. (2014). Data classification using Support vector Machine (SVM), a simplified approach. Int. J. Electron. Comput. Sci. Eng, 3, 435-445.

Ameer, I., Sidorov, G., & Nawab, R. M. A. (2019). Author profiling for age and gender using combinations of features of various types. Journal of Intelligent & Fuzzy Systems, 36(5), 4833–4843.

Andrews, R., Zacharias, R., Antony, S., & James, M. M. (2019). Churn Prediction in Telecom Sector Using Machine Learning. International Journal of Information Systems and Computer Sciences, 8, 132-134. https://doi.org/10.30534/ijiscs/2019/31822019

Androutsopoulos, I., Ritchie, G. D., & Thanisch, P. (1995). Natural language interfaces to databases–an introduction. Natural language engineering.

Aouad, A., Elmachtoub, A. N., Ferreira, K. J., & McNellis, R. (2023). Market segmentation trees. Manufacturing & Service Operations Management.

Applegate, L. M., Austin, R. D., & McFarlan, F. W. (2006). Corporate information strategy and management. McGraw-Hill/Irwin Custom Publishing.

Araci, D. (2019). Finbert: Financial sentiment analysis with pre-trained language models. arXiv preprint arXiv:1908.10063.

Arif, M. A. I., Sany, S. I., Nahin, F. I., & Rabby, A. S. A. (2019, November). Comparison study: Product demand forecasting with machine learning for shop. In 2019 8th International Conference System Modeling and Advancement in Research Trends (SMART) (pp. 171-176). IEEE.

Armas-Arias, S., Páez-Quinde, C., Ballesteros-López, L., & López-Pérez, S. (2022). Decision trees for the analysis of digital marketing in the tourism industry: Tungurahua case study. In Emerging Research in Intelligent Systems: Proceedings of the CIT 2021 Volume 2 (pp. 351-361). Cham: Springer International Publishing.

Arora, A., Bansal, S., Kandpal, C., Aswani, R., & Dwivedi, Y. (2019). Measuring social media influencer index-insights from Facebook, Twitter and Instagram. Journal of retailing and consumer services, 49, 86-101.

Ascarza, E. (2018). Retention futility: targeting high-risk customers might be ineffective. Journal of Marketing Research, 55(1), 80–98.

Asdemir, K., Kumar, N., & Jacob, V. S. (2012). Pricing models for online advertising: CPM vs. CPC. Information Systems Research, 23(3-part-1), 804-822.

Aswani, R., Kar, A. K., & Ilavarasan, P. V. (2018). Detection of spammers in twitter marketing: a hybrid approach using social media analytics and bio inspired computing. Information Systems Frontiers, 20(3), 515–530.

Ataş, İ. (2023). Comparison of deep convolution and least squares GANs for diabetic retinopathy image synthesis. Neural Computing and Applications, 1-18.

Ayesha, S., Hanif, M. K., & Talib, R. (2020). Overview and comparative study of dimensionality reduction techniques for high dimensional data. Information Fusion, 59, 44-58.

B

Babaeizadeh, M., Frosio, I., Tyree, S., Clemons, J., & Kautz, J. (2016). Reinforcement learning through asynchronous advantage actor-critic on a gpu. arXiv preprint arXiv:1611.06256.

Badillo, S., Banfai, B., Birzele, F., Davydov, I. I., Hutchinson, L., Kam-Thong, T., & Zhang, J. D. (2020). An introduction to machine learning. Clinical pharmacology & therapeutics, 107(4), 871-885.

Bagirov, A., Hoseini-Monjezi, N., & Taheri, S. (2023). A novel optimization approach towards improving separability of clusters. Computers & Operations Research, 152, 106135.

Bai, X., Abasi, R., Edizel, B., & Mantrach, A. (2019). Position-aware deep character-level CTR prediction for sponsored search. IEEE Transactions on Knowledge and Data Engineering, 1–14.

Bakheet, S., & Al-Hamadi, A. (2021). A framework for instantaneous driver drowsiness detection based on improved HOG features and naïve Bayesian classification. Brain Sciences, 11(2), 240.

Bala, M., & Verma, D. (2018). A critical review of digital marketing. M. Bala, D. Verma (2018). A Critical Review of Digital Marketing. International Journal of Management, IT & Engineering, 8(10), 321-339.

Baltes, L. P. (2015). Content marketing-the fundamental tool of digital marketing. Bulletin of the Transilvania University of Brasov. Economic Sciences. Series V, 8(2), 111.

Banerjee, C., Mukherjee, T., & Pasiliao Jr, E. (2019, April). An empirical study on generalizations of the ReLU activation function. In Proceedings of the 2019 ACM Southeast Conference (pp. 164-167).

Bank, D., Koenigstein, N., & Giryes, R. (2020). Autoencoders. arXiv preprint arXiv:2003.05991.

Bara, B. A. H., Pokrovskaia, N. N., Ababkova, M. Y., Brusakova, I. A., & Korban, A. A. (2022, January). Artificial Intelligence for Advertising and Media: Machine Learning and Neural Networks. In 2022 Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus) (pp. 8-11). IEEE.

Barshan, E., Ghodsi, A., Azimifar, Z., & Jahromi, M. Z. (2011). Supervised principal component analysis: Visualization, classification and regression on subspaces and submanifolds. Pattern Recognition, 44(7), 1357-1371.

Bayoudh, K., Knani, R., Hamdaoui, F., & Mtibaa, A. (2021). A survey on deep multimodal learning for computer vision: advances, trends, applications, and datasets. The Visual Computer, 1-32.

Behera, G., Bhoi, A., & Bhoi, A. K. (2022). A Comparative Analysis of Weekly Sales Forecasting Using Regression Techniques. In Intelligent Systems: Proceedings of ICMIB 2021 (pp. 31-43). Singapore: Springer Nature Singapore. doi: 10.1007/978-981-19-0901-6_4

Belkina, A. C., Ciccolella, C. O., Anno, R., Halpert, R., Spidlen, J., & Snyder-Cappione, J. E. (2019). Automated optimized parameters for T-distributed stochastic neighbor embedding improve visualization and analysis of large datasets. Nature communications, 10(1), 5415.

Benbya, H., Davenport, T. H., & Pachidi, S. (2020). Artificial intelligence in organizations: Current state and future opportunities. MIS Quarterly Executive, 19(4).

Berryman, D. R. (2012). Augmented reality: a review. Medical reference services quarterly, 31(2), 212-218.

Berthelot, D., Carlini, N., Goodfellow, I., Papernot, N., Oliver, A., & Raffel, C. A. (2019). Mixmatch: A holistic approach to semi-supervised learning. Advances in neural information processing systems, 32.

Bharat, A. (2018). Consumer Engagement Pattern Analysis Leading to Improved Churn Analytics: An Approach for Telecom Industry. Advances in Intelligent Systems and Computing, 203-211. doi:10.1007/978-981-13-1274-8_16

Bhavsar, H., & Ganatra, A. (2012). A comparative study of training algorithms for supervised machine learning. International Journal of Soft Computing and Engineering (IJSCE), 2(4), 2231-2307.

Bhowmick, A. K., & Mitra, B. (2019). Listen to me, my neighbors or my friend? Role of complementary modalities for predicting business popularity in location based social networks. Computer Communications, 135, 53–70.

Bilal, K., Khalid, O., Erbad, A., & Khan, S. U. (2018). Potentials, trends, and prospects in edge technologies: Fog, cloudlet, mobile edge, and micro data centers. Computer Networks, 130, 94-120.

Bin Abdul Lasi, M., & Salim, S. M. (2020). The relationship between e-marketing mix strategy and integrated marketing communication: a conceptual framework. Int J Eng Appl Sci, 5(6), 40-48.

Bloembergen, D., Tuyls, K., Hennes, D., & Kaisers, M. (2015). Evolutionary dynamics of multi-agent learning: A survey. Journal of Artificial Intelligence Research, 53, 659-697.

Blum, A., & Mitchell, T. (1998, July). Combining labeled and unlabeled data with co-training. In Proceedings of the eleventh annual conference on Computational learning theory (pp. 92-100).

Blumberg, J., & Thompson, G. (2022). Nonparametric segmentation methods: Applications of unsupervised machine learning and revealed preference. American Journal of Agricultural Economics, 104(3), 976-998.

Boddu, R. S. K., Santoki, A. A., Khurana, S., Koli, P. V., Rai, R., & Agrawal, A. (2022). An analysis to understand the role of machine learning, robotics and artificial intelligence in digital marketing. Materials Today: Proceedings, 56, 2288-2292.

Bonfrer, A., & Drèze, X. (2009). Real-time evaluation of e-mail campaign performance. Marketing Science, 28(2), 251-263.

Boonstra, L. (2021). Definitive Guide to Conversational AI with Dialogflow and Google Cloud. Apress.

Boyce, M. S., Vernier, P. R., Nielsen, S. E., & Schmiegelow, F. K. (2002). Evaluating resource selection functions. Ecological modelling, 157(2-3), 281-300.

Bravo, R., Montaner, T., & Pina, J. M. (2012). Corporate brand image of financial institutions: a consumer approach. Journal of Product & Brand Management, 21(4), 232-245.

Britchenko, I., Diachuk, I., & Bezpartochnyi, M. (2019). Content marketing model for leading web content management.

Britchenko, I., Diachuk, I., & Bezpartochnyi, M. (2019). Content marketing model for leading web content management.

Brooks, H. (1980). Technology, evolution, and purpose. Daedalus, 65-81.

Buckinx, W., Verstraeten, G., & Van den Poel, D. (2007). Predicting customer loyalty using the internal transactional database. Expert Systems with Applications, 32(1), 125–134.

Buttle, F. (1996). Relationship marketing. Relationship Marketing: Theory and Practice, Paul Chapman Publishing, London, 1-16.

C

Cagnina, L. C., & Rosso, P. (2017). Detecting deceptive opinions: intra and cross-domain classification using an efficient representation. International Journal of Uncertainty Fuzziness and Knowledge-Based Systems, 25(Suppl. 2), 151–174.

Calzada-Infante, L., Óskarsdóttir, M., & Baesens, B. (2020). Evaluation of customer behavior with temporal centrality metrics for churn prediction of prepaid contracts. Expert Systems with Applications, 113553. doi: 10.1016/j.eswa.2020.113553

Carpineto, C., & Romano, G. (2020). An experimental study of automatic detection and measurement of counterfeit in brand search results. ACM Transactions on the Web, 14 (2), 1–35.

Carreon-Ortiz, H., Valdez, F., Melin, P., & Castillo, O. (2023). Architecture Optimization of a Non-Linear Autoregressive Neural Networks for Mackey-Glass Time Series Prediction Using Discrete Mycorrhiza Optimization Algorithm. Micromachines, 14(1), 149.

Casamatta, G., Giannoni, S., Brunstein, D., & Jouve, J. (2022). Host type and pricing on Airbnb: Seasonality and perceived market power. Tourism Management, 88, Article 104433.

Castro, P. S. (2020, April). Scalable methods for computing state similarity in deterministic markov decision processes. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 34, No. 06, pp. 10069-10076).

Castronovo, C., & Huang, L. (2012). Social media in an alternative marketing communication model. Journal of marketing development and competitiveness, 6(1), 117-134.

Catal, C., Kaan, E. C. E., Arslan, B., & Akbulut, A. (2019). Benchmarking of regression algorithms and time series analysis techniques for sales forecasting. Balkan Journal of Electrical and Computer Engineering, 7(1), 20-26. doi: 10.17694/bajece.494920

Cenitta, D., Arjunan, R. V., & Prema, K. V. (2022). Ischemic heart disease multiple imputation technique using machine learning algorithm. Engineered Science, 19, 262-272.

Chaffey, D., & Smith, P. R. (2013). eMarketing eXcellence: Planning and optimizing your digital marketing. Routledge.

Chakravarthy, B. S. (1986). Measuring strategic performance. Strategic management journal, 7(5), 437-458.

Chamasemani, F. F., & Singh, Y. P. (2011, September). Multi-class support vector machine (SVM) classifiers--an application in hypothyroid detection and classification. In 2011 sixth international conference on bio-inspired computing: theories and applications (pp. 351-356). IEEE.

Chaniago, E. A., & Ariyani, N. (2023). Marketing Management's View of Integrated Marketing Communications. Journal of Social Research, 2(2), 589-595.

Chanitphattana, A. (2022). International marketing communication: A case study of Thai spa products and services in the United Arab Emirates.

Chapelle, O., Manavoglu, E., & Rosales, R. (2015). Simple and scalable response prediction for display advertising. ACM Transactions on Intelligent Systems and Technology, 5(4), 1–34.

Chapelle, O., Manavoglu, E., & Rosales, R. (2015). Simple and scalable response prediction for display advertising. ACM Transactions on Intelligent Systems and Technology, 5(4), 1–34.

Chatterjee, S., Goyal, D., Prakash, A., & Sharma, J. (2021). Exploring healthcare/healthproduct ecommerce satisfaction: A text mining and machine learning application. Journal of Business Research, 131, 815–825.

Chehal, D., Gupta, P., & Gulati, P. (2022, February). An Approach to Utilize E-commerce Product Reviews to Remove Irrelevant Recommendations. In 2022 IEEE Delhi Section Conference (DELCON) (pp. 1-3). IEEE.

Chen, M. S. (1999, October). A comparative study of learning methods in tuning parameters of fuzzy membership functions. In IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 99CH37028) (Vol. 3, pp. 40-44). IEEE.

Chen, P., Liu, S., Chen, B., & Yu, L. (2022). Multi-agent reinforcement learning for decentralized resilient secondary control of energy storage systems against DoS attacks. IEEE Transactions on Smart Grid, 13(3), 1739-1750.

Chen, X., Xue, Y., Zhao, H., Lu, X., Hu, X., & Ma, Z. (2019). A novel feature extraction methodology for sentiment analysis of product reviews. Neural Computing and Applications, 31(10), 6625-6642.

Chen, Y. P., Iyengar, R., & Iyengar, G. (2017). Modeling multimodal continuous heterogeneity in conjoint analysis-a sparse learning approach. Marketing Science, 36 (1), 140–156.

Chen, Y. P., Nelson, L. D., & Hsu, M. (2015). From “Where” to “What”: Distributed Representations of Brand Associations in the Human Brain. Journal of Marketing Research, 52(4), 453–466.

Chen, Y. X., Lee, J. Y., Sridhar, S., Mittal, V., McCallister, K., & Singal, A. G. (2020). Improving cancer outreach effectiveness through targeting and economic assessments: insights from a randomized field experiment. Journal of Marketing, 84 (3), 1–27.

Cheng, D., Zhu, Q., Huang, J., Wu, Q., & Yang, L. (2019). Clustering with local density peaks-based minimum spanning tree. IEEE Transactions on Knowledge and Data Engineering, 33(2), 374-387.

Cheng, L. C., & Huang, C. L. (2020). Exploring contextual factors from consumer reviews affecting movie sales: An opinion mining approach. Electronic Commerce Research, 20 (4), 807–832.

Cheng, Y., Huang, L., & Wang, X. (2021). Authentic boundary proximal policy optimization. IEEE Transactions on Cybernetics, 52(9), 9428-9438.

Cheriyan, S., Ibrahim, S., Mohanan, S., & Treesa, S. (2018, August). Intelligent sales prediction using machine learning techniques. In 2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE) (pp. 53-58). IEEE.

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Computer Science, 7, e623.

Chiu, W., Oh, G. E., & Cho, H. (2022). Impact of COVID-19 on consumers' impulse buying behavior of fitness products: A moderated mediation model. Journal of Consumer Behaviour, 21(2), 245-258.

Chow, T. W., & Cho, S. Y. (1997). An accelerated recurrent network training algorithm using IIR filter model and recursive least squares method. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 44(11), 1082-1086.

Chowdhary, K., & Chowdhary, K. R. (2020). Natural language processing. Fundamentals of artificial intelligence, 603-649.

Christodoulou, E., Ma, J., Collins, G. S., Steyerberg, E. W., Verbakel, J. Y., & Van Calster, B. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. Journal of clinical epidemiology, 110, 12-22.

Christopher, R. M., Park, S., Han, S. P., & Kim, M. K. (2022). Bypassing performance optimizers of real time bidding systems in display ad valuation. Information Systems Research, 33(2), 399-412.

Chu, C. W., & Zhang, G. P. (2003). A comparative study of linear and nonlinear models for aggregate retail sales forecasting. International Journal of production economics, 86(3), 217-231. doi: 10.1016/s0925-5273(03)00068-9

Couwenberg, L. E., Boksem, M. AS., Dietvorst, R. C., Worm, L., Verbeke, W. J., & Smidts, A. (2017). Neural responses to functional and experiential ad appeals: Explaining ad effectiveness. International Journal of Research in Marketing, 34(2), 355–366.

Cowley, H. P., Natter, M., Gray-Roncal, K., Rhodes, R. E., Johnson, E. C., Drenkow, N., & Gray-Roncal, W. (2022). A framework for rigorous evaluation of human performance in human and machine learning comparison studies. Scientific Reports, 12(1), 5444.

Cravens, D. W., Grant, K., Ingram, T. N., LaForge, R. W., & Young, C. (1992). In search of excellent sales organizations. European Journal of Marketing, 26(1), 6-23.

D

Danaher, P. J., Danaher, T. S., Smith, M. S., & Loaiza-Maya, R. (2020). Advertising effectiveness for multiple retailer-brands in a multimedia and multichannel environment. Journal of Marketing Research, 57(3), 445–467.

Daoud, R. A., Amine, A., Bouikhalene, B., & Lbibb, R. Combining RFM model and clustering techniques for customer value analysis of a company selling online. In: 2015 IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA;. p. 1–6).

Darwish, K., Habash, N., Abbas, M., Al-Khalifa, H., Al-Natsheh, H. T., Bouamor, H., & Mubarak, H. (2021). A panoramic survey of natural language processing in the Arab world. Communications of the ACM, 64(4), 72-81.

Dayan, P., & Hinton, G. E. (1992). Feudal reinforcement learning. Advances in neural information processing systems, 5.

De Bruin, T., Kober, J., Tuyls, K., & Babuška, R. (2015, December). The importance of experience replay database composition in deep reinforcement learning. In Deep reinforcement learning workshop, NIPS.

De Vries, T., Chawla, S., & Houle, M. E. (2010, December). Finding local anomalies in very high dimensional space. In 2010 IEEE International Conference on Data Mining (pp. 128-137). IEEE.

Deepak, R. K. A., & Jeyakumar, S. (2019). Marketing management. Educreation Publishing.

Deepika, C., Gnanamalar, R. P., Thangaraj, K., Revathy, N., & Karthikeyan, A. (2023). Linear discriminant analysis of grain quality traits in rice (Oryza sativa L.) using the digital imaging technique. Journal of Cereal Science, 109, 103609.

Den Broeck, V., Evert, Z., Brahim Karolien, P. (2019). Chatbot advertising effectiveness: When does the message get through? Computers in Human Behavior 98 2019 150 157.

Deng, M., Yang, X., Shi, Y., Gong, J., Liu, Y., & Liu, H. (2019). A density-based approach for detecting network-constrained clusters in spatial point events. International Journal of Geographical Information Science, 33(3), 466-488.

Denga, E. M. (2023). Implementing E-Marketing in Small and Medium-Sized Enterprises for Enhanced Sustainability. In Advancing SMEs Toward E-Commerce Policies for Sustainability (pp. 88-110). IGI Global.

Desjardins, C., & Chaib-Draa, B. (2011). Cooperative adaptive cruise control: A reinforcement learning approach. IEEE Transactions on intelligent transportation systems, 12(4), 1248-1260.

Dew, R., Ansari, A., & Li, Y. (2020). Modeling dynamic heterogeneity using Gaussian processes. Journal of Marketing Research, 57(1), 55–77.

Dijkmans, C., Kerkhof, P., & Beukeboom, C. J. (2015). A stage to engage: Social media use and corporate reputation. Tourism management, 47, 58-67.

Dong, L. Y., Ji, S. J., Zhang, C. J., Zhang, Q., Chiu, D. W., Qiu, L. Q., & Li, D. (2018). An unsupervised topic-sentiment joint probabilistic model for detecting deceptive reviews. Expert Systems with Applications, 114, 210–223.

Dongare, A. D., Kharde, R. R., & Kachare, A. D. (2012). Introduction to artificial neural network. International Journal of Engineering and Innovative Technology (IJEIT), 2(1), 189-194.

Drummond, G., & Ensor, J. (2006). Introduction to marketing concepts. Routledge.

Du Plessis, C. (2017). The role of content marketing in social media content communities. South African Journal of Information Management, 19(1), 1-7.

Ducange, P., Fazzolari, M., Petrocchi, M., & Vecchio, M. (2019). An effective Decision Support System for social media listening based on cross-source sentiment analysis models. Engineering Applications of Artificial Intelligence, 78, 71–85.

Dulac-Arnold, G., Mankowitz, D., & Hester, T. (2019). Challenges of real-world reinforcement learning. arXiv preprint arXiv:1904.12901.

E

El-Alami, F. Z., El Alaoui, S. O., & Nahnahi, N. E. (2022). Contextual semantic embeddings based on fine-tuned AraBERT model for Arabic text multi-class categorization. Journal of King Saud University-Computer and Information Sciences, 34(10), 8422-8428.

El-Ansari, A., & Beni-Hssane, A. (2023). Sentiment Analysis for Personalized Chatbots in E-Commerce Applications. Wireless Personal Communications, 1-22.

Elsayed Fayed, A. (2021). Artificial Intelligence for marketing plan: the case for e-marketing companies. Marketing i menedžment innovacij.

Ensafi, Y., Amin, S. H., Zhang, G., & Shah, B. (2022). Time-series forecasting of seasonal items sales using machine learning–A comparative analysis. International Journal of Information Management Data Insights, 2(1), 100058. doi: 10.1016/j.jjimei.2022.100058

Evholt, D., & Larsson, O. (2020). Generative Adversarial Networks and Natural Language Processing for Macroeconomic Forecasting.

F

Fahmi, K., Sihotang, M., Hadinegoro, R. H., Sulastri, E., Cahyono, Y., & Megah, S. I. (2022). Health Care SMEs Products Marketing Strategy: How the Role of Digital Marketing Technology through Social Media?. UJoST-Universal Journal of Science and Technology, 1(1), 16-22.

Fallucchi, F., Coladangelo, M., Giuliano, R., & William De Luca, E. (2020). Predicting employee attrition using machine learning techniques. Computers, 9(4), 86.

Fan, G., Zhang, C., Chen, J., Li, B., Xu, Z., Li, Y., ... & Gong, Z. (2022, May). Field-aware variational autoencoders for billion-scale user representation learning. In 2022 IEEE 38th International Conference on Data Engineering (ICDE) (pp. 3413-3425). IEEE.

Fan, T. K., & Chang, C. H. (2010). Sentiment-oriented contextual advertising. Knowledge and Information Systems, 23(3), 321–344.

Fan, T. K., & Chang, C. H. (2011). Blogger-centric contextual advertising. Expert Systems with Applications, 38(3), 1777–1788.

Fang, X., & Hu, P. J. H. (2018). Top persuader prediction for social networks. MIS Quarterly, 42(1), 63–82.

Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. Journal of Big Data, 2(1), 1-14.

Farzadnia, S., & Vanani, I. R. (2022). Identification of opinion trends using sentiment analysis of airlines passengers' reviews. Journal of Air Transport Management, 103, 102232.

Fedus, W., Ramachandran, P., Agarwal, R., Bengio, Y., Larochelle, H., Rowland, M., & Dabney, W. (2020, November). Revisiting fundamentals of experience replay. In International Conference on Machine Learning (pp. 3061-3071). PMLR.

Feng, Y., Chen, H., & Kong, Q. (2020). An expert with whom I can identify: The role of narratives in influencer marketing. International Journal of Advertising, 1–22.

Ferreira, L. B., da Cunha, F. F., & Fernandes Filho, E. I. (2022). Exploring machine learning and multi-task learning to estimate meteorological data and reference evapotranspiration across Brazil. Agricultural Water Management, 259, 107281.

Fiore, M., Gallo, C., Tsoukatos, E., & La Sala, P. (2017). Predicting consumer healthy choices regarding type 1 wheat flour. British Food Journal, 119(11), 2388–2405.

Florez-Lopez, R., & Ramon-Jeronimo, J. M. (2009). Marketing segmentation through machine learning models an approach based on customer relationship management and customer profitability accounting. Social Science Computer Review, 27(1), 96–117.

Fuertes, G., Alfaro, M., Vargas, M., Gutierrez, S., Ternero, R., & Sabattin, J. (2020). Conceptual framework for the strategic management: a literature review—descriptive. Journal of Engineering, 2020, 1-21.

G

Gan, Y., Ji, Y., Jiang, S., Liu, X., Feng, Z., Li, Y., & Liu, Y. (2021). Integrating aesthetic and emotional preferences in social robot design: An affective design approach with Kansei Engineering and Deep Convolutional Generative Adversarial Network. International Journal of Industrial Ergonomics, 83, 103128.

Gao, S., Greene, A. S., Constable, R. T., & Scheinost, D. (2019). Combining multiple connectomes improves predictive modeling of phenotypic measures. Neuroimage, 201, 116038.

Gautier, P. (2021). Contribution des méthodes d'apprentissage à la distribution de tâches dans un cluster robotique (Doctoral dissertation, Université de Bretagne Sud).

George, P. M., Ganguly, S., & Yasa, V. R. (2023). Machine Learning and Artificial Intelligence-Based Tools in Digital Marketing: An Integrated Approach. In Machine Learning for Business Analytics (pp. 97-106). Productivity Press.

Gokkoeva, M. (2020). Trends in Search Engine Optimisation: The Role of Voice Search.

Gold, M. S., & Bentler, P. M. (2000). Treatments of missing data: A Monte Carlo comparison of RBHDI, iterative stochastic regression imputation, and expectation-maximization. Structural equation modeling, 7(3), 319-355.

Gong, C., Tao, D., Liu, W., Liu, L., & Yang, J. (2016). Label propagation via teaching-to-learn and learning-to-teach. IEEE transactions on neural networks and learning systems, 28(6), 1452-1465.

Gong, S., Xie, Y., Xu, J., Niyato, D., & Liang, Y. C. (2020). Deep reinforcement learning for backscatter-aided data offloading in mobile edge computing. IEEE Network, 34(5), 106-113.

Gong, T., Lee, T., Stephenson, C., Renduchintala, V., Padhy, S., Ndirango, A., & Elibol, O. H. (2019). A comparison of loss weighting strategies for multi task learning in deep neural networks. IEEE Access, 7, 141627-141632.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2020). Generative adversarial networks. Commun. Acm, 63(11), 139-144.

Goodrich, K., Schiller, S.u. Z., & Galletta, D. (2015). Consumer reactions to intrusiveness of online-video advertisements: Do length, informativeness, and humor help (or hinder) marketing outcomes? Journal of Advertising Research, 55(1), 37–50.

Grandini, M., Bagli, E., & Visani, G. (2020). Metrics for multi-class classification: an overview. arXiv preprint arXiv:2008.05756.

Gu, S., Lillicrap, T., Sutskever, I., & Levine, S. (2016, June). Continuous deep q-learning with model-based acceleration. In International conference on machine learning (pp. 2829-2838). PMLR.

Gul, R. F., Liu, D., Jamil, K., Baig, S. A., Awan, F. H., & Liu, M. (2021). Linkages between market orientation and brand performance with positioning strategies of significant fashion apparels in Pakistan. Fashion and Textiles, 8(1), 1-19.

Gunawan, G., & Sulaeman, M. (2020). Determining Factors in the Use of Digital Marketing and Its Effect on Marketing Performance in the Creative Industries in Tasikmalaya. Budapest International Research and Critics Institute-Journal (BIRCI-Journal), 3(3), 2543-2550.

Gupta, V., Sadana, R., & Moudgil, S. (2019). Image style transfer using convolutional neural networks based on transfer learning. International journal of computational systems engineering, 5(1), 53-60.

Gurumurthy, K., & Mukherjee, A. (2020). The Bass Model: a parsimonious and accurate approach to forecasting mortality caused by COVID-19. International Journal of Pharmaceutical and Healthcare Marketing, 14(3), 349-360.

Gönen, M., & Alpaydın, E. (2013). Localized algorithms for multiple kernel learning. Pattern Recognition, 46(3), 795-807.

H

Hai Ming, L., Gang, L., Hua, H., & Waqas, M. (2022). Modeling the influencing factors of electronic word-of-mouth about CSR on social networking sites. Environmental Science and Pollution Research, 29(44), 66204-66221.

Haider, C. M. R., Iqbal, A., Rahman, A. H., & Rahman, M. S. (2018). An ensemble learning based approach for impression fraud detection in mobile advertising. Journal of Network and Computer Applications, 112, 126–141.

Hajarian, M., Camilleri, M. A., Díaz, P., & Aedo, I. (2021). A taxonomy of online marketing methods. In Strategic corporate communication in the digital age (pp. 235-250). Emerald Publishing Limited.

Halder, R., Fidkowski, K. J., & Maki, K. J. (2022). Non-intrusive reduced-order modeling using convolutional autoencoders. International Journal for Numerical Methods in Engineering, 123(21), 5369-5390.

Hanna, R., Rohm, A., & Crittenden, V. L. (2011). We're all connected: The power of the social media ecosystem. Business horizons, 54(3), 265-273.

Haque, T. U., Saber, N. N., & Shah, F. M. (2018, May). Sentiment analysis on large scale Amazon product reviews. In 2018 IEEE international conference on innovative research and development (ICIRD) (pp. 1-6). IEEE.

Harsoor, A. S., & Patil, A. (2015). Forecast of sales of Walmart store using big data applications. International Journal of Research in Engineering and Technology, 4(6), 51-59.

Hassenzahl, M. (2018). The thing and I: understanding the relationship between user and product. Funology 2: from usability to enjoyment, 301-313.

Hazim, M., Anuar, N. B., Ab Razak, M. F., & Abdullah, N. A. (2018). Detecting opinion spams through supervised boosting approach. PloS One, 13(6), Article e0198884.

He, W., He, Y., Li, B., & Zhang, C. (2019). A naive-Bayes-based fault diagnosis approach for analog circuit by using image-oriented feature extraction and selection technique. IEEE Access, 8, 5065-5079.

Herpich, F., Guarese, R. L. M., & Tarouco, L. M. R. (2017). A comparative analysis of augmented reality frameworks aimed at the development of educational applications. Creative Education, 8(9), 1433-1451.

Hill, T., & Marquez, L. (1994). M. O, Connor, and W. Remus,"Artificial Neural Networks for Forecasting and Decision Marking,". International Journal of Forecasting, 10, 5.doi: 10.1016/0169-2070(94)90045-0

Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. Science, 349(6245), 261-266.

Ho, C. I., Liu, Y., & Chen, M. C. (2022). Factors influencing watching and purchase intentions on live streaming platforms: From a 7Ps marketing mix perspective. Information, 13(5), 239.

Ho, H. N., & Lee, E. (2015, January). Model-based reinforcement learning approach for planning in self-adaptive software system. In Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication (pp. 1-8).

Holcomb, S. D., Porter, W. K., Ault, S. V., Mao, G., & Wang, J. (2018, March). Overview on deepmind and its alphago zero ai. In Proceedings of the 2018 international conference on big data and education (pp. 67-71).

Hosseinpour, F., Amoli, P. V., Farahnakian, F., Plosila, J., & Hämäläinen, T. (2014). Artificial immune system-based intrusion detection: innate immunity using an unsupervised learning approach. International Journal of Digital Content Technology and its Applications, 8(5), 1.

Hou, Z. H., Ma, K., Wang, Y. F., Yu, J., Ji, K., Chen, Z. X., & Abraham, A. (2021). Attention-based learning of self-media data for marketing intention detection. Engineering Applications of Artificial Intelligence, 98, 1–9.

Houman Andersen, P. (2001). Relationship development and marketing communication: an integrative model. Journal of business & industrial marketing, 16(3), 167-183.

Hsieh, C. C., Karkoub, M., Lai, W. R., & Lin, P. H. (2015). Visual people counting using gender features and LRU updating scheme. Multimedia Tools and Applications, 74(6), 1741–1759.

Hu, J. L., Li, T. H., Zhuang, Y., Huang, S., & Dong, S. B. (2020). GFD: A weighted heterogeneous graph embedding based approach for fraud detection in mobile advertising. Security and Communication Networks, 2020, 1–12.

Hu, J., Li, X., Hu, G., Lyu, Y., Susztak, K., & Li, M. (2020). Iterative transfer learning with neural network for clustering and cell type classification in single-cell RNA-seq analysis. Nature machine intelligence, 2(10), 607-618.

Huang, F., Zhu, M., Yuan, K., & Deng, E.O. (2015). Telco churn prediction with big data. ACM SIGMOD international conference on management of data, 607-618.

Huang, G. (2022, April). E-Commerce Intelligent Recommendation System Based on Deep Learning. In 2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC) (pp. 1154-1157). IEEE.

I

Ibarz, J., Tan, J., Finn, C., Kalakrishnan, M., Pastor, P., & Levine, S. (2021). How to train your robot with deep reinforcement learning: lessons we have learned. The International Journal of Robotics Research, 40(4-5), 698-721.

Infante, A., & Mardikaningsih, R. (2022). The Potential of social media as a Means of Online Business Promotion. Journal of Social Science Studies (JOS3), 2(2), 45-49.

Ing, P., Osman, Z., & Tze-Yin, L. (2020). Marketing Mix and STP Strategies: An Exploratory Study into Grocery Retailers in Malaysia. Asian Journal of Entrepreneurship, 1(4), 129-143.

Islam, M. R., & Nahiduzzaman, M. (2022). Complex features extraction with deep learning model for the detection of COVID19 from CT scan images using ensemble-based machine learning approach. Expert Systems with Applications, 195, 116554.

Ištvanić, M., Crnjac Milić, D., & Krpić, Z. (2017). Digital marketing in the business environment. International journal of electrical and computer engineering systems, 8(2.), 67-75.

J

Jaas, A. (2022). E-Marketing and Its Strategies: Digital Opportunities and Challenges. Open Journal of Business and Management, 10(2), 822-845.

Jagabathula, S., Subramanian, L., & Venkataraman, A. (2018). A model-based embedding technique for segmenting customers. Operations Research, 66(5), 1247–1267.

Jagdale, R. S., Shirsat, V. S., & Deshmukh, S. N. (2019). Sentiment analysis on product reviews using machine learning techniques. In Cognitive informatics and soft computing (pp. 639-647). Springer, Singapore.

Jai, T.-M., Fang, D., Bao, F. S., James, R. N., Chen, T. W., & Cai, W. D. (2021). Seeing it is like touching it: Unraveling the effective product presentations on online apparel purchase decisions and brain activity (An fMRI Study). Journal of Interactive Marketing, 53, 66–79.

Jain, S., Seth, G., Paruthi, A., Soni, U., & Kumar, G. (2022). Synthetic data augmentation for surface defect detection and classification using deep learning. Journal of Intelligent Manufacturing, 1-14.

Jamshido'g'li, R. D., RakhmonjonZokirjono'g'li, O., & Kholdorovna, R. F. (2020). Changes in market relations in the period of market economy and their classification. South Asian Journal of Marketing & Management Research, 10(4), 120-129.

Jang, Y., & Park, E. (2022). Satisfied or not: user experience of mobile augmented reality in using natural language processing techniques on review comments. Virtual Reality, 1-10.

Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. Electronic Markets, 31(3), 685-695.

Jaradat, Y., Masoud, M., Jannoud, I., Manasrah, A., & Alia, M. (2021, July). A Tutorial on Singular Value Decomposition with Applications on Image Compression and Dimensionality Reduction. In 2021 International Conference on Information Technology (ICIT) (pp. 769-772). IEEE.

Jeon, Y., Jeon, S. G., & Han, K. S. (2020). Better targeting of consumers: Modeling multifactorial gender and biological sex from Instagram posts. User Modeling and User-Adapted Interaction, 30(5), 833–866.

Jeshurun, S. B. (2018). A study on the effectiveness of email marketing. Management, 6(S1), 84-86.

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., & Fung, P. (2023). Survey of hallucination in natural language generation. ACM Computing Surveys, 55(12), 1-38.

Jia, J., & Wang, W. (2020, October). Review of reinforcement learning research. In 2020 35th Youth Academic Annual Conference of Chinese Association of Automation (YAC) (pp. 186-191). IEEE.

Jiang, J., Chen, J., Gu, T., Choo, K. K. R., Liu, C., Yu, M., ... & Mohapatra, P. (2019, November). Anomaly detection with graph convolutional networks for insider threat and fraud detection. In MILCOM 2019-2019 IEEE Military Communications Conference (MILCOM) (pp. 109-114). IEEE.

Jiang, W. (2022, September). Applications of singular value decomposition in data reduction. In International Conference on Applied Statistics, Computational Mathematics, and Software Engineering (ASCMSE 2022) (Vol. 12345, pp. 138-148). SPIE.

Jiao, Y., Wang, P., Niyato, D., Lin, B., & Kim, D. I. (2020). Toward an automated auction framework for wireless federated learning services market. IEEE Transactions on Mobile Computing, 20(10), 3034-3048.

Joachims, T. (1998). Text categorization with support vector machines: Learning with many relevant features. Machine learning: ECML-98, 137-142.

K

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. Journal of artificial intelligence research, 4, 237-285.

Kaiser, C., Ahuvia, A., Rauschnabel, P. A., & Wimble, M. (2020). Social media monitoring: What can marketers learn from Facebook brand photos? Journal of Business Research, 117, 707–717.

Kanei, F., Chiba, D., Hato, K., Yoshioka, K., Matsumoto, T., & Akiyama, M. (2020). Detecting and understanding online advertising fraud in the wild. IEICE Transactions on Information and Systems, E103d(7), 1512–1523.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. Business horizons, 53(1), 59-68.

Kapoor, R., & Kapoor, K. (2021). The transition from traditional to digital marketing: a study of the evolution of e-marketing in the Indian hotel industry. Worldwide Hospitality and Tourism Themes, 13(2), 199-213.

Kar, M. K., Neog, D. R., & Nath, M. K. (2023). Retinal vessel segmentation using multi-scale residual convolutional neural network (MSR-Net) combined with generative adversarial networks. Circuits, Systems, and Signal Processing, 42(2), 1206-1235.

Karjaluoto, H., & Leinonen, H. (2009). Advertisers' perceptions of search engine marketing. International Journal of Internet Marketing and Advertising, 5(1-2), 95-112.

Kasabov, N. K. (2019). Time-space, spiking neural networks and brain-inspired artificial intelligence. Berlin, Heidelberg: Springer Berlin Heidelberg.

Kassabgi, G. (2017, January 11). Text Classification using Algorithms. Chatbotslife. Retrieved from https://chatbotslife.com/textclassification-using-algorithms-e4d50dcba45/

Kassabgi, G. (2017, January 24). Text Classification using Neural Networks. Machinelearnings. Retrieved from https://machinelearnings.co/text-classification-using-neural-networksf5cd7b8765c6/

Kaur, R., Singh, R., Gehlot, A., Priyadarshi, N., & Twala, B. (2022). Marketing Strategies 4.0: Recent Trends and Technologies in Marketing. Sustainability, 14(24), 16356.

Kazantseva, A., & Szpakowicz, S. (2011, July). Linear text segmentation using affinity propagation. In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (pp. 284-293).

Keikhosrokiani, P., & Pourya Asl, M. (Eds.). (2022). Handbook of Research on Opinion Mining and Text Analytics on Literary Works and Social Media. IGI Global.

Kendall, A., & Cipolla, R. (2017). Geometric loss functions for camera pose regression with deep learning. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 5974-5983).

Kernbach, J. M., & Staartjes, V. E. (2022). Foundations of machine learning-based clinical prediction modeling: Part II—Generalization and overfitting. Machine Learning in Clinical Neuroscience: Foundations and Applications, 15-21.

Key, T. M. (2017). Domains of digital marketing channels in the sharing economy. Journal of Marketing Channels, 24(1-2), 27-38.

Khan, L. U., Han, Z., Saad, W., Hossain, E., Guizani, M., & Hong, C. S. (2022). Digital twin of wireless systems: Overview, taxonomy, challenges, and opportunities. IEEE Communications Surveys & Tutorials.

Khurshid, F., Zhu, Y., Xu, Z., Ahmad, M., & Ahmad, M. (2019). Enactment of ensemble learning for review spam detection on selected features. International Journal of Computational Intelligence Systems, 12(1), 387–394.

Kim, J., Tae, D., & Seok, J. (2020, February). A survey of missing data imputation using generative adversarial networks. In 2020 International conference on artificial intelligence in information and communication (ICAIIC) (pp. 454-456). IEEE.

King, M. A., Abrahams, A. S., & Ragsdale, C. T. (2015). Ensemble learning methods for pay-per-click campaign management. Expert Systems with Applications, 42(10), 4818–4829.

Kiran, K. U., & Arumugam, T. (2020, December). Role of programmatic advertising on effective digital promotion strategy: A conceptual framework. In Journal of Physics: Conference Series (Vol. 1716, No. 1, p. 012032). IOP Publishing.

Klinger, R., & Tomanek, K. (2007). Classical probabilistic models and conditional random fields. TU, Algorithm Engineering.

Kocev, D., Vens, C., Struyf, J., & Džeroski, S. (2007). Ensembles of multi-objective decision trees. In Machine Learning: ECML 2007: 18th European Conference on Machine Learning, Warsaw, Poland, September 17-21, 2007. Proceedings 18 (pp. 624-631). Springer Berlin Heidelberg.

Koehn, D., Lessmann, S., & Schaal, M. (2020). Predicting online shopping behaviour from clickstream data using deep learning. Expert Systems with Applications, 150, 1–16.

Kong, X., Wu, Y., Wang, H., & Xia, F. (2022). Edge Computing for Internet of Everything: A Survey. IEEE Internet of Things Journal, 9(23), 23472-23485.

Kontschieder, P., Fiterau, M., Criminisi, A., & Bulo, S. R. (2015). Deep neural decision forests. In Proceedings of the IEEE international conference on computer vision (pp. 1467-1475).

Koul, S., & Philip, T. M. (2021, March). Customer Segmentation Techniques on E-Commerce. In 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) (pp. 135-138). IEEE.

Kovačević, M., Pasquato, M., Marelli, M., De Luca, A., Salvaterra, R., & Belfiore, A. (2022). Exploring X-ray variability with unsupervised machine learning-I. Self-organizing maps applied to XMM-Newton data. Astronomy & Astrophysics, 659, A66.

Krishna, A., Akhilesh, V., Aich, A., & Hegde, C. (2018, December). Sales-forecasting of retail stores using machine learning techniques. In 2018 3rd international conference on computational systems and information technology for sustainable solutions (CSITSS) (pp. 160-166). IEEE.

Krishna, S. H., Sargunam, S. S., Kulkarni, N., Nandal, N., Chellam, V. V., & Praveenkumar, S. (2023, January). Application of Artificial Intelligence in E-Marketing. In 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF) (pp. 1-7). IEEE.

Kritzinger, W. T., & Weideman, M. (2013). Search engine optimization and pay-per-click marketing strategies. Journal of Organizational Computing and Electronic Commerce, 23(3), 273-286.

Kumanova-Larde, D., & Hristov, P. (2019). Some modern aspects of e-marketing.

Kumar, H., Agarwal, A., Dasgupta, R., & Joshi, S. (2018, April). Dialogue act sequence labeling using hierarchical encoder with crf. In Proceedings of the AAAI conference on artificial intelligence, 32(1).

Kumar, N., & Naik, C. (2017). Comparative Analysis of Machine Learning Algorithms for Their Effectiveness in Churn Prediction in the Telecom Industry. International Research Journal of Engineering and Technology, 4, 485-489.

Kuo, R. J., & Cheng, H. R. (2022). A content-based recommender system with consideration of repeat purchase behavior. Applied Soft Computing, 127, 109361.

Kwok, L., Tang, Y., & Yu, B. (2020). The 7 Ps marketing mix of home-sharing services: Mining travelers' online reviews on Airbnb. International Journal of Hospitality Management, 90, 1–11.

Kübler, R. V., Colicev, A., & Pauwels, K. H. (2020). Social media's impact on the consumer mindset: When to use which sentiment extraction tool? Journal of Interactive Marketing, 50, 136–155.

Kühl, N., Mühlthaler, M., & Goutier, M. (2019). Supporting customer-oriented marketing with artificial intelligence: Automatically quantifying customer needs from social media. Electronic Markets, 30(2), 351–367.

L

Lai, G., Chang, W. C., Yang, Y., & Liu, H. (2018, June). Modeling long-and short-term temporal patterns with deep neural networks. In The 41st international ACM SIGIR conference on research & development in information retrieval (pp. 95-104).

Lambrecht, P., & Peter, M. K. (2022). The influence of digital assistants on search engine strategies: recommendations for voice search optimization. In Marketing and Smart Technologies: Proceedings of ICMarkTech 2021, Volume 2 (pp. 665-672). Singapore: Springer Singapore.

Lasi, M. B. A. (2021). The Relationship between E-Marketing Mix Strategy and Integrated Marketing Communication: A Conceptual Framework. International Journal of Economics and Management Systems, 6.

Lauriola, I., Lavelli, A., & Aiolli, F. (2022). An introduction to deep learning in natural language processing: Models, techniques, and tools. Neurocomputing, 470, 443-456.

Le Bihan, N., & Mars, J. (2004). Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing. Signal processing, 84(7), 1177-1199.

Le Meunier-FitzHugh, K. (2021). Marketing: A Very Short Introduction. Oxford University Press, USA.

Lee, D., Hosanagar, K., & Nair, H. S. (2018). Advertising content and consumer engagement on social media: Evidence from Facebook. Management Science, 64(11), 5105–5131.

Lee, I., & Shin, Y. J. (2020). Machine learning for enterprises: Applications, algorithm selection, and challenges. Business Horizons, 63(2), 157-170.

Lee, J., Xiao, L., Schoenholz, S., Bahri, Y., Novak, R., Sohl-Dickstein, J., & Pennington, J. (2019). Wide neural networks of any depth evolve as linear models under gradient descent. Advances in neural information processing systems, 32.

Lee, M. (2011). Google ads and the blindspot debate. Media, Culture & Society, 33(3), 433-447.

Lee, S. (2018). Deep structured learning: architectures and applications. International Journal of Advanced Culture Technology, 6(4), 262-265.

Lee, S., & Park, H. (2023, February). Effect of Optimization Techniques on Feedback Alignment Learning of Neural Networks. In 2023 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) (pp. 227-231). IEEE.

Leguina, J. R., Rumin, A. C., & Rumin, R. C. (2020). Digital marketing attribution: Understanding the user path. Electronics, 9(11), 1–25.

Lessmann, S., Haupt, J., Coussement, K., & De Bock, K. W. (2021). Targeting customers for profit: An ensemble learning framework to support marketing decision-making. Information Sciences, 557, 286–301.

Letheren, Kate, Glavas, Charmaine. (2017). Embracing the bots: How direct to consumer advertising is about to change forever. The Conversation, March 17.

Li, D., Chen, D., Jin, B., Shi, L., Goh, J., & Ng, S. K. (2019, September). MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks. In Artificial Neural Networks and Machine Learning–ICANN 2019: Text and Time Series: 28th International Conference on Artificial Neural

Networks, Munich, Germany, September 17–19, 2019, Proceedings, Part IV (pp. 703-716). Cham: Springer International Publishing.

Li, H., Wu, Y., Cao, D., & Wang, Y. (2021). Organizational mindfulness towards digital transformation as a prerequisite of information processing capability to achieve market agility. Journal of Business research, 122, 700-712.

Li, W., Prasad, S., Fowler, J. E., & Bruce, L. M. (2011). Locality-preserving discriminant analysis in kernel-induced feature spaces for hyperspectral image classification. IEEE Geoscience and Remote Sensing Letters, 8(5), 894-898.

Li, Y. R., & Zhang, L. Y. (2020). Do online reviews truly matter? A study of the characteristics of consumers involved in different online review scenarios. Behaviour & Information Technology, 1–19.

Liang, X., Wang, C., & Zhao, G. (2019). Enhancing content marketing article detection with graph analysis. IEEE Access, 7, 94869–94881.

Lies, J. (2019). Marketing intelligence and big data: Digital marketing techniques on their way to becoming social engineering techniques in marketing.

Ligthart, A., Catal, C., & Tekinerdogan, B. (2021). Analyzing the effectiveness of semi-supervised learning approaches for opinion spam classification. Applied Soft Computing, 101, 107023.

Lin, S. L. (2021). Application of machine learning to a medium Gaussian support vector machine in the diagnosis of motor bearing faults. Electronics, 10(18), 2266.

Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008, December). Isolation forest. In 2008 eighth ieee international conference on data mining (pp. 413-422). IEEE.

Liu, L., Dzyabura, D., Mizik, N. (2020). Visual Listening In: Extracting Brand Image Portrayed on Social Media. Marketing Science, 39(4), 669-686.

Liu, N., Ren, S., Choi, T. M., Hui, C. L., & Ng, S. F. (2013). Sales forecasting for fashion retailing service industry: a review. Mathematical Problems in Engineering, 2013. doi: 10.1155/2013/738675

Liu, P., An, K., Lei, J., Liu, W., Sun, Y., Zheng, G., & Chatzinotas, S. (2023). SCMA-Enabled Multi-Cell Edge Computing Networks: Design and Optimization. IEEE Transactions on Vehicular Technology.

Liu, P., Qiu, X., & Huang, X. (2017). Adversarial multi-task learning for text classification. arXiv preprint arXiv:1704.05742.

Liu, R., Wu, T., & Mozafari, B. (2020). Adam with bandit sampling for deep learning. Advances in Neural Information Processing Systems, 33, 5393-5404.

Lopes, A. R., & Casais, B. (2022). Digital Content marketing: conceptual review and recommendations for practitioners. Academy of Strategic Management Journal, 21(2), 1-17.

Loshchilov, I. (2017). LM-CMA: An alternative to L-BFGS for large-scale black box optimization. Evolutionary computation, 25(1), 143-171.

Loukili, M., Messaoudi, F. (2023). Machine Learning, Deep Neural Network and Natural Language Processing Based Recommendation System. In: Kacprzyk, J., Ezziyyani, M., Balas, V.E. (eds) International Conference on Advanced Intelligent Systems for Sustainable Development. AI2SD 2022. Lecture Notes in Networks and Systems, vol 637. Springer, Cham. https://doi.org/10.1007/978-3-031-26384-2_7

Loukili, M., Messaoudi, F., & El Ghazi, M. (2022). Supervised Learning Algorithms for Predicting Customer Churn with Hyperparameter Optimization, International Journal of Advances in Soft Computing and its Application, 14(3), 49-63.

Loukili, M., Messaoudi, F., & El Ghazi, M. (2023). Defending Against Digital Thievery: A Machine Learning Approach to Predict E-Payment Fraud. International Journal of Management Practice.

Loukili, M., Messaoudi, F., & El Ghazi, M. (2023). Machine Learning Based Recommender System for E-Commerce. IAES International Journal of Artificial Intelligence (IJ-AI), 12(4), 1803-1811. doi: 10.11591/ijai.v12.i4

Loukili, M., Messaoudi, F., & El Ghazi, M. (2023). Sentiment Analysis of Product Reviews for E-Commerce Recommendation based on Machine Learning, International Journal of Advances in Soft Computing and its Application, 15(1), 1-13.

Luaces, O., Diez, J., Joachims, T., & Bahamonde, A. (2015). Mapping preferences into Euclidean space. Expert Systems with Applications, 42(22), 8588–8596.

Luo, X. M., Tong, S. L., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. Marketing Science, 38(6), 937–947.

Luo, Y., & Xu, X. W. (2019). Predicting the helpfulness of online restaurant reviews using different machine learning algorithms: A case study of yelp. Sustainability, 11(19), 1–17

Luo, Y., Tao, D., Ramamohanarao, K., Xu, C., & Wen, Y. (2015). Tensor canonical correlation analysis for multi-view dimension reduction. IEEE transactions on Knowledge and Data Engineering, 27(11), 3111-3124.

López García, J. J., Lizcano, D., Ramos, C. M., & Matos, N. (2019). Digital marketing actions that achieve a better attraction and loyalty of users: An analytical study. Future Internet, 11(6), 130.

M

Madhu, G., Kautish, S., Alnowibet, K. A., Zawbaa, H. M., & Mohamed, A. W. (2023). NIPUNA: A Novel Optimizer Activation Function for Deep Neural Networks. Axioms, 12(3), 246.

Mahdavi, K., Labarta, J., Gimenez, J., Mousavinia, A., & Mousavinia, A. (2022, July). Feature Space Curvature Map: A Method to Homogenize Cluster Densities. In 2022 International Joint Conference on Neural Networks (IJCNN) (pp. 1-10). IEEE.

Mahesh, B. (2020). Machine learning algorithms-a review. International Journal of Science and Research (IJSR). [Internet], 9, 381-386.

Maind, S. B., & Wankar, P. (2014). Research paper on basic of artificial neural network. International Journal on Recent and Innovation Trends in Computing and Communication, 2(1), 96-100.

Mak, M. W., Ku, K. W., & Lu, Y. L. (1999). On the improvement of the real time recurrent learning algorithm for recurrent neural networks. Neurocomputing, 24(1-3), 13-36.

Mamaghani, M. T., & Hong, Y. (2020). Intelligent trajectory design for secure full-duplex MIMO-UAV relaying against active eavesdroppers: A model-free reinforcement learning approach. IEEE Access, 9, 4447-4465.

Maryani, I., Riana, D. Astuti, R. D., Ishaq, A., Sutrisno, & Pratama, E. A. “Customer segmentation based on RFM model and clustering techniques with k-means algorithm,” in 2018 Third International Conference on Informatics and Computing (ICIC). IEEE, Oct. 2018. [Online]. Available: https://doi.org/10.1109/iac.2018.8780570

Masui, K., Okada, G., & Tsumura, N. (2020). Measurement of advertisement effect based on multimodal emotional responses considering personality. ITE Transactions on Media Technology and Applications, 8(1), 49–59.

Matz, S. C., Segalin, C., Stillwell, D., Muller, S. R., & Bos, M. W. (2019). Predicting the personal appeal of marketing images using computational methods. Journal of Consumer Psychology, 29(3), 370–390.

McNeish, D. M. (2015). Using lasso for predictor selection and to assuage overfitting: A method long overlooked in behavioral sciences. Multivariate behavioral research, 50(5), 471-484.

Mendonça, M. F. T. F. D. (2022). Content marketing: how product presentation and search engine optimization techniques can improve the online customer experience: case study of Mercadão (Doctoral dissertation).

Merisavo, M., & Raulas, M. (2004). The impact of e-mail marketing on brand loyalty. Journal of Product & Brand Management, 13(7), 498-505.

Millstein, F. (2020). Natural language processing with python: natural language processing using NLTK. Frank Millstein.

Miralles-Pechu´ an, L., Ponce, H., & Martínez-Villasenor, ˜ L. (2018). A novel methodology for optimizing display advertising campaigns using genetic algorithms. Electronic Commerce Research and Applications, 27, 39–51.

Mishu, S. Z., & Rafiuddin, S. M. (2016, December). Performance analysis of supervised machine learning algorithms for text classification. In Computer and Information Technology (ICCIT), 2016 19th International Conference on (pp. 409-413). IEEE.

Miyato, T., Maeda, S. I., Koyama, M., & Ishii, S. (2018). Virtual adversarial training: a regularization method for supervised and semi-supervised learning. IEEE transactions on pattern analysis and machine intelligence, 41(8), 1979-1993.

Molenaar, C. (2013). E-Marketing: Applications of information technology and the internet within marketing. Routledge.

Montanés, E., Suarez-V azquez, A., & Quevedo, J. R. (2014). Ordinal classification/ regression for analyzing the influence of superstars on spectators in cinema marketing. Expert Systems with Applications, 41(18), 8101–8111.

Moon, J., Noh, Y., Jung, S., Lee, J., & Hwang, E. (2023). Anomaly detection using a model-agnostic meta-learning-based variational auto-encoder for facility management. Journal of Building Engineering, 68, 106099.

Moon, Y., & Armstrong, D. J. (2020). Service quality factors affecting customer attitudes in online-to-offline commerce. Information Systems and e-Business Management, 18(1), 1-34.

Morette, N., Heredia, L. C., Ditchi, T., Mor, A. R., & Oussar, Y. (2020). Partial discharges and noise classification under HVDC using unsupervised and semi-supervised learning. International Journal of Electrical Power & Energy Systems, 121, 106129.

Mostafa, M. M. (2009). Shades of green: A psychographic segmentation of the green consumer in Kuwait using self-organizing maps. Expert Systems with Applications, 36 (8), 11030–11038.

Mousavi, S. S., Schukat, M., & Howley, E. (2017). Traffic light control using deep policy-gradient and value-function-based reinforcement learning. IET Intelligent Transport Systems, 11(7), 417-423.

Mousavi, S. S., Schukat, M., & Howley, E. (2018). Deep reinforcement learning: an overview. In Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016(2), 426-440. Springer International Publishing.

Mowle, N., Georgia, E., Doss, B., & A. Updegraff, J. (2014). Application of regulatory focus theory to search advertising. Journal of Consumer Marketing, 31(6/7), 494-502.

Mujika, A., Meier, F., & Steger, A. (2018). Approximating real-time recurrent learning with random kronecker factors. Advances in Neural Information Processing Systems, 31.

Mukherjee, S., & Bhattacharyya, P. (2012, March). Feature specific sentiment analysis for product reviews. In International conference on intelligent text processing and computational linguistics (pp. 475-487). Springer, Berlin, Heidelberg.

Mulyana, M., & Azka, M. (2022). Market Sensing Capability and Customer Relationship Management on Marketing Performance: Role E-Marketing Strategy. Jurnal Manajemen Indonesia, 22(3), 279-288.

N

Nadube, P. M., & Ordah, J. M. (2023). E-CUSTOMER RELATIONSHIP MANAGEMENT AND E-MARKETING PERFORMANCE OF DEPOSIT MONEY BANKS IN PORT HARCOURT. BW Academic Journal, 19-19.

Naeem, A., Rehman, M., Anjum, M., & Asif, M. (2019). Development of an efficient hierarchical clustering analysis using an agglomerative clustering algorithm. Current Science, 117(6), 1045-1053.

Nagabandi, A., Kahn, G., Fearing, R. S., & Levine, S. (2018, May). Neural network dynamics for model-based deep reinforcement learning with model-free fine-tuning. In 2018 IEEE international conference on robotics and automation (ICRA) (pp. 7559-7566). IEEE.

Nagarajan, S. M., & Gandhi, U. D. (2019). Classifying streaming of Twitter data based on sentiment analysis using hybridization. Neural Computing and Applications, 31(5), 1425-1433.

Najm, H., Al-Mafrachi, M., & Pascal, L. (2021). The Impact of Electronic Marketing on Economic Development-A Theoretical Approach. Scientific Papers: Management, Economic Engineering in Agriculture & Rural Development, 21(2), 407-415.

Nam, N., & Duong, T. (2021). Comparison of Two Main Approaches for Handling Imbalanced Data in Churn Prediction Problem. Journal of Advances in Information Technology, 12(1), 29-35. doi: 10.12720/jait.12.1.29-35

Namatēvs, I. (2017). Deep convolutional neural networks: Structure, feature extraction and training. Information Technology and Management Science, 20(1), 40-47.

Namugenyi, C., Nimmagadda, S. L., & Reiners, T. (2019). Design of a SWOT analysis model and its evaluation in diverse digital business ecosystem contexts. Procedia Computer Science, 159, 1145-1154.

Nanne, A. J., Antheunis, M. L., Van Der Lee, C. G., Postma, E. O., Wubben, S., & Van Noort, G. (2020). The use of computer vision to analyze brand-related user generated image content. Journal of Interactive Marketing, 50(1), 156-167.

Ng, Y. C., Colombo, N., & Silva, R. (2018). Bayesian semi-supervised learning with graph gaussian processes. Advances in Neural Information Processing Systems, 31.

Ngai, E. W. (2003). Selection of web sites for online advertising using the AHP. Information & management, 40(4), 233-242.

Nguyen, T. K. D. (2015). A study of real-time content marketing: formulating real-time content marketing based on content, search and social media.

Ni, P., Li, Y., Li, G., & Chang, V. (2020). Natural language understanding approaches based on joint task of intent detection and slot filling for IoT voice interaction. Neural Computing and Applications, 32, 16149-16166.

Nie, F., Zhu, W., & Li, X. (2020). Decision Tree SVM: An extension of linear SVM for non-linear classification. Neurocomputing, 401, 153-159.

Nikam, S. S. (2015). A comparative study of classification techniques in data mining algorithms. Oriental Journal of Computer Science and Technology, 8(1), 13-19.

Nikolaev, A. S., Maximova, T. G., Sakhno, I. E., Antipov, A. A., & Murashova, S. V. (2023). Facial Recognition Technologies Patent Landscape. In Software Engineering Application in Systems Design: Proceedings of 6th Computational Methods in Systems and Software 2022, Volume 1 (pp. 568-583). Cham: Springer International Publishing.

Nilashi, M., Mardani, A., Liao, H. C., Ahmadi, H., Manaf, A. A., & Almukadi, W. (2019). A hybrid method with TOPSIS and machine learning techniques for sustainable development of green hotels considering online reviews. Sustainability, 11(21), 1–21.

Nosrati, M., Karimi, R., Mohammadi, M., & Malekian, K. (2013). Internet marketing or modern advertising! How? Why. International Journal of economy, management and social sciences, 2(3), 56-63.

O

O'Halloran, S., & Nowaczyk, N. (2019). An artificial intelligence approach to regulating systemic risk. Frontiers in Artificial Intelligence, 2, 7.

Obulesu, O., Mahendra, M., & ThrilokReddy, M. (2018, July). Machine learning techniques and tools: A survey. In 2018 international conference on inventive research in computing applications (ICIRCA) (pp. 605-611). IEEE.

Onyutha, C. (2022). A hydrological model skill score and revised R-squared. Hydrology Research, 53(1), 51-64.

Oruh, J., Viriri, S., & Adegun, A. (2022). Long short-term Memory Recurrent neural network for Automatic speech recognition. IEEE Access, 10, 30069-30079.

Osmond, A. B., Hidayat, F., & Supangkat, S. H. (2021). Electronic Commerce Product Recommendation using Enhanced Conjoint Analysis. International Journal of Advanced Computer Science and Applications, 12(11).

Oubedli, A., & Nait Ali, K. (2021). Les leviers du marketing digital et leur Contribution dans la gestion et suivi de la clientèle Cas: El-Baraka banque (Doctoral dissertation, Université Mouloud Mammeri).

P

Pantano, E., Giglio, S., & Dennis, C. (2019). Making sense of consumers' tweets Sentiment outcomes for fast fashion retailers through Big Data analytics. International Journal of Retail & Distribution Management, 47(9), 915–927.

Paralič, J., Kaszoni, T., & Mačina, J. (2020). Predicting suitable time for sending marketing emails. In Information Systems Architecture and Technology: Proceedings of 40th Anniversary International Conference on Information Systems Architecture and Technology–ISAT 2019: Part II (pp. 189-196). Springer International Publishing.

Patel, N., & Trivedi, S. (2020). Leveraging Predictive Modeling, Machine Learning Personalization, NLP Customer Support, and AI Chatbots to Increase Customer Loyalty. Empirical Quests for Management Essences, 3(3), 1-24.

Patel, N., & Trivedi, S. (2020). Leveraging Predictive Modeling, Machine Learning Personalization, NLP Customer Support, and AI Chatbots to Increase Customer Loyalty. Empirical Quests for Management Essences, 3(3), 1-24.

Pateria, S., Subagdja, B., Tan, A. H., & Quek, C. (2021). Hierarchical reinforcement learning: A comprehensive survey. ACM Computing Surveys (CSUR), 54(5), 1-35.

Patrutiu-Baltes, L. (2016). Inbound Marketing-the most important digital marketing strategy. Bulletin of the Transilvania University of Brasov. Economic Sciences. Series V, 9(2), 61.

Peng, L., Cui, G., & Chung, Y. H. (2020). Do the pieces fit? Assessing the configuration effects of promotion attributes. Journal of Business Research, 109, 337–349.

Peter, M. K., & Dalla Vecchia, M. (2021). The digital marketing toolkit: a literature review for the identification of digital marketing channels and platforms. New trends in business information systems and technology: Digital innovation and digital business transformation, 251-265.

Peyravi, B., Nekrošienė, J., & Lobanova, L. (2020). Revolutionised technologies for marketing: Theoretical review with focus on artificial intelligence. Business: Theory and Practice, 21(2), 827-834.

Philp, M., Jacobson, J., & Pancer, E. (2022). Predicting social media engagement with computer vision: An examination of food marketing on Instagram. Journal of Business Research, 149, 736-747.

Pisner, D. A., & Schnyer, D. M. (2020). Support vector machine. In Machine learning (pp. 101-121). Academic Press.

Pistol, L., & Bucea-Manea, R. Ţ. (2017). The „7Ps" &" 1G" that rule in the digital world the marketing mix. In Proceedings of the International Conference on Business Excellence (Vol. 11, No. 1, pp. 759-769).

Poecze, F., Ebster, C., & Strauss, C. (2019). Let's play on Facebook: Using sentiment analysis and social media metrics to measure the success of YouTube gamers' post types. Personal and Ubiquitous Computing, 1–10.

Pogorelova, I. (2019). Optimization of e-mail marketing of JUSTFAB in European countries.

Polce, E. M., & Kunze, K. N. (2023). A guide for the application of statistics in biomedical studies concerning machine learning and artificial intelligence. Arthroscopy: The Journal of Arthroscopic & Related Surgery, 39(2), 151-158.

Polosukhin, I. (2016, November 19). TensorFlow— Text Classification. Medium. Retrieved from https://medium.com/@ilblackdragon/tensorflow-text-classification615198df9231/

Porter, M. E. (2008). On competition. Harvard Business Press.

Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems, 9, 181-199.

Priya Varshini, A. G., & Anitha Kumari, K. (2020). Predictive analytics approaches for software effort estimation: A review. Indian J. Sci. Technol, 13, 2094-2103.

Provotar, O. I., Linder, Y. M., & Veres, M. M. (2019, December). Unsupervised anomaly detection in time series using lstm-based autoencoders. In 2019 IEEE International Conference on Advanced Trends in Information Theory (ATIT) (pp. 513-517). IEEE.

Prudnikov, Y., & Nazarenko, A. (2021). The role of content marketing in the promotion of medical goods and services.

Puggini, L., & McLoone, S. (2018). An enhanced variable selection and Isolation Forest based methodology for anomaly detection with OES data. Engineering Applications of Artificial Intelligence, 67, 126-135.

Páramo, Á. J. L., García, Á. H., & Peláez, J. C. (2021). Modelling e-mail marketing effectiveness–An approach based on the theory of hierarchy-of-effects. Cuadernos de Gestión, 21(1).

Q

Qiao, M., Bian, W., Da Xu, R. Y., & Tao, D. (2015). Diversified hidden Markov models for sequential labeling. IEEE Transactions on Knowledge and Data Engineering, 27(11), 2947-2960.

Quelch, J. A., & Klein, L. R. (2007). The Internet and international marketing. Readings in modern marketing, 381-385.

Quesenberry, K. A., & Coolsen, M. K. (2019). Drama goes viral: Effects of story development on shares and views of online advertising videos. Journal of Interactive Marketing, 48, 1–16.

R

Rafiq, M., & Ahmed, P. K. (1995). Using the 7Ps as a generic marketing mix: An exploratory survey of UK and European marketing academics. Marketing Intelligence & Planning., 13(9), 4–15.

Rain, C. (2013). Sentiment analysis in amazon reviews using probabilistic machine learning. Swarthmore College.

Rajesh, M., Senapati, B., Das, R., & Martha, S. (2023). Identifying Colorectal Tumor For Single Cell RNA Sequence Using Rectified Linear Unit With Stochastic Gradient Descent. Procedia Computer Science, 218, 189-198.

Ramya, R., Anandh, A., Muthulakshmi, K., & Venkatesh, S. (2022). Gender recognition from facial images using multichannel deep learning framework. In Machine Learning for Biometrics (pp. 105-128). Academic Press.

Rathore, B. (2022). Impact of Green Marketing on Sustainable Business Development. Cardiff Metropolitan University. Presentation.

Rathore, B. (2023). Digital Transformation 4.0: Integration of Artificial Intelligence & Metaverse in Marketing. Eduzone: International Peer Reviewed/Refereed Multidisciplinary Journal, 12(1), 42-48.

Rauschnabel, P. A., Babin, B. J., tom Dieck, M. C., Krey, N., & Jung, T. (2022). What is augmented reality marketing? Its definition, complexity, and future. Journal of Business Research, 142, 1140-1150.

Ray, S. (2019, February). A quick review of machine learning algorithms. In 2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon) (pp. 35-39). IEEE.

Renigier-Biłozor, M., Janowski, A., Walacik, M., & Chmielewska, A. (2021). Human emotion recognition in the significance assessment of property attributes. Journal of Housing and the Built Environment, 1–34.

Reza Chowdhury, S. B. (2022). Robi's Digital solution analysis and customer relationship management.

Rietveld, R., Dolen, W. V., Mazloom, M., & Worring, M. (2020). What you feel, is what you like influence of message appeals on customer engagement on Instagram. Journal of Interactive Marketing, 49, 20–53.

Rizki, B., Ginasta, N. G., Tamrin, M. A., & Rahman, A. (2020). Customer Loyality Segmentation On Point Of Sale System Using Recency-Frequency-Monetary (RFM) and K-Means. Jurnal online informatika, 5(2), 130-136.

Rosenbaum, R. (2022). On the relationship between predictive coding and backpropagation. Plos one, 17(3), e0266102.

Rosenberg, L. (2022, March). Regulation of the Metaverse: A Roadmap: The risks and regulatory solutions for largescale consumer platforms. In Proceedings of the 6th International Conference on Virtual and Augmented Reality Simulations (pp. 21-26).

Ruder, S. (2017). An overview of multi-task learning in deep neural networks. arXiv preprint arXiv:1706.05098.

Rutz, O. J., Bucklin, R. E., & Sonnier, G. P. (2012). A latent instrumental variables approach to modeling keyword conversion in paid search advertising. Journal of Marketing Research, 49(3), 306-319.

Ryoo, J. H., Wang, X., & Lu, S. J. (2020). Do spoilers really spoil? Using topic modeling to measure the effect of spoiler reviews on box office revenue. Journal of Marketing, 85 (2), 70–88.

S

Sachdev, R. (2020, April). Towards security and privacy for edge AI in IoT/IoE based digital marketing environments. In 2020 Fifth International Conference on Fog and Mobile Edge Computing (FMEC) (pp. 341-346). IEEE.

Sadeq, N., Nassreddine, G., & Younis, J. (2023). Impact of Artificial Intelligence on E-marketing.

Saha, T., Gupta, D., Saha, S., & Bhattacharyya, P. (2020). Towards integrated dialogue policy learning for multiple domains and intents using hierarchical deep reinforcement learning. Expert Systems with Applications, 162, 113650.

Sakas, D. P., Reklitis, D. P., Terzi, M. C., & Vassilakis, C. (2022). Multichannel Digital Marketing Optimizations through Big Data Analytics in the Tourism and Hospitality Industry. Journal of Theoretical and Applied Electronic Commerce Research, 17(4), 1383-1408.

Salminen, J., Yoganathan, V., Corporan, J., Jansen, B. J., & Jung, S. G. (2019). Machine learning approach to auto-tagging online content for content marketing efficiency: A comparative analysis between methods and content type. Journal of Business Research, 101, 203–217.

Sanghi, N., & Sanghi, N. (2021). Combining Policy Gradient and Q-Learning. Deep Reinforcement Learning with Python: With PyTorch, TensorFlow and OpenAI Gym, 251-303.

Saputra, T. R., & Sunardi, L. (2022, September). Web Mobile Based on E-Marketing at Jumiran Stores using the Customer Relationship Management (CRM) Method. In Adpebi Science Series, Proceedings of Adpebi International Conference on Management, Education, Social Science, Economics and Technology (AICMEST), 1(1).

Saravanakumar, M., & SuganthaLakshmi, T. (2012). Social media marketing. Life science journal, 9(4), 4444-4451.

Sarkar, S., Ejaz, N., Maiti, J., & Pramanik, A. (2022). An integrated approach using growing self-organizing map-based genetic K-means clustering and tolerance rough set in occupational risk analysis. Neural computing and applications, 34(12), 9661-9687.

Sattigeri, P., Hoffman, S. C., Chenthamarakshan, V., & Varshney, K. R. (2019). Fairness GAN: Generating datasets with fairness properties using a generative adversarial network. IBM Journal of Research and Development, 63(4/5), 3-1.

Saunders, W., Sastry, G., Stuhlmueller, A., & Evans, O. (2017). Trial without error: Towards safe reinforcement learning via human intervention. arXiv preprint arXiv:1707.05173.

Saura, J. R. (Ed.). (2021). Advanced digital marketing strategies in a data-driven era. IGI Global.

Schaefer, J. W. (2018). Content Marketing: Essential Guide to Learn Step-by-Step the Best Content Marketing Strategies to Attract your Audience and Boost Your Business. EC Publishing via PublishDrive.

Schaeffer, S. E., & Sanchez, S. V. R. (2020). Forecasting client retention - A machinelearning approach. Journal of Retailing and Consumer Services, 52, 1–9.

Schalkwyk, J., Beeferman, D., Beaufays, F., Byrne, B., Chelba, C., Cohen, M., & Strope, B. (2010). “Your word is my command”: Google search by voice: A case study. Advances in Speech Recognition: Mobile Environments, Call Centers and Clinics, 61-90.

Schein, R., Wilson, K., & Keelan, J. E. (2011). Literature review on effectiveness of the use of social media: a report for peel public health. [Region of Peel], Peel Public Health.

Schmidt, A., Kabir, M. W. U., & Hoque, M. T. (2022). Machine learning based restaurant sales forecasting. Machine Learning and Knowledge Extraction, 4(1), 105-130. doi: 10.3390/make4010006

Schoettler, G., Nair, A., Luo, J., Bahl, S., Ojea, J. A., Solowjow, E., & Levine, S. (2020, October). Deep reinforcement learning for industrial insertion tasks with visual inputs and natural rewards. In 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 5548-5555). IEEE.

Schrouff, J., Mourão-Miranda, J., Phillips, C., & Parvizi, J. (2016). Decoding intracranial EEG data with multiple kernel learning method. Journal of neuroscience methods, 261, 19-28.

Schwartz, E. M., Bradlow, E. T., & Fader, P. S. (2017). Customer acquisition via display advertising using multi-armed bandit experiments. Marketing Science, 36(4), 500–522.

Scott, D. M. (2009). The new rules of marketing and PR: how to use social media, blogs, news releases, online video, and viral marketing to reach buyers directly. John Wiley & Sons.

Seymour, T., Frantsvog, D., & Kumar, S. (2011). History of search engines. International Journal of Management & Information Systems (IJMIS), 15(4), 47-58.

Shabazz, D. (2005). Toward a better understanding of e-marketing strategy: Past and present. Services Marketing Quarterly, 26(2), 117-130.

Shah, D., & Murthi, B. P. S. (2021). Marketing in a data-driven digital world: Implications for the role and scope of marketing. Journal of Business Research, 125, 772-779.

Shah, K., Patel, H., Sanghvi, D., & Shah, M. (2020). A comparative analysis of logistic regression, random forest and KNN models for the text classification. Augmented Human Research, 5, 1-16.

Shakya, P. B., Sinha, P., Nair, P. P., Rao, S. R., & Banerjee, T. (2022). Role of Content Marketing as a Sustainable Business Strategy. ECS Transactions, 107(1), 13221.

Shaltoni, A. M. (2016). E-marketing education in transition: An analysis of international courses and programs. The International Journal of Management Education, 14(2), 212-218.

Shandilya, P., Murphy, L., & Perales, F. (2023). Artificial intelligence for smart bidding. Journal of AI, Robotics & Workplace Automation, 2(2), 153-164.

Shandilya, P., Murphy, L., & Perales, F. (2023). Artificial intelligence for smart bidding. Journal of AI, Robotics & Workplace Automation, 2(2), 153-164.

Shen, B. (2021, January). E-commerce Customer Segmentation via Unsupervised Machine Learning. In The 2nd International Conference on Computing and Data Science (pp. 1-7).

Shetty, P., & Singh, S. (2021). Hierarchical clustering: a survey. International Journal of Applied Research, 7(4), 178-181.

Shetty, P., & Singh, S. (2021). Hierarchical clustering: a survey. International Journal of Applied Research, 7(4), 178-181.

Shih, B. Y., Chen, C. Y., & Chen, Z. S. (2013). Retracted: an empirical study of an internet marketing strategy for search engine optimization. Human Factors and Ergonomics in Manufacturing & Service Industries, 23(6), 528-540.

Shihab, S. H., Afroge, S., & Mishu, S. Z. (2019, February). RFM based market segmentation approach using advanced k-means and agglomerative clustering: a comparative study. In 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE) (pp. 1-4). IEEE.

Shuai, H., & He, H. (2020). Online scheduling of a residential microgrid via Monte-Carlo tree search and a learned model. IEEE Transactions on Smart Grid, 12(2), 1073-1087.

Siddika, A., Faruque, A., & Masum, A. K. M. (2021). Comparative Analysis of Churn Predictive Models and Factor Identification in Telecom Industry. In 2021 24th International Conference on Computer and Information Technology. ICCIT (pp. 1-6). doi: 10.1109/ICCIT54785.2021.9689881

Sideratos, G., Ikonomopoulos, A., & Hatziargyriou, N. D. (2020). A novel fuzzy-based ensemble model for load forecasting using hybrid deep neural networks. Electric Power Systems Research, 178, 106025.

Siegel, E. (2013). Predictive analytics: The power to predict who will click, buy, lie, or die. John Wiley & Sons.

Simmonds, L., Bellman, S., Kennedy, R., Nenycz-Thiel, M., & Bogomolova, S. (2020). Moderating effects of prior brand usage on visual attention to video advertising and recall: An eye-tracking investigation. Journal of Business Research, 111, 241–248.

Sindhu, C., & Mukherjee, D. (2021, April). A Joint Sentiment-Topic Model for Product Review Analysis of Electronic Goods. In 2021 5th International Conference on Computing Methodologies and Communication (ICCMC) (pp. 574-578). IEEE.

Singh, M., Ghutla, B., Jnr, R. L., Mohammed, A. F., & Rashid, M. A. (2017, December). Walmart's Sales Data Analysis-A Big Data Analytics Perspective. In 2017 4th Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE) (pp. 114-119). IEEE.

Smirnov, D., & Huchzermeier, A. (2020). Analytics for labor planning in systems with load-dependent service times. European Journal of Operational Research, 287(2), 668–681.

Soedarsono, D. K., Mohamad, B., Adamu, A. A., & Pradita, K. A. (2020). Managing Digital Marketing Communication of Coffee Shop Using Instagram. International Journal of Interactive Mobile Technologies, 14(5).

Sohony, I., Pratap, R., & Nambiar, U. (2018, January). Ensemble learning for credit card fraud detection. In Proceedings of the ACM India Joint International Conference on Data Science and Management of Data (pp. 289-294).

Soleimani, E., & Nazerfard, E. (2021). Cross-subject transfer learning in human activity recognition systems using generative adversarial networks. Neurocomputing, 426, 26-34.

Somvanshi, M., Chavan, P., Tambade, S., & Shinde, S. V. (2016, August). A review of machine learning techniques using decision tree and support vector machine. In 2016 international conference on computing communication control and automation (ICCUBEA) (pp. 1-7). IEEE.

Song, Y., & Ermon, S. (2020). Improved techniques for training score-based generative models. Advances in neural information processing systems, 33, 12438-12448.

Srivastava, R. K., Shervani, T. A., & Fahey, L. (1999). Marketing, business processes, and shareholder value: an organizationally embedded view of marketing activities and the discipline of marketing. Journal of marketing, 63(4_suppl1), 168-179.

Srivastava, V., & Kalro, A. D. (2019). Enhancing the helpfulness of online consumer reviews: the role of latent (content) factors. Journal of Interactive Marketing, 48, 33–50.

Stikic, M., Larlus, D., & Schiele, B. (2009, September). Multi-graph based semi-supervised learning for activity recognition. In 2009 international symposium on wearable computers (pp. 85-92). IEEE.

Struye, J., & Latré, S. (2020). Hierarchical temporal memory and recurrent neural networks for time series prediction: An empirical validation and reduction to multilayer perceptrons. Neurocomputing, 396, 291-301.

Stucke, M. E., & Ezrachi, A. (2016). When competition fails to optimize quality: A look at search engines. Yale JL & Tech., 18, 70.

Su, Xueping, Gao, Meng, Ren, Jie, Li, Yunhong, R¨ atsch, Matthias. (2020). Personalized clothing recommendation based on user emotional analysis. Discrete Dynamics in Nature and Society, 2020.

Subramanian, M., Shanmugavadivel, K., & Nandhini, P. S. (2022). On fine-tuning deep learning models using transfer learning and hyper-parameters optimization for disease identification in maize leaves. Neural Computing and Applications, 34(16), 13951-13968.

Sueyoshi, T., & Tadiparthi, G. R. (2005). A wholesale power trading simulator with learning capabilities. IEEE Transactions on Power Systems, 20(3), 1330–1340.

Sueyoshi, T., & Tadiparthi, G. R. (2008). An agent-based decision support system for wholesale electricity market. Decision Support Systems, 44(2), 425–446.

Suh, T., & Chow, T. E. (2021). Developing a digital marketing tool for ethnic ventures' mixed business model and market-shaping: A design scientific approach of web demographics. Industrial Marketing Management, 93, 10-21.

Sun, Z. L., Choi, T. M., Au, K. F., & Yu, Y. (2008). Sales forecasting using extreme learning machine with applications in fashion retailing. Decision Support Systems, 46(1), 411-419. doi: 10.1016/j.dss.2008.07.009

Sutton, C. D. (2005). Classification and regression trees, bagging, and boosting. Handbook of statistics, 24, 303-329.

Sümer, E., Sümer, S. I., & Oğul, H. (2018). A Novel Computer Vision-Based Advertisement System for Individual Marketing. Tehnički vjesnik, 25(5), 1485-1491.

Škare, V., & Gospić, D. (2015). Dynamic pricing and customers' perceptions of price fairness in the airline industry. Tourism: An International Interdisciplinary Journal, 63(4), 515-528.

T

Taheri, S., & Bouyer, A. (2020). Community detection in social networks using affinity propagation with adaptive similarity matrix. Big data, 8(3), 189-202.

Tallec, C. (2019). Recurrent Neural Networks and Reinforcement Learning: Dynamic Approaches (Doctoral dissertation, Université Paris-Saclay).

Tamposis, I. A., Tsirigos, K. D., Theodoropoulou, M. C., Kontou, P. I., & Bagos, P. G. (2019). Semi-supervised learning of Hidden Markov Models for biological sequence analysis. Bioinformatics, 35(13), 2208-2215.

Tan, C. W., Bergmeir, C., Petitjean, F., & Webb, G. I. (2021). Time series extrinsic regression: Predicting numeric values from time series data. Data Mining and Knowledge Discovery, 35, 1032-1060.

Tarwani, K. M., & Edem, S. (2017). Survey on recurrent neural network in natural language processing. Int. J. Eng. Trends Technol, 48(6), 301-304.

Thomas, T., P. Vijayaraghavan, A., Emmanuel, S., Thomas, T., P. Vijayaraghavan, A., & Emmanuel, S. (2020). Applications of decision trees. Machine learning approaches in cyber security analytics, 157-184.

Thomassey, S., Happiette, M., & Castelain, J. M. (2005). A global forecasting support system adapted to textile distribution. International Journal of Production Economics, 96(1), 81-95. doi: 10.1016/j.ijpe.2004.03.001

Thudumu, S., Branch, P., Jin, J., & Singh, J. (2020). A comprehensive survey of anomaly detection techniques for high dimensional big data. Journal of Big Data, 7, 1-30.

Timoshenko, A., & Hauser, J. R. (2019). Identifying customer needs from user-generated content. Marketing Science, 38(1), 1–20.

Tirunillai, S., & Tellis, G. J. (2014). Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent dirichlet allocation. Journal of Marketing Research, 51(4), 463–479.

Torfi, A., Shirvani, R. A., Keneshloo, Y., Tavaf, N., & Fox, E. A. (2020). Natural language processing advancements by deep learning: A survey. arXiv preprint arXiv:2003.01200.

Trappey, C. V., Trappey, A. J., & Lin, S. C. C. (2020). Intelligent trademark similarity analysis of image, spelling, and phonetic features using machine learning methodologies. Advanced Engineering Informatics, 45, 1-12.

Trardi, Y., Ananou, B., Tchatchoua, P., & Ouladsine, M. (2022, July). Ensemble Machine Learning Algorithms for Anomaly Detection in Multivariate Time-Series. In 2022 International Conference on Control, Automation and Diagnosis (ICCAD) (pp. 1-6). IEEE.

Tsao, H. Y., Campbell, C. L., Sands, S., Ferraro, C., Mavrommatis, A., & Lu, S. (2019). A machine-learning based approach to measuring constructs through text analysis. European Journal of Marketing, 54(3), 511–524.

Tsimonis, G., & Dimitriadis, S. (2014). Brand strategies in social media. Marketing Intelligence & Planning, 32(3), 328-344.

Tsiotsou, R., & Ratten, V. (2010). Future research directions in tourism marketing. Marketing intelligence & planning, 28(4), 533-544.

Tu, H., & Menzies, T. (2022). DebtFree: minimizing labeling cost in self-admitted technical debt identification using semi-supervised learning. Empirical Software Engineering, 27(4), 80.

Tulshan, A. S., & Dhage, S. N. (2019). Survey on virtual assistant: Google assistant, siri, cortana, alexa. In Advances in Signal Processing and Intelligent Recognition Systems: 4th International Symposium SIRS 2018, Bangalore, India, September 19–22, 2018, Revised Selected Papers 4 (pp. 190-201). Springer Singapore.

Turow, J. (1997). Breaking up America: Advertisers and the new media world. University of Chicago Press.

Turow, J. (2005). Audience construction and culture production: Marketing surveillance in the digital age. The ANNALS of the American Academy of Political and Social Science, 597(1), 103-121.

Turow, J. (2008). Niche envy: Marketing discrimination in the digital age. MIT Press.

Tyagi, B., & Bhatnagar, V. (2021). Hyper-Personalized Recommendation Systems: A Systematic Literature Mapping. Disruptive Technologies for Society 5.0, 69-86.

U

Ullah, I., Raza, B., Malik, A. K., Imran, M., Ul Islam, S., & Kim, S. W. (2019). A churn prediction model using random forest: Analysis of machine learning techniques for churn prediction and factor identification in telecom sector. IEEE Access, 7, 60134–60149.

Ur Rehman, S., Tu, S., Waqas, M., Huang, Y., Ur Rehman, O., Ahmad, B., & Ahmad, S. (2019). Unsupervised pre-trained filter learning approach for efficient convolution neural network. Neurocomputing, 365, 171-190.

V

Vakeel, A., Vantari, N. R., Reddy, S. N., Muthyapu, R., & Chavan, A. (2022). Machine Learning Models for Predicting and Clustering Customer Churn Based on Boosting Algorithms and Gaussian Mixture Model. In 2022 International Conference for Advancement in Technology. ICONAT (pp. 1-5). doi: 10.1109/ICONAT53423.2022.9725957

Valdez Mendia, J. M., & Flores-Cuautle, J. D. J. A. (2022). Toward customer hyper-personalization experience—A data-driven approach. Cogent Business & Management, 9(1), 2041384.

Van Hasselt, H., Guez, A., & Silver, D. (2016, March). Deep reinforcement learning with double q-learning. In Proceedings of the AAAI conference on artificial intelligence (Vol. 30, No. 1).

Vavliakis, K. N., Kotouza, M. T., Symeonidis, A. L., & Mitkas, P. A. (2018). Recommendation Systems in a Conversational Web. In WEBIST (pp. 68-77).

Vavliakis, K. N., Kotouza, M. T., Symeonidis, A. L., & Mitkas, P. A. (2019). Personalization and the Conversational Web. In Web Information Systems and Technologies: 14th International Conference, WEBIST 2018, Seville, Spain, September 18–20, 2018, Revised Selected Papers 14 (pp. 56-77). Springer International Publishing.

Verhoef, P. C., Kooge, E., & Walk, N. (2016). Creating value with big data analytics: Making smarter marketing decisions. Routledge.

Vermeer, S. A. M., Araujo, T., Bernritter, S. F., & van Noort, G. (2019). Seeing the wood for the trees: How machine learning can help firms in identifying relevant electronic word-of-mouth in social media. International Journal of Research in Marketing, 36(3), 492–508.

Vezhnevets, A. S., Osindero, S., Schaul, T., Heess, N., Jaderberg, M., Silver, D., & Kavukcuoglu, K. (2017, July). Feudal networks for hierarchical reinforcement learning. In International Conference on Machine Learning (pp. 3540-3549). PMLR.

Viglia, G., & Viglia, G. (2014). Online Marketing Communication Channels. Pricing, Online Marketing Behavior, and Analytics, 23-38.

Vij, S. R., More, A., Mukhopadhyay, D., & Agrawal, A. J. (2015). An e-negotiation agent using rule based and case based approaches: A comparative study with bilateral e-negotiation with prediction. Journal of Software Engineering and Applications, 8(10), 521.

Vinaykarthik, B. C. (2022, October). Design of Artificial Intelligence (AI) based User Experience Websites for E-commerce Application and Future of Digital Marketing. In 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC) (pp. 1023-1029). IEEE.

W

Wan, C., & Jones, D. T. (2020). Protein function prediction is improved by creating synthetic feature samples with generative adversarial networks. Nature Machine Intelligence, 2(9), 540-550.

Wang, B., Tu, Z., & Tsotsos, J. K. (2013). Dynamic label propagation for semi-supervised multi-class multi-label classification. In Proceedings of the IEEE international conference on computer vision (pp. 425-432).

Wang, S., Sun, J., & Xu, Z. (2019, July). Hyperadam: A learnable task-adaptive adam for network training. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, No. 01, pp. 5297-5304).

Wang, X., & Ji, H. (2020). Semi-supervised hyperspectral image classification based on label propagation via selected path. IEEE Access, 8, 221225-221234.

Wang, X., Nguyen, K., & Nguyen, B. P. (2020). Churn Prediction using Ensemble Learning. In Proceedings of the 4th International Conference on Machine Learning and Soft Computing. doi:10.1145/3380688.3380710

Wang, Y. F., Ma, K., Garcia-Hernandez, L., Chen, J., Hou, Z. H., Ji, K., … Abraham, A. (2020). A CLSTM-TMN for marketing intention detection. Engineering Applications of Artificial Intelligence, 91, 1–9.

Wang, Z., Maalla, A., & Liang, M. (2021, December). Research on e-commerce personalized recommendation system based on big data technology. In 2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA) (Vol. 2, pp. 909-913). IEEE.

Wang, Z., O'Dwyer, K., Muddiman, R., Ward, T., Camp Jr, C. H., & Hennelly, B. M. (2022). VECTOR: Very deep convolutional autoencoders for non-resonant background removal in broadband coherent anti-Stokes Raman scattering. Journal of Raman Spectroscopy, 53(6), 1081-1093.

Weideman, M., & Kritzinger, W. (2017). Parallel search engine optimisation and pay-per-click campaigns: A comparison of cost per acquisition. South African Journal of Information Management, 19(1), 1-13.

Weng, L. (2019). From gan to wgan. arXiv preprint arXiv:1904.08994.

Wiatrak, M., Albrecht, S. V., & Nystrom, A. (2019). Stabilizing generative adversarial networks: A survey. arXiv preprint arXiv:1910.00927.

Wind, Y., & Mahajan, V. (2002). Convergence marketing. Journal of Interactive Marketing, 16(2), 64-79.

Wolkenfelt, M. R. J., & Situmeang, F. B. I. (2020). Effects of app pricing structures on product evaluations. Journal of Research in Interactive Marketing, 14(1), 89–110.

Wu, L., Dodoo, N. A., Wen, T. J., & Ke, L. (2022). Understanding Twitter conversations about artificial intelligence in advertising based on natural language processing. International Journal of Advertising, 41(4), 685-702.

Wu, S., Yau, W. C., Ong, T. S., & Chong, S. C. (2021). Integrated churn prediction and customer segmentation framework for telco business. IEEE Access, 9, 62118-62136.

Wu, Y. Y., Ngai, E. W. T., Wu, P. K., & Wu, C. (2020). Fake online reviews: Literature review, synthesis, and directions for future research. Decision Support Systems, 132, Article 113280.

Wu, Y. Y., Ngai, E. W. T., Wu, P. K., & Wu, C. (2020). Fake online reviews: Literature review, synthesis, and directions for future research. Decision Support Systems, 132, Article 113280.

Wyffels, F., Schrauwen, B., & Stroobandt, D. (2008). Stable output feedback in reservoir computing using ridge regression. In Artificial Neural Networks-ICANN 2008: 18th International Conference, Prague, Czech Republic, September 3-6, 2008, Proceedings, Part I 18 (pp. 808-817). Springer Berlin Heidelberg.

X

Xia, Y., Liu, F., Yang, D., Cai, J., Yu, L., Zhu, Z., & Roth, H. (2020). 3d semi-supervised learning with uncertainty-aware multi-view co-training. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 3646-3655).

Xia, Y., Yang, D., Yu, Z., Liu, F., Cai, J., Yu, L., & Roth, H. (2020). Uncertainty-aware multi-view co-training for semi-supervised medical image segmentation and domain adaptation. Medical image analysis, 65, 101766.

Xie, D., & He, Y. (2022). Marketing Strategy of Rural Tourism Based on Big Data and Artificial Intelligence. Mobile Information Systems, 2022.

Xie, Y., Lin, T., Chen, Z., Xiong, W., Ran, Q., & Shang, C. (2022). A lightweight ensemble discriminator for Generative Adversarial Networks. Knowledge-Based Systems, 250, 108975.

Xu, J., Zhang, H., & Qiu, J. (2022). A deep deterministic policy gradient algorithm based on averaged state-action estimation. Computers and Electrical Engineering, 101, 108015.

Xu, R. B., Wang, M. L., & Xie, Y. (2018). Optimally connected deep belief net for click through rate prediction in online advertising. IEEE Access, 6, 43009–43020.

Xu, Y., Wu, X., & Wang, Q. (2015). Sentiment Analysis of Yelp ‘s Ratings Based on Text Reviews. Stanford University, 17, 117-120.

Xu, Z., Kakde, D., & Chaudhuri, A. (2019, December). Automatic hyperparameter tuning method for local outlier factor, with applications to anomaly detection. In 2019 IEEE International Conference on Big Data (Big Data) (pp. 4201-4207). IEEE.

Xuecong, C., Zhaoming, L., & Sisi, C. (2021, April). Design and Implementation of E-commerce Recommendation System Model Based on Cloud Computing. In 2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC) (pp. 1100-1103). IEEE.

Y

Yan, R. (2010). Cooperative advertising, pricing strategy and firm performance in the e-marketing age. Journal of the Academy of Marketing Science, 38, 510-519.

Yang, L., Li, Y., Wang, J., & Sherratt, R. S. (2020). Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. IEEE access, 8, 23522-23530.

Yang, W., Sun, S., Hao, Y., & Wang, S. (2022). A novel machine learning-based electricity price forecasting model based on optimal model selection strategy. Energy, 238, Article 121989.

Yang, Y., Xu, B., Shen, S., Shen, F., & Zhao, J. (2020). Operation-aware Neural Networks for user response prediction. Neural Networks, 121, 161–168.

Yang, Y., Yeh, H. G., Zhang, W., Lee, C. J., Meese, E. N., & Lowe, C. G. (2020). Feature extraction, selection, and K-nearest neighbors algorithm for shark behavior classification based on imbalanced dataset. IEEE Sensors Journal, 21(5), 6429-6439.

Yang, Y., Zhang, W., & Li, P. (2021, July). Backpropagated neighborhood aggregation for accurate training of spiking neural networks. In International Conference on Machine Learning (pp. 11852-11862). PMLR.

Ye, S., Jiang, T., Qin, S., Zou, W., & Deng, C. (2018, November). Speech enhancement based on a new architecture of wasserstein generative adversarial networks. In 2018 11th International Symposium on Chinese Spoken Language Processing (ISCSLP) (pp. 399-403). IEEE.

Yildirim, S. (2022). Natural Language Processing Application in Engineering.

Yolcu, G., Oztel, I., Kazan, S., Oz, C., & Bunyak, F. (2020). Deep learning-based face analysis system for monitoring customer interest. Journal of Ambient Intelligence and Humanized Computing, 11(1), 237–248.

Yost, E., Zhang, T., & Qi, R. (2021). The power of engagement: Understanding active social media engagement and the impact on sales in the hospitality industry. Journal of Hospitality and Tourism Management, 46, 83-95.

You, A., Kim, J. K., Ryu, I. H., & Yoo, T. K. (2022). Application of generative adversarial networks (GAN) for ophthalmology image domains: a survey. Eye and Vision, 9(1), 1-19.

Yu, C. M., Zuo, Y. H., Feng, B. L., An, L., & Chen, B. Y. (2019). An individual-groupmerchant relation model for identifying fake online reviews: An empirical study on a Chinese e-commerce platform. Information Technology & Management, 20(3), 123–138.

Yun, C., Shun, M., Junta, U., & Browndi, I. (2022). Predictive Analytics: A Survey, Trends, Applications, Opportunities' and Challenges for Smart City Planning. International Journal of Computer Science and Information Technology, 23(56), 226-231.

Z

Zadoo, A., Jagtap, T., Khule, N., Kedari, A., & Khedkar, S. (2022). A review on Churn Prediction and Customer Segmentation using Machine Learning. In 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing. COM-IT-CON (pp. 174-178). doi: 10.1109/COM-IT-CON54601.2022.9850924

Zaman, H. R. R., & Gharehchopogh, F. S. (2022). An improved particle swarm optimization with backtracking search optimization algorithm for solving continuous optimization problems. Engineering with Computers, 38(Suppl 4), 2797-2831.

Zamfirache, I. A., Precup, R. E., Roman, R. C., & Petriu, E. M. (2022). Reinforcement Learning-based control using Q-learning and gravitational search algorithm with experimental validation on a nonlinear servo system. Information Sciences, 583, 99-120.

Zarouali, B., Van den Broeck, E., Walrave, M., & Poels, K. (2018). Predicting consumer responses to a Chatbot on Facebook. Cyberpsychology Behavior and Social Networking, 21(8), 491–497

Zeng, H., & Pan, D. (2010, July). A knowledge discovery and data mining process model in E-marketing. In 2010 8th World Congress on Intelligent Control and Automation (pp. 3960-3964). IEEE.

Zhan, X., Liu, Z., Yan, J., Lin, D., & Loy, C. C. (2018). Consensus-driven propagation in massive unlabeled data for face recognition. In Proceedings of the European conference on computer vision (ECCV) (pp. 568-583).

Zhang, B., Zhu, J., & Su, H. (2023). Toward the third-generation artificial intelligence. Science China Information Sciences, 66(2), 1-19.

Zhang, D. S., Zhou, L., Kehoe, J. L., & Kilic, I. Y. (2016). What online reviewer behaviors really matter? Effects of verbal and nonverbal behaviors on detection of fake online reviews. Journal of Management Information Systems, 33(2), 456–481.

Zhang, H., Sun, Y., Liu, L., Wang, X., Li, L., & Liu, W. (2020). ClothingOut: a category-supervised GAN model for clothing segmentation and retrieval. Neural computing and applications, 32, 4519-4530.

Zhang, S., & Cabage, N. (2017). Search engine optimization: Comparison of link building and social sharing. Journal of Computer Information Systems, 57(2), 148-159.

Zhang, W., & Li, P. (2019). Spike-train level backpropagation for training deep recurrent spiking neural networks. Advances in neural information processing systems, 32.

Zhang, X., Zhang, Z., Liang D., & Jin, H. (2018). A Novel Decision Tree Based on Profit Variance Maximization Criterion for Customer Churn Problem. 2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), 20-23, doi: 10.1109/IHMSC.2018.00013.

Zhang, Y., & Haghani, A. (2015). A gradient boosting method to improve travel time prediction. Transportation Research Part C: Emerging Technologies, 58, 308-324.

Zhang, Y., Bai, R., Qu, R., Tu, C., & Jin, J. (2022). A deep reinforcement learning based hyper-heuristic for combinatorial optimisation with uncertainties. European Journal of Operational Research, 300(2), 418-427.

Zhao, R., & Li, C. (2021, July). Research on E-commerce Customer Segmentation Based on RFAC Model. In 2021 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS) (pp. 439-444). IEEE.

Zhao, X., Guo, J., Nie, F., Chen, L., Li, Z., & Zhang, H. (2019). Joint principal component and discriminant analysis for dimensionality reduction. IEEE transactions on neural networks and learning systems, 31(2), 433-444.

Zhao, Y., Jin, Z., Qi, G. J., Lu, H., & Hua, X. S. (2018). An adversarial approach to hard triplet generation. In Proceedings of the European conference on computer vision (ECCV) (pp. 501-517).

Zhong, C., Wang, H., & Yang, Q. (2022). Hydrochemical interpretation of groundwater in Yinchuan basin using self-organizing maps and hierarchical clustering. Chemosphere, 309, 136787.

Zhou, G. B., Wu, J., Zhang, C. L., & Zhou, Z. H. (2016). Minimal gated unit for recurrent neural networks. International Journal of Automation and Computing, 13(3), 226-234.

Zhou, H., Huang, J., Lu, F., Thiyagalingam, J., & Kirubarajan, T. (2018). Echo state kernel recursive least squares algorithm for machine condition prediction. Mechanical Systems and Signal Processing, 111, 68-86.

Zhou, M., Liu, Z., Sui, P., Li, Y., & Chung, Y. Y. (2020). Learning implicit credit assignment for cooperative multi-agent reinforcement learning. Advances in neural information processing systems, 33, 11853-11864.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., & He, Q. (2020). A comprehensive survey on transfer learning. Proceedings of the IEEE, 109(1), 43-76.

Zhuĺ, X., & Ghahramaniĺн, Z. (2002). Learning from labeled and unlabeled data with label propagation.

Zouaoui, I., & Allouche, M. (2023). Le marketing digital au service de la performance commerciale de l'entreprise tunisienne en période de la crise sanitaire du COVID-19. International Journal of Accounting, Finance, Auditing, Management and Economics, 4(1-1), 436-464.

Zumstein, D., & Hundertmark, S. (2018). Chatbots: An interactive technology for personalized communication and transaction. IADIS International Journal on WWW/ Internet, 15(1), 96–109.