N° d'ordre :

Université Sultan Moulay Slimane
Faculté des Sciences et Techniques
Beni Mellal

Centre d'Études Doctorales : Faculté des Sciences et Techniques
Formation Doctorale : **Mathématiques et Physiques Appliquées**

THÈSE

Présentée par

Ouidad AKHRIF

Pour l'obtention du grade de
DOCTEUR

Spécialité: **Informatique**

Smart collaborative learning based on Semantic Educational data.

Soutenue le 25/05/2022 devant la commission d'examen :

Pr. Hafida HANINE	Professeur, Université Sultan Moulay Slimane, F.S.T. Béni-Mellal, Maroc.	Président
Pr. Younes TABII	Professeur, Université Mohammed V, ENSIAS. Rabat, Maroc.	Rapporteur
Pr. Ilham OUMAIRA	Professeur Habilité, Université Ibn Tofail, ENSA. Kénitra, Maroc.	Rapporteur
Pr. Hanaa HACHIMI	Professeur Habilité, Université Sultan Moulay Slimane, Béni-Mellal, Maroc.	Rapporteur
Pr. Belaid BOUIKHALENE	Professeur, Université Sultan Moulay Slimane, Faculté Polydisciplinaire. Béni-Mellal, Maroc.	Examinateur
Pr. Mostapha EL JAI	Professeur Assistant, Université Moulay Ismail, E.N.S.A.M. Meknès, Maroc.	Invité
Pr.Younès EL BOUZEKRI EL IDRISSI	Professeur Habilité, Université Ibn Tofail, ENSA. Kénitra, Maroc.	Co-Directeur de thèse
Pr. Nabil HMINA	Professeur, Université Sultan Moulay Slimane, Béni-Mellal, Maroc.	Directeur de thèse

Dedicace

To my dear parents,

No words can express my feelings towards you.

God keep and protect you.

To my husband Anass,

No words can express my deep affection and my immense gratitude to you, for all the support you offered me.

To my sons Mohammed Yassine,Taha et Oussama,

No work is more important than my love for you

To my dear sisters and brother,

I don't know how to thank you for everything you have done for me.

To all my family,

All my friends

CONTENTS

List of Figures

List of tables

List of equations

List of acronyms

AI	Artificial Intelligence	**LD**	Learning Design
EBS	Elastic Block Store	**LOM**	Learning Object Metadata
EC2	Elastic Compute Cloud	**ML**	Machine Learning
MSK	Managed Streaming for Kafka	**MOOC**	Massive Open Online Course
AMI	Amazon Machine Instance	**OMS**	Opinion Mining System
CART	Classification and Regression Trees	**RFID**	Radio Frequency Identification Devices
CL	Collaborative Learning	**SAWSDL**	Semantic Annotations for WSDL and XML Schema
EDM	Educational Data Mining	**SC**	Smart City
FOAF	Friend of a friend	**SCT**	Smart Computing Technologies
GPS	Global Positioning Systems	**SOA**	service-oriented architecture
HDFS	Hadoop Distributed File System	**SOC**	service-oriented computing
HTTP	Hypertext Transfer Protocol	**SPARQL**	Simple Protocol and RDF Query Language
IBM	International Business Machines	**SPOC**	Small Private Online Course
ICT	information and communication technologies	**SQL**	Structured Query Language
IEC	International Electrotechnical Commission	**SWSF**	Semantic Web Services Framework
IEEE PAPI	Public and Private Information	**UN**	United Nations
IMS LIP	Learner Information Package	**UNESCO**	United Nations Educational, Scientific and Cultural Organization
IoT	Internet of Things	**URI**	Uniform Ressource Identifier

ISO	International Organization for Standardization	**VPC**	Virtual Private Cloud
OWL	Web Ontology Language	**XML**	Extensible Markup Language
QoS	Quality of Service		
RDF	Resource Description Framework		
REST API	Representational State Transfer Application Programming Interface		
KB	knowledge bases		
KNIME	Konstanz Information Miner		

Acknowledgment

Praise be to God. Prayers and peace be upon our Prophet Muhammad, his family and all of his companions.

Upon completion of this thesis, I am pleased to express my sincere gratitude to my supervisor, Prof. Nabil HMINA, for his immense support, motivation, as well as for the guidance he gave me throughout this research. I warmly appreciate his steering methodology, reflecting his exceptional wisdom and knowledge.

A debt of gratitude is also due to Prof. Younès El Bouzekri El Idrissi, my esteemed co-supervisor, for advising me through relevant guidelines leading to the realization of this Ph.D. I thank him for his patience, kindness, and endless support. His rich experience encouraged me throughout my academic research and daily life.

I would like to extend my sincere thanks to the members of my committee,who kindly accepted to review my thesis work. I thank them a lot for their detailed report, constructive questions, and thoughtful suggestions.

I would like to express my sincere thanks to my parents for their tremendous understanding and encouragement in the past few years, I don't have the words to tell you how lucky I feel to have you in my life. I am thankful for all the prayers sent my way for achieving my success. A heartfelt thank you to my husband and my children for supporting me in good and bad times. I want to thank my sisters and brother, You are great motivators!

Finally, I would like to thank all the people who supported me in completing this Ph.D.

Abstract

Nowadays, the development of so-called “smart” cities is a key focus orientation of growth policy in most countries, through the deployment of pilot projects combining the use of a new generation of information technologies and a responsive governance strategy. As a result, this new urban ecosystem fosters the “smartness” of capacities for planning, construction and digital management of urban services.

To attain this goal, it is vital to pique the enthusiasm of all municipal stakeholders so how they can work together, around a clear vision, to build their future urban landscape. Education, in particular, is highlighted as a key indicator of urban competitiveness and, as a result, is positioned as a critical component of any "smart" transformation of the urban ecosystem.

As a corollary, "Smart Education City" has become a necessary connotation of the smart city. Smart education is a new form of learning in the digital age, as well as a new stage in the computerization of education to encourage change. Smart education, under the auspices of the "Smart University," improves the quality and efficiency of teaching and provides the optimal environment for developing future talent. The Smart University (SU) is a multi-system project with the underlying overall structure: "smart student, smart learning, smart education, smart knowledge, and smart interactions." The "Smart University's" main mission is to smartly impart knowledge to learners who can collaborate internally or externally with the university ecosystem to acquire new skills.

Our research axis is founded on these facts, and it requires first establishing an algorithm-based semantic data that models interdisciplinary collaboration while incorporating a prediction approach in order to optimize complementary team selection. The algorithm will then be servitized in a second phase in order to standardize it as a Smart University (SU) intelligent service.

Keywords: Smart City, Smart University, Smart Service, Interdisciplinary collaboration, Completeness Team, Heuristic, Ontology, Educational Data Mining, Classification, Prediction, Continuous Deployment, Cloud Computing

Résumé

À l'heure actuelle, le développement des villes dites 'intelligentes' constitue une orientation de choix dans la politique d'essor de la plupart des pays, moyennant le déploiement de projets pilotes qui conjuguent l'usage d'une nouvelle génération de technologies de l'information et d'un modèle de gouvernance adapté. Ainsi, ce nouvel écosystème urbain promeut la 'smartness' de la planification, de la construction et de la capacité de gestion numérique des services urbains.

Pour atteindre cet objectif, il est nécessaire de stimuler l'enthousiasme de toutes les parties prenantes, afin qu'elles puissent collaborer, autour d'une vision univoque, pour bâtir leur nouvel environnement urbain. Plus particulièrement, l'éducation est identifiée comme un indicateur important servant à mesurer la compétitivité urbaine, et de ce fait, se positionne comme l'un des composants substantiels dans toute transformation 'Smart' de l'écosystème urbain.

Ainsi, "Smart Éducation City" est une connotation indispensable de la Smart City. L'éducation intelligente est une nouvelle sublimation de l'éducation à l'ère du digital et une nouvelle étape de son informatisation afin de promouvoir la réforme de l'éducation. Sous l'égide de la 'Smart Université', l'éducation intelligente améliore la qualité et l'efficacité de l'éducation et de l'enseignement et offre le climat propice pour la formation des talents qui vont exercer les métiers d'avenir. L'Université intelligente est un grand projet multi-systèmes, et son architecture globale peut être résumée comme : "apprenant intelligent, enseignement intelligent, pédagogie intelligente, connaissance intelligente et interactions intelligentes". La mission principale de 'Smart Université' c'est de transmettre intelligemment la connaissance aux apprenants, lesquels peuvent collaborer en interne ou de façon étendu quant à l'écosystème universitaire pour acquérir de nouvelles compétences.

Partant de ces faits, l'axe de recherche adopté dans le cadre de notre travail consiste, dans un premier temps, à développer un algorithme, basé sur des données sémantiques qui permettra de modéliser la collaboration interdisciplinaire intégrant une étape de prédiction afin d'optimiser les choix des équipes complémentaires. Puis, dans un second temps, la servicisation dudit algorithme dans une perspective de sa standardisation en tant que service intelligent de l'université intelligente.

Mots-clés: Ville Intelligente, Université Intelligente, Service Intelligent, Collaboration interdisciplinaire, Équipe Complémentaire, Heuristique, Ontologie, Exploration de données éducatives, Classification, Prédiction, Déploiement continu, Cloud Computing

ملخص

يشكل ما يسمى في الوقت الراهن بالمدن " الذكية"، تطويرا وتوجها رئيسيا للسياسات الإنمائية لمعظم البلدان، وذلك من خلال إنجاز مشاريع تجريبية، تزاوج ما بين استخدام جيل جديد من تقنيات المعلومة، ونموذج حكامة مناسبة، وبالتالي فإن هذا النظام الحضري الجديد يعزز "ذكاء" التخطيط، والبناء والقدرة على الإدارة الرقمية للخدمات الحضرية.

ولتحقيق هذا الهدف، لابد من تحفيز جميع الجهات المستفيدة، حتى تتمكن من التعاون لبلورة رؤية واضحة، قصد بناء بيئة حضرية جديدة. وبشكل أدق، فالتعليم يعتبر مؤشرا هاما لقياس التنافسية والتطور الحضري. ومن هذا المنطلق فإنه يعتبر أحد المكونات الأساسية لأي تحول "ذكي" للمجال الحضري.

تشكل "مدينة التعليم الذكي" دلالة مهمة لا غنى عنها بالنسبة للمدينة الذكية، كما أن التعليم الذكي يعتبر قفزة نوعية وجديدة للتعليم في العصر الرقمي، ويشكل كذلك مرحلة جديدة من حوسبته لأجل إصلاح التعليم. في إطار " الجامعة الذكية"، وبفضل التعليم الذكي وجودته ونجاعته الذي يمكن أن يوفر مناخا مناسبا لتكوين الكفاءات التي ستمارس مهن المستقبل. فالجامعة الذكية هي عبارة عن ورش كبير متعدد النظم، يمكن تلخيص بنيته على الشكل التالي: " المتعلم الذكي، التعليم الذكي، البيداغوجية الذكية، المعرفة الذكية، وأخيرا التفاعلات الذكية". وتكمن المهمة الرئيسية "للجامعة الذكية" في نقل المعرفة للمتعلمين بطريقة ذكية، كما أن المتعلمين يمكنهم التعاون داخليا أو بشكل أوسع مع البيئة الجامعية بهدف اكتساب مهارات جديدة.

وانطلاقا مما ذكر، يتمحور البحث الذي اعتمدنا في هذا العمل كمرحلة أولى، على تطوير خوارزمية معتمدة على بيانات دلالية، بحيث تمكننا من تحديد نموذج تعاون متعدد الاختصاصات، مدمجا فيها ذلك مرحلة تنبؤية من أجل تحقيق اختيار أمثل للمجموعات المتكاملة. ثم في مرحلة ثانية، تطبيق هاته الخوارزمية في مختلف المصالح، في أفق تبنيها كخدمة ذكية داخل الجامعة الذكية.

الكلمات الرئيسية: المدينة الذكية، الجامعة الذكية، التعاون متعدد التخصصات، الفريق المتكامل،الاستدلال،الانطولوجيا، البحث في البيانات التعليمية، التصنيف،التنبؤ، النشر المستمر، الحوسبة السحابية.

Liste of publications

Articles

- **AKHRIF Ouidad,** BENFARES Chaymae, El BOUZEKRI EL IDRISSI Younès, HMINA Nabil. Smart collaborative learning: A recommended building team approach. International Journal of Smart Security Technologies (IJSST), 2019, vol. 6, no 2, p. 52-66.
- **AKHRIF, Ouidad,** BENFARESS Chaymae, EL JAI Mostapha, El BOUZEKRI EL IDRISSI Younès, HMINA Nabil . Completeness based classification algorithm: a novel approach for educational semantic data completeness assessment. Interactive Technology and Smart Education, 2021.
- **AKHRIF Ouidad,** BENFARES Chaymae, El BOUZEKRI EL IDRISSI Younès, HMINA Nabil. An architecture for continuous deployment of the Smart Collaborative Learning Service (SCLS) based on a predictive model to build complementary teams. International Journal of Pervasive Computing and Communications. Submitted for review (pending) 2022.
- BENFARES Chaymae, **AKHRIF Ouidad,** El BOUZEKRI EL IDRISSI Younès, Hamid Karim. "Diagnosis and Prediction of Depression using Machine Learning: A Systematic Review”. Artificial Intelligence In Medicine. Submitted for review (pending) 2021.
- BENFARES Chaymae, **AKHRIF Ouidad,** El BOUZEKRI EL IDRISSI Younès, Hamid Karim. A clinical support system for classification and prediction of depression using machine learning methods. Computational Intelligence, 2021, vol. 37, no 4, p. 1619-1632.
- BENFARES Chaymae, **AKHRIF Ouidad,** El BOUZEKRI EL IDRISSI Younès, Hamid Karim. Multi-Criteria Decision Making Semantic for Mental Healthcare. International Journal of Smart Security Technologies (IJSST), 7(1), 58-71.

Book Chapters

- **AKHRIF Ouidad,** El BOUZEKRI EL IDRISSI Younès, HMINA Nabil. Service Oriented Computing and Smart University. In : The Proceedings of the Third International Conference on Smart City Applications. Springer, Cham, 2018. p. 437-449.
- **AKHRIF Ouidad,** BENFARES Chaymae, El BOUZEKRI EL IDRISSI Younès, HMINA Nabil. Smart University services for collaborative learning. In : The Proceedings of the Third International Conference on Smart City Applications. Springer, Cham, 2019. p. 131-142

International Conferences and Workshops

- **AKHRIF Ouidad,** El BOUZEKRI EL IDRISSI Younès, HMINA Nabil (October 10-11) - Smart university: SOC-based study- The 3rd International Conference on Smart City Applications, SCA2018- Tetouan-Morocco.

- **AKHRIF Ouidad,** El BOUZEKRI EL IDRISSI Younès, HMINA Nabil (October 02-03-04)- Collaborative learning services in the Smart University (SU) environment - The 4th International Conference on Smart City Applications, SCA2019- Casablanca-Morocco.

- **AKHRIF Ouidad,** El BOUZEKRI EL IDRISSI Younès, HMINA Nabil (2019, October). Enabling Smart Collaboration with Smart University (SU) Services. In Proceedings of the 3rd International Conference on Computer Science and Application Engineering (pp. 1-6)- Sanya- China.

- **AKHRIF Ouidad,** BENFARES Chaymae, El BOUZEKRI EL IDRISSI Younès, HMINA Nabil (October 22-23) -Collaborative Approaches in Smart Learning Environment: A case study -The Second International Workshop on Emerging Networks and Communications (IWENC2020) August 9-12, 2020, Leuven, Belgium.

General introduction

1. Overview

Smart and innovative abilities have recently defined a city that uses information and communication technologies (ICT) to deploy sustainable services that meet the needs of a gradually growing population. Improving the quality of life of its citizens, anticipating their needs in terms of education, transport and health, solving energy and environmental problems are the main challenges when overhauling today's cities. We believe that the Smart City (SC) must have sufficient resources, be intelligent, recyclable, flexible, collaborative and be able to predict the scarcity of its resources in order to protect them in the future. The Smart City (SC) project extracts knowledge from the huge volume of data and information, gathered by an Internet of Things (IoT) infrastructure to design the Smart City Services (SCS) layers. Therefore, Building smart cities require extensive technical support, and Big Data is the cornerstone of knowledge analysis which is considered the driving force behind the development of the so-called “knowledge city”.The specific objectives of the construction of Smart Cities (SC) include accessibility of public services, sophisticated urban management, a livable living environment, smart infrastructure and long-term network security. Education is one of the important public services in the city, and smart education is also the basic pilar of a smart city (SC). Many intelligent education systems will be built. So, how to assess the intelligence of the city's educational services that meet the requirements of the Smart City building?

Traditional education is the essential learning method holding from the industrial revolution to the present day. It is a practice adopted in most colleges, research institutes and universities. Wherein, The teaching model is based on the idea that students should be passive receivers of information. This not only removes creativity but also eliminates the role of differences between students and diversity which is the most valuable asset of human society. Fortunately, The new generation of education and learning methods are evolving. Entering the 21st century, many changes have been added to the development of education due to technological advances. The predominance of the Internet and the audiovisual industry has made the dissemination of information even more extensive and traditional education has gradually changed toward online education. Online learning is a generic term that refers to technology-supported learning, rather than similar terms such as electronic learning (e-learning), web-based learning, distributed learning and technology-mediated learning (McGill and Klobas 2009). Online learning is no longer limited to this, smart technologies continue to drive the optimization and upgrading of the education industry to become a learner-centered smart education.

Smart education is providing smart knowledge to smart learners, with the guidance of smart teachers and the support of smart parents, through smart interactions, to create a smart learning environment, using smart pedagogy methods to formulate a Smart University (SU) framework. It is a readjustment of education for student-centered learning. Relying on its main component, the Smart University (SU) brings new opportunities for Smart learners for well-being and well known in the academic environment. Key among these opportunities is the implementation of learner-centered teaching, students can acquire knowledge and solve problems through active exploration and self-construction of teaching methods in learning activities. As a student-centered teaching strategy, problem-solving learning refers to grouping smart learners around a project according to their interests, which are presented as a description of the project. These aspects motivate students to work together toward a common goal, generating positive interdependence within the team and creating individual responsibilities for each student to benefit the group’s progress (Ramírez-Donoso, Pérez-Sanagustín, and Neyem 2018). It is a great opportunity to ensure true collaboration in the learning process. Collaboration is a powerful pedagogy method for learners to harness new ideas and synthesize information if it is properly managed. Simply put, collaborative learning is affected by the quality of interactions and the assignment of learners to successful teams building. Therefore, envisioning permanent communication channels personalized according to the skills and background of the students, creating complementary student groups, and designing open and interdisciplinary collaboration are the main challenges of smart collaborative learning.

2. Research interests

Smart University (SU) aims to democratize learning and provide a level playing field for all students, regardless of their accessibility, background, or communication skills. The learning process is based on teaching and receiving knowledge while ensuring full integration and communication between learners in this setting. Catalyzing learner engagement and communication necessitates a collaborative strategy that will allow students to expand their learning opportunities while also maintaining the successful relationships between teammates that are required to meet their training goals. While implementing collaborative learning, it is critical to handle the organizational aspect of the team's learners as well as the content management. In this research, We have identified the primary issues encountered in order to build a successful interdisciplinary collaboration.

Issue 1: Building a collaborative team fosters cooperation and enthusiasm while assisting students in further working in a complementary manner, allowing them to capitalize on their knowledge and experience. It's an educational method for getting students to work together in heterogeneous groups. As a result, An algorithm that govern the assignment of learners to complementary teams while maintaining full participation and integration of students in the learning process is required for the team building.

Issue 2: Smart institutions must maintain an emphasis on communication and coordination, as well as foster a culture of knowledge transfer among students based on cooperation, diversification, and tight cross-collaboration. Students working in interdisciplinary projects can be inspired by the atmosphere of various disciplines and develops the ability to innovate and overcome problems. Interdisciplinary collaboration must be engaged at this point to broaden the breadth of information retrieval. In such a heterogeneous field, finding the right collaborator as well as the specialties within a partnership's scope is challenging. As a result, multidisciplinary collaboration demands an appropriate conceptual model for modeling student profiles and disciplines in order to perform accurate and meaningful research.

Issue 3: Through communication and mutual aid between learners, cross - disciplinary collaboration in education and research is a very successful way of sharing knowledge and bringing out new ones. Universities have indeed been pushed to collaborate with possible partners in ways that lead to successful collaborative project completion as a result of the combination of collaborative disciplines and education. Universities must have a clear vision of the effectiveness and success of collaborative projects before beginning any collaboration. A bad working environment lowers students' performance, which is a significant disadvantage for students who are constantly striving to learn new abilities. In order to help overcome this issue, universities must have decision support that allows them to analyze the success of allocating students to complementary teams while assuring their variety in terms of abilities and homogeneity in terms of collaboration skills.

Issue 4: Students' participation in collaborative projects is permanently generating information that represents traces of their collaborative behaviors. This information plays an important role in the precision of future collaborative suggestions and the prediction of effective student assignments to complementary teams. As part of its strategy, the Smart University (SU) must synchronize its decision-making support with changes in data reflecting learners to improve the quality of its decision-making level.

3. The Thesis Objectives

In the previous section, we discussed various challenges related to building complementary teams. These issues motivate the objectives of this thesis. We listed them as follows:

- A comprehensive review of the literature is conducted to identify the main components of the Smart University, as well as opportunities and challenges, in order to better comprehend and assess intelligent learning environments. This approach led us to the creation of a taxonomy that will aid in the development of the Smart University's ontology;
- We generated an ontology of domains reflecting the main components of the Smart University, as well as these semantic relationships, based on the acquired taxonomy. This ontology aids in recognizing the Smart University (SU) field in order to perform meaningful processing;

- Developing a heuristic allows learners to be assigned to complimentary teams. Using statistical procedures such as sorting, search by criterion, and Boolean algebra blocks on data from the Smart University (SU) ontology;
- Using this heuristic in a decision support system to smartly create complementary teams, provides for more accurate outcomes in forecasting the performance of each combination of these teams. Accordingly, we developed a predictive model based on the complementarity heuristic's classes for classifying and predicting the students' assignment to complementary teams.
- We covered how to deploy the Smart Collaborative Learning Service (SCLS) in a cloud architecture (Amazon Web Service) to provide continuous and on-demand building complementary team service through a REST API interface. As a result of our experimentation, we designed a prototype architecture that allows for continuous deployment of a predictive model in a cloud environment using streaming and integrated deployment architectures.

4. Thesis Layout

The thesis is organized as follows:

Chapter I : The smart city ecosystem. The smart city is a new approach of managing urban environments in order to accommodate rapid population growth, which requires a lot of resources and needs. A smart city's management is based on the ability to connect various urban systems, collect, analyze, and interpret data extracted from them, and provide services adapted to the needs of citizens. Many projects have supported the creation of smart cities to address issues arising from each of the city's restrictions. For this, we found a variety of definitions of smart cities in the literature that are closely relevant to the problem encountered; furthermore, we were able to extract the essential definitions and explain the ideas related to the definition of smart cities. The smart city, as a system of systems, is viewed as an ecosystem characterized by the synergy of these several pillars. The deployment of this ecosystem calls for the development of an architecture that comprises the key layers of sensors, data, and services. The Smart University (SU) subsystem exists at the core of the smart city ecosystem as an educational institution structured under a smart system architecture, aimed to generate the right learning environment and train learners who will have future professions. In this approach, a new ecosystem of learner-centered education may be built.

Chapter II : Universities and Smart City: new concepts,opportunities and challenges. The performed literature review examined the concept "Smart University" and its various opportunities to obtain better outcomes. In this chapter, we emphazid the main components of the Smart University (SU) that support the adoption of the smart collaborative learning environment . In reality, collaborative learning is critical for effectively acquiring and sharing knowledge through intelligent interactions among learners. This concept requires a smart service architecture that ensures a functional aspect while also touching on additional treatments, as well as an adapted deployment, a perfect composition, and more relevant selection based on several contextual parameters such as user profile, interaction history, and preferences.

Chapter III : Ontologies in Educational Data Mining. In this chapter, we present an overview of Educational Data Mining (EDM) methodologies and methods in education . We've also talked of how to incorporate semantic educational data into the EDM process for meaningful processing.

Chapter IV : Completeness based classification algorithm: The purpose of this paper is to reveal the Smart Collaborative Learning Service (SCLS). This approach intends to build teams of learners based on the complementarity of their skills, allowing for flexible participation, and providing opportunities for interdisciplinary cooperation for all students. The success of this environment is determined by the ability to predict efficient collaboration the different teammates, allowing for smart knowledge exchange in the Smart University (SU) environment. A Random Forest technique has been proposed, which is based on a semantic modelization of the learner and the problem-solving and a heuristic for building complementary teams (HBCT). To provide it, we created the KNIME (Konstanz Information Miner) workflow, which combines the main steps for building and evaluating the Random Forest classifier. The workflow is divided into three parts: extracting knowledge from the Smart Collaborative Learning ontology, calculating completeness using a novel heuristic, and building the Random

Forest classifier. Using a semantic decision support system, the Smart Collaborative Learning Service (SCLS) facilitates efficient collaboration and democratized knowledge sharing between learners. This service solves a frequent issue related to the composition of learning groups in order to serve pedagogical perspectives.

Chapter V : An architecture for continuous deployment of the Smart Collaborative Learning Service (SCLS) based on a predictive model to build complementary teams. In this chapter, We presented a novel cloud architecture for continuously deploying the Smart Collaborative Learning Service (SCLS) that contains a predictive algorithm for building complementary learning teams. Invocating this smart service in a collaborative learning platform creates a flexible building team of learners, based on the completeness of learner skills required by a collaborative project. This intelligent service satisfies the intelligence requirements that we specified in our approach to evaluate its ability to continuously adapt to environmental changes and data updates , as well as efficiently predict complementary teams. We successfully deployed the Smart Collaborative Learning Service (SCLS) and used its REST API to make it powerful and on-demand sharing. The suggested architecture incorporates intelligent layers such as semantic data representation, heuristic preprocessing, predictive model creation, and stream processing updates.

Finally, we conclude the manuscript with a summary of the research's contributions and future prospects.

Chapter I : The smart city ecosystem

1. Introduction

Today's cities' urbanization is correlated with the establishment of sustainable development and ecological urbanism concerning their economic, societal and environmental activities. The increasing concentration of people in cities prompted them to implement a strategy aimed at reducing expenses, improving organization, and insuring the well-being of citizens, in order ensure a viable, livable, and equitable environment for an urban population that is constantly growing. Currently, 50% of the world's population lives in cities. By the year 2050 that percentage will rise to 70% (UN 2008), by that time, many countries will face challenges to meet the needs of their growing urban populations, including housing, transportation, energy systems and other infrastructures. Any strategy referring to the urbanization of emerging cities is closely linked to the logic of each of these cities, which is determined by their culture, geography, economic, and industrial activity. These factors identify the impending issues that will be encountered throughout popular expansion, which might include pollution, energy consumption, crime, transportation, and housing. To do this, designing a city model that respects the urbanization rules of emerging countries remains restricted. However, the technological revolution allows the emergence of new systems resistant to changes in urban territories and the demands of citizens' needs. Indeed, the integration of information technologies in traditional infrastructures improves the performance of the services provided to citizens, optimizes the use of existing infrastructure and consolidates collaboration between the various actors of the city to create strong synergy and cohabitation between the activities of the urban environment. All these citizen-oriented assets are materialized by building a foundation for an innovative and sustainable city per the “SMART CITY”.

"SMART CITY," an Anglo-Saxon concept, is not new. It has gotten a lot of attention since 2009, when IBM launched the Smarter Planet corporate project, which has gained wide support from governments, businesses, universities, and other connected communities all across the world (C. Harrison et al. 2010). The term "SMART" refers to transformational and innovative developments driven about by new technologies. It brings up notions of data-driven decisions and technology-enabled data-sharing, plus communications and collaboration, all leading to continuous improvement (Kitchin 2016). A "smart city" is often defined as a city that is able of responding to the changing or growing demands of institutions, enterprises, and citizens on several levels, including economically, socially, and ecologically. It helps to the resolution of development issues by enhancing citizens' lifestyles and ensures effective functioning of their infrastructure. According to (Albino, Berardi, and Dangelico 2015),The concept of the smart city is far from being limited to the application of technologies to cities. In fact, the term is being used in a variety of contexts with no agreed-upon meanings. This has caused consternation among urban leaders who want to implement measures that would make their cities "SMART".

This thesis tries to explain the cutting edge knowledge about a “SMART CITY”, what are its key dimensions and how its performance can be assessed. It is based on a review of the literature, including relevant researchers in this field. This paper is organized as follows: The next section presents the Smart cities at the digital transformation. Then, it explores the smart city characteristics. This is followed by a talk about Smart city dimensions followed by the Smart city architecture patterns and the last one is concerned with Collaborative impacts on Smart Cities.

2. Smart Cities in the digital transformation era

The term “SMART CITY” is often associated with a city qualified to achieve a digital transformation at its infrastructure, business and activities to improve the quality of its public services. Of course, the integration of new information and communication technologies in urban areas has greatly accelerated its development and modernization process, but this digital transformation is only a catalyst for the development of smart cities. For that, we aim to clarify the meaning of the word “SMART” in the context of cities through a bibliographic study that focuses on the main definitions and dimensions of the smart city:

This term “smart city” is generally attached to any urban phenomenon based on a cybernetic effect where the action is corrected by the feedback of the effect on the cause, generating a cumulative learning effect. With digital convergence, there is an amplification of these phenomena, which allows new applications (Rochet 2014). It is about a high-tech intensive and advanced city that connects people, information and city elements using new technologies to create a sustainable, greener city, competitive and innovative commerce, and an increased life quality (Bakıcı, Almirall, and Wareham 2013). Being a smart city means orchestrating and optimizing the intervention of all the technologies and resources available in a city to develop integrated, livable and sustainable urban centers. Thus, the smarter city connects the physical infrastructure, the IT infrastructure, the social infrastructure, and the business infrastructure to leverage the collective intelligence of the city (C. Harrison et al. 2010).On the physical infrastructure side, the smart city is based on its smart digital infrastructure, which is characterized by the integration of new interconnection media such as the Internet of Things, allowing an optimal flow of information from machine to the machine which eliminates human intervention. It uses IoT to listen and understand what is happening in the city, which enables them to make better decisions and to provide information and services adapted to their citizens (Trilles, Calia, Belmonte, Torres-Sospedra, et al. 2017).

An indicative Smart City definition comes from ISO / IEC 8 and recognizes the smart and sustainable city as “an innovative city that uses ICT and other means to improve quality of life, the efficiency of urban operation and services, and competitiveness while ensuring that it meets the needs of present and future generations for economic, social, and environmental aspects. ”

The smart city social infrastructure is based on intelligent exchanges of information that circulate between its many different subsystems. This flow of information is analyzed and translated into citizen-centered and commercial services. The city acts on this flow of information to make its wider ecosystem more efficient and more sustainable. Information exchange is based on an operational governance framework designed to make cities sustainable (Vinod Kumar and Dahiya 2017). In the urban planning field, the smart city, as a new concept of urban development, aims to improve the quality of life of city dwellers by making the city more adaptive and efficient, using new technologies that are based on an ecosystem of objects and services. The perimeter covering this new mode of city management includes in particular:

public infrastructures (buildings, urban furniture, home automation, etc.), networks (water, electricity, gas, telecommunications); transport (public transport, smart roads and cars, carpooling, so-called soft mobility - by bike, on foot, etc.); e-services and e-administrations. Moreover, a smart city should not be considered a technological artifact. It is an ecological and cultural system. The ecological system defines the environmental issues and can be diverse at every location within the city and may differ from other cities. A cultural system that interacts with the environment arises out of the way of life which is influenced by religious practices (Vinod Kumar and Dahiya 2017).

In this regard, a smart city is understood as a certain intellectual ability that addresses several innovative socio-technical and socio-economic aspects of growth. These aspects lead to smart city conceptions as “green” referring to urban infrastructure for environmental protection and reduction of CO2 emission, “interconnected” related to the revolution of broadband economy, “intelligent” declaring the capacity to produce added value information from the processing of city's real-time data from sensors and activators, whereas the terms “innovating”, “knowledge” cities interchangeably refer to the city's ability to raise innovation based on knowledgeable and creative human capital (Zygiaris 2013). Also, "Smart city" means an instrumented, interconnected and intelligent city (C. Harrison et al. 2010). who can respond to the changing or emerging needs of institutions, businesses and citizens, both economically, socially and environmentally? It helps solve development problems by improving the lives of citizens and the optimal functioning of their infrastructures (Ouidad Akhrif and El n.d.).

The definition of the term “SMART CITY” is accomplished by disclosing the main keywords used to better identify the different metrics of urban intelligence, these concepts represent the main characteristics and dimensions of a “SMART CITY”.

3. Smart city characteristics

The "SMART CITY," with its various identities, is a difficult idea to describe. As a result, we recommend citing the primary terms that are usually associated with the description of the Smart City, which are technological or business concepts associated with the term "intelligent" in the urban context. Thanks to a study of the literature defining the concepts related to the smart city, we established:

Sustainable city: A city that uses information and communication technology and other tools, such as cyber-physical systems, to improve people's quality of life and the efficiency of urban public services. It ensures the city's competitiveness for the participation of coming generations on economic, social, and environmental plans (Tokody and Mezei 2017).

Instrumented city: Refers to the capacity to capture and integrate live real-world data via sensors, meters, appliances, personal gadgets, and other similar sensors (Albino, Berardi, and Dangelico 2015). Sensors of all kinds will need to be distributed throughout the instrumented city's infrastructure in order to provide continuous diagnostics on buried pipes, the status of the transportation network, street pollution, automotive traffic and public lighting. Two more critical issues must be addressed: the supply of energy to the sensors and the transmission of the data collected.

Technological city: Smart cities are technological solutions designed not only to absorb growing population pressure, but also to deliver better and more efficient services and processes, as well as to promote long-term economic growth ,as a corollary, to give citizens a better quality of life. The use of Smart Computing technology makes a city's critical infrastructure components and services more intelligent, interconnected, and efficient, including municipal administration, education, healthcare, public safety, real estate, transportation (Washburn and Sindhu 2009).

Interconnected city: This means integrating data into a computing platform that allows such information to be communicated among the many city services (Albino, Berardi, and Dangelico 2015). This level is divided into two parts. The physical phase starts with the network, which combines the systems and links the devices. Second, the logical phase is characterized by the integration of data from diverse sources and the creation of meaningful relationships between that data (Alshawish, Alfagih, and Musbah 2016).

Collaborative city: The main logic of a participatory and collaborative city is that it is the people's fundamental right to participate in establishing the direction and aims of topics that involve them. As a result, city inhabitants may participate in the design process alongside experts (Tokody and Mezei 2017). In this context, IBM's role is to support smart city projects by collaborating closely with industrial partners and service providers, the notion of the collaborative city, that of a community that involves all of its public and private actors in the search for more effective urban solutions in management, advertising, sustainable development, or services to residents (T. M. Harrison et al. 2012; Vinod Kumar and Dahiya 2017).

Digital city: The concept of digital cities creates a space for people from different groups to communicate and exchange their knowledge, experiences, and shared interests. It refers to the use of ICT to assist the establishment of a wired, ubiquitous, linked network of citizens and organizations, sharing data and information, and participating in online services, which is aided by public policies such as e-government and e-democracy (Ishida 2002).

Ubiquitous city: An extension of the digital city concept in terms of ubiquitous accessibility and infrastructure (C. Harrison et al. 2010). In terms of great accessibility, a "Ubiquitous City" is a digital city. It makes ubiquitous computing available to all urban aspects (Ishida 2002). Its defining characteristic is the construction of an ecosystem in which any citizen may obtain any service at any time and from any device.

Knowledge city: It concerns policies aimed at ensuring that the data, information and knowledge available and produced in cities are respected and valued, in particular through their cultural institutions, but also produced and used by businesses, innovative neighborhoods, technological parks (C. Harrison et al. 2010). Developing knowledge-based city extraordinary abilities to be self-aware, how it works 24/7 and selectively communicate in real-time knowledge to citizen end-users for satisfying lifestyle with easy public service delivery, comfortable mobility, conserve energy, the environment and other natural resources, and create vibrant face-to-face communities and a vibrant urban economy even at a time when there are national economic downturns (Vinod Kumar and Dahiya 2017).

Green city: It regards an ecological vision of urban space that is founded on the notion of sustainable development. Green policies in the city focus on decreasing the city's environmental impact, lowering pollution, waste, and energy consumption, and conserving or constructing public green spaces such as parks and gardens (Benevolo, Dameri, and Auria 2016).

Smart city: Uses advanced analytic techniques to provide insight for city events and visualization tools to visualize the city's behavior (Alshawish, Alfagih, and Musbah 2016). It refers to the inclusion of complex analysis, modeling, optimization and visualization services to make better operational decisions. To give more detail, we conferred a clear smart city clarification in the Definitions section.

In this light, Every city has issues relating to its urban ecology, which primarily affects the economy, society, and environment. Thus, the notion of the smart city is defined by each city's strategy to ensure a favorable environment for development. Thus, the concept of the smart city is defined by the policy of each city to ensure a favorable context for development. We used Google trends to analyze the change in interest in various terms related to the smart cities definitions as shown in Figure 1:

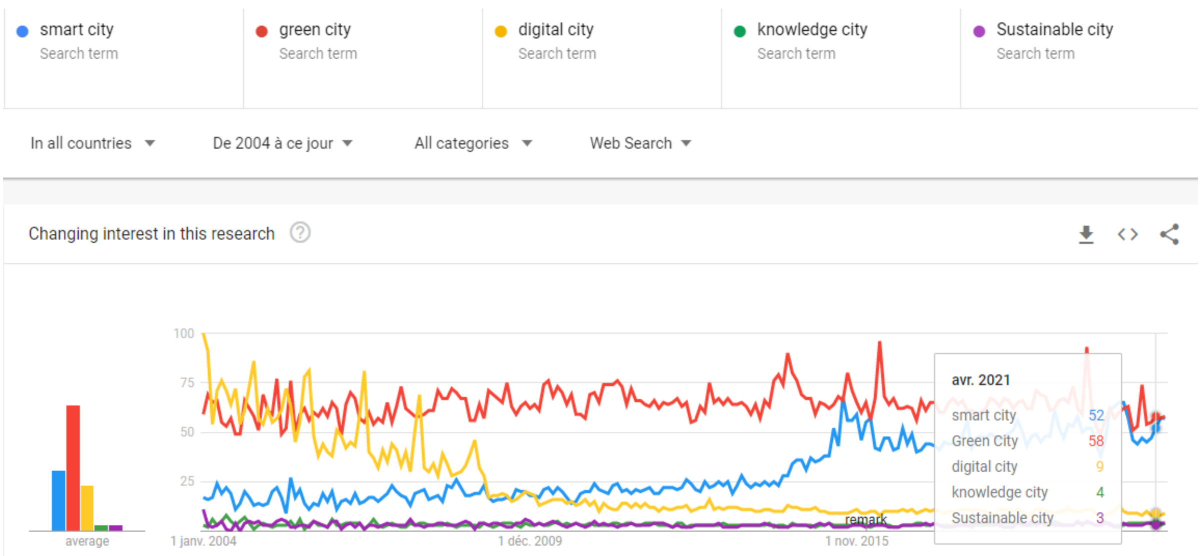


Figure 1. Smart city characteristics

These statistics show that every city in the world is implementing the smart city idea in order to handle its challenges, but we can presume that the most important area right now is addressing environmental issues.

4. Smart city dimensions

The Smart City is a fuzzy concept (Nam and Pardo 2011), it has several elements, components, and measurements . A smart city must satisfy six requirements, according to Rudolf Giffinger, an expert in analytical research on urban and regional development at the Technical University of Vienna., as shown in Figure 2:

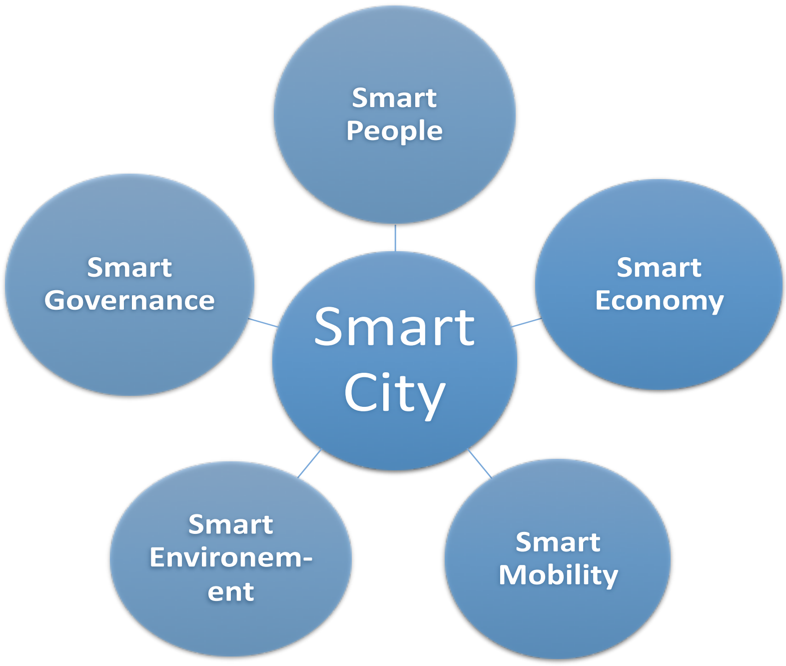


Figure 2. Smart city dimensions

Smart economy: A smart city's economy is characterised by its capacity to overcome challenges, generate new jobs, start new businesses, and boost attractiveness and competitiveness (Alawadhi et al. 2012). The smart economy is defined by the use of ICT in all aspects of economic activity.

Smart mobility: Smart Mobility is only one of the topics covered in the Smart City implementation (Nam and Pardo 2011). It is, therefore, a critical subject, affecting various dimensions of the smart city, several aspects of people' quality of life, and all the potential stakeholders expecting benefits from the smart city implementation (Arena et al. 2013).

Smart environment: The smart environment is a knowledge-based environment that fosters great skills to be aware of its 24/7 day operations and selectively send its information to end-users in real-time for a fulfilling living, with efficient public service delivery, convenient mobility, conservation of energy, the environment, and other natural resources (Vinod Kumar and Dahiya 2017). Smart Environment, as a smart city system, may also be conceived of as a six-component system, such as for the water element of the environment, Smart Water Community, Smart Water Mobility, Smart Water Economy, Smart Water Governance, Smart Water Environment, and Smart Water entered Living (Vinod Kumar and Dahiya 2017).

Smart People: is the primary building component of a Smart City System, and it requires a number of critical characteristics, which are listed below (Vinod Kumar and Dahiya 2017):

- Intelligent people excel at what they do professionally;
- A smart city incorporates universities and colleges into all aspects of daily life;
- It promotes high-level human resources, such as knowledge workers;
- A smart city has a high graduate enrollment ratio and people that choose lifelong learning and use e-learning methods;
- Residents of smart cities excel at being creative and coming up with novel solutions to complicated issues;
- Smart people are cosmopolitan, open-minded, and have a multicultural outlook;
- Intelligent citizens are actively involved in their city's long-term growth, effective and seamless operation, maintenance and administration, and making it more livable.

Smart Governance: To attain Smart City, cities are needed to have effective governance, which in this instance demands good collaboration between the government (as the authority) and the community. Transparency in government operations, openness, community support for continuous governance, and active engagement from the community and government are the primary components to achieving Smart Governance (Nasution, Erwin, and Risanty 2020).

Smart Living: Smart Cities are intended to achieve the process of a better living (based on technological information), which includes the community's quality of life and culture that has been operated on society. To achieve these features, it is possible to provide support infrastructure (electricity, internet, roadway), deal with societal problems (social, health, environment), and preserve cultural heritage by employing technological information. (Nasution, Erwin, and Risanty 2020). Smart Living aims to enhance living circumstances, particularly access to culture and education, while also promoting social cohesion, health, and safety.

5. Smart city architecture patterns

Smart cities are becoming very sophisticated ecosystems integrating innovative solutions and smart services. These sub-ecosystems make Smart cities engines of collecting, producing and sharing pertinent data, which sets new challenges for building effective Smart City architectures and new services. To design the infrastructures and systems of Smart Cities well, we need to design them in context - that is, with an understanding of the environment in which they will exist, and the other elements of that environment with which they will interact. In other words, the layers of this architecture depend on the problems to be solved and the constraints encountered during the implementation of this architecture and that requires a deep understanding of the urban environment and the other elements with which it interacts. As a result, a large number of smart city architecture can be found in the literature focusing on different aspects, such as technology, human-system interaction, and logic. Here we present a list of prior architectures to help better understand the fundamental layers of a smart city.

Table 1. Prior studies on smart city's architectures

Articles	Research Area	Architectures	Technologies
(Anthony Jnr 2020)	Energy sustainability	Context layer Service layer Business layer Big data Physical infrastructure layer	Smart grid, MongoDB Not Only SQL Database (NoSQL) and Hadoop Distributed File System

			(HDFS),Energy sensors, Metering devices.
(Anthopoulos, Janssen, and Weerakkody 2016)	Generic multi-tier ICT architecture for smart city	Networking Infrastructure and Communications Protocol Applications Business Management Services	Datacenters, supercomputers and servers, networks, IoT, Sensors.
(Cha et al. 2021)	Privacy data transaction	Device layer Inter-Network layer Fog layer Cloud layer	Blockchain Internet of thing Cloud computing
(X. Chen et al. 2017)	Energy management Reduce traffic congestion	Data creation and collection layer Data processing and management layer Event and decision management layer Smart service layer	Big Data analytics RESTful API Internet of thing Web od thing RESTful Smart Gateway
(Schleicher et al. 2016)	Buildings, hard infrastructure (traffic lights, bridges, and roads), energy grids, or public transportation	Infrastructure layer Applications layer Citizen participation and engagement layer(Smart service)	Internet of thing Internet of thing cloud application API Web-based front ends
(Gavrilović and Mishra 2021)	Smart City, healthcare, and agriculture	Sensors and actuators Internet gateways and data acquisition systems. Edge Information Technology. Datacenter and cloud. Service layer	Internet of thing Cloud computing Big data

(Gaur et al. 2015)	Smart health, smart environment, smart energy, smart security, smart office and residential buildings, smart administration, smart transport and smart industries	Data collection Data processing Data integration and reasoning Device control and alerts Communication Services Customized Services	Internet of thing Semantic web Cloud computing
(Costa and Santos 2016)	Smart transportation	High volume and variety data Data extraction and loading mechanisms Big data storage Data output mechanisms Big data analytics Data services and applications	Internet of thing Big data/Open data Cloud computing Smart service
(Cheng et al. 2015)	Big data platform for SmartSantander	Data Collection Data Repository Data Processing CityModel API-Services	Internet of thing Big data REST API Cloud computing
(Chamoso et al. 2020)	Smart Home	Citizen and administration applications Information management/storage Developed modules providing the main functionalities.	Internet of thing Big data Cloud computing Kafka Connect and Apache Arrow
(Khan and Kiani 2012)	Urban governance	Data sources Data Acquisition and Analysis Layer Thematic Layer Service Composition Layer Application Service Layer	Internet of thing Artificiel intelligence Cloud computing Service-oriented architecture

Typically, smart cities are urban environments that exploit heterogeneous data received from sensors to deliver smart services for the benefit of users. In essence, this blueprint consists of three main parts.

5.1. Sensors "Internet of Things"

Sensors are objects capable of receiving and transmitting data via standardized wireless electronic identification systems. There are a large number of them in the field of health, industry, leisure, or home automation, but the best known remains the Smartphone. These are entities (Citizen, RFID, GPS, IR, camera, laser scanners, etc.) that are interconnected under an "Internet of Things" (IoT) infrastructure; for example, environmental sensors can collect data about the environment, such as noise and air quality (Khan et al. 2015). While many connected objects, including smartphones, are useful and allow the development of the smart city, others are specifically designed and developed for the smart city. These objects, street lamps, trash cans, pollution sensors, roads, car parks make it possible to facilitate access to information for users and communities, to detect street places, optimize waste collection and the distribution of energy, simplify mobility and bring well-being and safety to users (Trilles, Calia, Belmonte, and Torres-sospedra 2017).The application of IoT requires the following layers: Perception (made up of sensors, embedded communication hardware and control unit), Network (sending storage requests, setting up data analytics tools), Presentation (tools for visualization and interpretation that can be applied to different platforms and used in different applications) (Bhabad and Bagade 2015; Gavrilović and Mishra 2021):

- The Perception layer includes various devices capable of discovering and monitoring things and communicating information over the Internet, as well as the information in its inventory and how things are organized (Bhabad and Bagade 2015). Radio Frequency Identification Devices (RFID), cameras, sensors, Global Positioning Systems (GPS) are some examples of perception layer devices (Shao 2012).
- At the Network layer, the data gathered by sensors used to be sent to the internet with the help of computers, wireless/ wired network and other components. Hence network

layer is mainly responsible for the transmission of information with the feature of reliable delivery this layer also includes the functionality of the transport layer (Bhabad and Bagade 2015).

- The latest level of the Presentation layer is where analyzing the received information and making the control decisions to achieve its feature of intelligent processing by connection, identification and control between objects and devices. Intelligence means making use of intelligent computing technology such as cloud computing and processing the information for intelligent control like what to do and when to do things hence this layer is also called as process layer (Shen et al. 2007).

5.2. Data “big data”

A Smart City represents a rich environment with multiple potential data sources, such as urban resources, buildings, mobility, energy, among others. These data sources can generate two types of data: data with a low degree of velocity and concurrency, incorporating files (CSV, TXT, JSON, XML…) and scheduled readings (periodic readings from databases or historical web data, such as news feeds, for example); data with a high degree of velocity and concurrency, representing data streams from the web (tweets, blog posts…) or electronic devices (smart meters and other sensors, smartphones, meteorological stations, geolocation capable devices…) (Costa and Santos 2016). Due to the diversity of types and sources of data, the Smart City has to implement a Big data strategy to ensure its permanent intelligence. Collecting and storing a mass of data is only of interest if we can make it efficient, operational and valued. Adopting a big data strategy is an asset that not only helps to better understand the functioning of the city and the behavior of its citizens but also to decompartmentalize the various actors and operators and to create new services responding to new uses thanks to a reservoir of existing data (Dobre and Xhafa 2014). Big data systems are stored, processed, and mined in smart cities efficiently to produce information to enhance different smart city services. In addition, big data can help decision-makers plan for any expansion in smart city services, resources, or areas. The various characteristics of big data demonstrate its considerable potential for gains and advancements (Abaker et al. 2016).To know the state of a smart city, it is not only necessary to monitor and measure everything at all times, and also to be able to quickly interpret the data collected, to act in the best way. For a city to be smart, this data is likely to move from data to information and then to knowledge before reaching the supreme level which is wisdom. This theory is commonly referred to as "DIKW" (Data, Information, Knowledge and Wisdom). The pyramid above represents the four phases of transforming data to Wisdom through Information and Knowledge (Alshawish, Alfagih, and Musbah 2016).

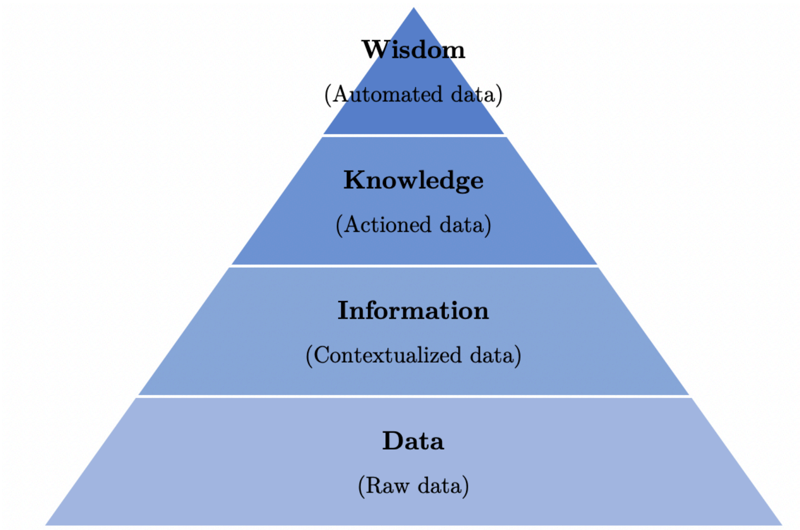


Figure 3. Data, information, knowledge and wisdom pyramid

The layers of the smart city perform the following operations: performs watch, listen, learn, connect, predict and correct On its Big Data to unlock the pyramid stages. Hence, intelligence is obtained (Alshawish, Alfagih, and Musbah 2016).

5.3. Smart City based Services

Affording a better quality of life to its citizens, anticipating their needs in terms of education, transport and health, solving energy and environmental problems and ensuring sustainable services: are the main challenges of the Smart City. To achieve this, the Smart City (SC) extracts knowledge from the enormous volume of data and information, collected by the Internet of Things (IoT) infrastructure to design a Smart City Services (SCS), to promote its ability to learn, respond and recommend relevant services for these citizens. Therefore, the smart city (CS) will have to develop efficient services for the benefit of its citizens in order to meet their needs. This requires more detailed management of these different services, coordination between them, communication, integration and greater involvement of citizens who will also contribute to enriching the services, their applications and their functionalities.

"Service-Oriented Computing (SOC) is the computing paradigm that utilizes services as fundamental elements for developing applications/solutions" (Papazoglou et al. 2007) which are fast, inexpensive, interoperable, scalable and massively distributed. The relevance of Service-Oriented Computing (SOC) is reflected in the flexible integration of contextual data collection, anticipation, prediction and recommendation services into a city system to make it smart and proactive.

5.3.1. Citizen-centric Smart Service

Smart City Services is the application layer covering the innovative functionalities provided by a Smart City to its citizens. Through this layer, the citizen interacts as a consumer, by exploiting the various Smart City services that improve their quality of life, or as a partner that contributes to the production of intelligence in the city. Designing smart services is an asset that allows citizens to discover their environment and have a participatory and collaborative life. Its role is to create an environment where any citizen can get any service anywhere and anytime via any device.

5.3.2. Smart City-centric Smart Service

The Smart City (SC) is a collection of Smart Computing Technologies (SCT) applied to the critical infrastructure components and services, which include city administration, education, healthcare, public safety, real estate, transportation, and utilities to make it more intelligent, interconnected, and efficient (Washburn and Sindhu 2009). The combination of these smart services brings along Smart City applications for active and autonomous adaptation by using the benefits of contextual information.

5.3.3. Smart service properties

In a Smart City System (SCS), Smart Services are characterized by many properties that are related to the environment, mobility, and invocation. Therefore, smart services will have specific features as (O. Akhrif, El Idrissi, and Hmina 2018):

- User-centric: based on the specific context and the preferences of the users;
- Ubiquitous: reachable everywhere and from any devices;
- Highly integrated: based on the integration of services and data from several and different applications or the social cooperation of multiple users;(Petrolo 2016)
- Adaptive: a smart service can flexibly sense, understand, and adapt to the user's needs. To achieve a successful personalized service, two fundamental requirements are needed. The first is the ability to understand the behavior of the users and the second is the ability to adapt efficiently, to the user's changing behavior over time;(Witten, Frank, and Hall n.d.)
- Context-awareness: To provide services to occupants, context data is also needed;(Sciarretta, Carbone, and Ranise 2016)
- Open: To realize open services, the developer site is disclosed to the public. Everyone can bring their innovation to utilize the Internet of things (IoT) devices and develops new applications, by opening the device API to the public, developers can bring their innovation and develop new applications to provide service to occupants. (Sciarretta, Carbone, and Ranise 2016)

5.3.4. Service-oriented architecture for Smart Cities

The smartness concept is a need as previously explained in smart cities, through these environments, many levels of smartness are concretized. In our study, we will be interested in the smartness-oriented service called "Smart Service". The smart services are a concept that emerges from the classic terms of SOA architecture that ensures a functional aspect and touches additional treatments by allowing a reinforced description, an adapted deployment, a perfect composition, and more relevant selection according to several contextual parameters as User Profile, Interaction History as well as preferences. In a smart service-oriented architecture, a smart service is characterized by additional processing besides the invocation and boosted by a dynamics layer that achieves smartness. Indeed, its description is enriched by additional metadata ontologies that facilitate the classification and the selection, and by an agent that analyzes the request to adapt to the intention of the user. Or Adapted Architecture by adding steps in the service lifecycle. Several works in summer give an overview of the main mechanisms for developing smart services according to architectural, technical, semantic, contextual and intentional orientations:

Architectural. Define 'smartness' means that services are implemented as RESTful APIs, using URIs for resource identification, HTTP for communication and message transmission, as well as standard formats for data exchange, such as XML, JSON, and RDF, which ensures good accessibility and integration in terms of using of standardized access mechanisms. RESTful interfaces are enhanced with semantics by creating a semantic service description and defining the inputs and outputs in terms of Linked Data concepts (Gu, Pung, and Zhang 2005).

Technical. (Guo, Zhu, and Yang 2017) Propose a smart service system architecture for IoT that supports heterogeneity, reconfiguration, and scalability of networks, devices, and services. It consists of three Smart services layers: i) smart service terminal (SST), ii) smart service network (SSN), and iii) smart service system (SSS) platform, all of these layers are context-aware, in order to elaborate a smart system that supports heterogeneity in an IoT environment. On the other hand, propose one recite to achieve a successful personalized service, two fundamental requirements are needed. The first is the ability to understand the behavior of the users and the second is the ability to adapt efficiently, to the user's changing behavior over time, personalized service making is based on a combination of Machine Learning, IoT (Internet of Things) and Big Data technologies (Chin, Callaghan, and Lam 2017).

Semantic. The appearance of the web semantic has combined service concept with a semantic description and ontology specifications based on Web Ontology Language (OWL), followed by several semantic service description and composition languages (eg OWL-S2, SWSF4 and SAWSDL5) (Urbieta et al. 2017). Most of these languages allow specifying services as a set of provided capabilities, which are the set of functionalities provided by the services of the smart environment or required for the realization of the user tasks6. Semantic Service Oriented Architecture can build semantic systems that can take advantage of semantic description, composition, and discovery of services, in order to support the specification of services in terms of their semantic signature, context-aware behavioral specification, and conversion (Urbieta et al. 2017).

Contextual. Context-awareness is one of the main properties of smart service, characterizing the service's ability to take the environment into account, collect information about it, and react accordingly. It also enables services to dynamically adapt to the current situation, such as the current physical location and/or social activity; the complex deployment conditions of mobile environments (scarcity of resources, connectivity, etc.) and user context preferences end specified parameters. Integrating non-functional elements represented by ontology to define the context of a service, allows the smart service to reach a level of adaptation to the user profile and environment changes.

Intentional. Intentional services have been proposed to bridge the gap between low-level, technical software-service descriptions and high-level, strategic expressions of business needs for services. Semantic annotation for Web services (SAWSDL) added to the descriptor is based on intentional service ontology. This ontology is built upon the intentional service model and the goal model (Aljoumaa and Souveyet n.d.); intention service description permits to facilitate the mismatching between the user's intention and the service goal by integrating the notion of context, QoS and goal in OWL-S service description (Khanfir and Djmeaa 2016). The intentional service-oriented architecture (ISOA) allows the resolution of the conceptual mismatch problem between the user's intention expressed as natural language and the service goal that should be achieved (Khanfir and Djmeaa 2016), and represents services at a high level according to an intentional perspective, referring to the intention they can fulfill rather than the function they perform. Such services, named intentional services, are expressed in terms of intentions and strategies to achieve them. In iSOA, the business provider intentionally describes the service and publishes them in a registry of intentional service. According to this, the discovery, selection, and matching of service in iSOA are guided by the user's intention (Salma 2014).

Management / life cycle. the idea of a conglomeration of services as a goal that requires the generation of transversal platforms to manage the multiple services in- involved in the smart cities ecosystem (Vilajosana et al. 2013). To achieve smartness scalability of services, the composition and self-organizing of these services needs to be efficient and based on user request parameters and requirements (profile, context….) To provide composition plans that respond to create context awareness smart service composite. This mechanism is mainly based on the application of machine learning and genetic algorithm in order to select and chain basic services to compose new complex services.

The Smart City objectives are creating tomorrow together, based on provided data, new potentials for value co-creation and providing Smart City Services collaboratively can be exploited (Bullinger et al. 2017). In fact, The need for additional competencies and external data leads to highly collaborative value creation in complex service ecosystems (Polese 2012), moreover, designing a smart city ecosystem requires a new kind of interaction and communication between both policies, processes and technologies components of the smart cities.

6. Collaborative impacts on Smart Cities

Smart cities are ecosystems where sustainability is maintained through the interactions of the city subsystems (Rochet 2014). A smart city ecosystem involves a multitude of actors engaged in public and private consumption, production, education, research, entertainment and professional activities. This collaboration requires high levels of human and social capital, as the innovation process is based on knowledge and learning (Appio, Lima, and Paroutis 2019). From this perspective, smart cities are above all seen as “smart communities”, as collaborative ecosystems that facilitate innovation, by creating links between citizens, government, businesses and educational institutions (Appio, Lima, and Paroutis 2019). The evolution of the Smart City (SC) ecosystem that is driven by technological advances, leads to define novel kinds of interactions between Smart City (SC) components that are characterized by autonomy and smartness. Generally, Smart City (SC) activities and projects require collaborative strategies to achieve efficient outcomes and promote citizen participation in order to emerge benefits co-creations and truly values exchange. The appearance of open and networked infrastructure, shared information and ubiquitous computing allows to a combination of human-centric collaborative activities, organizational constraints and business challenges in a Smart City (SC) environment that is based on collaborative workflows (Ouidad Akhrif et al.

2020). A set of fundamental factors to achieve collaborative strategies in a Smart City (SC) ecosystem are:

6.1. Collaborative governance

Smart city environments are characterized by strong dynamics and involve many actors; leadership, and often collaborative government processes, address these institutional factors (Neuroni et al. 2019). Collaborative government can apply collective intelligence for innovative solutions to problems; it can also provide shared governance which ultimately fosters citizens' trust in governments (ae Chun et al. 2012).

6.2. Co-creation

Smart cities will see citizens and other city stakeholders actively engage in co-creation and collaborative activities to shape the city. By encouraging the exchange of value among stakeholders, such activities will create the basis for new local business and working models (Scekic and Nastic 2018).

6.3. Crowdsourcing

Most crowdsourcing initiatives solve problems and/or design products or services to benefit businesses, through non-experts and crowd enthusiasts, while open source relies on collections. experts who collaborate openly towards shared goals. It can be argued that open source facilitates a crowdsourcing process in which users with certain expertise or degree of primary users are attracted to participate like the tasks to be performed (Schuurman et al. 2012).

6.4. Transparency

The implementation of ICT in governments and in particular the intensive use of new technologies in so-called “smart cities”, aims to improve transparency, accountability and citizen participation that help societies to increase their public value (Rodríguez Bolívar 2019). Indeed, the principles of transparency, participation and collaboration of transparent government are best considered as strategies anchored in public values that offer the possibility of obtaining greater or additional value by integrating these democratic practices (T. M. Harrison et al. 2012).

The main challenge of a Smart City (SC) ecosystem is sharing and ensuring communication between stakeholders to create additional values, which requires collaboration in terms of knowledge and materials. This collaboration requires layers of knowledge transmission and acquisition as mentioned in Figure 4 (Ouidad Akhrif et al. 2020).

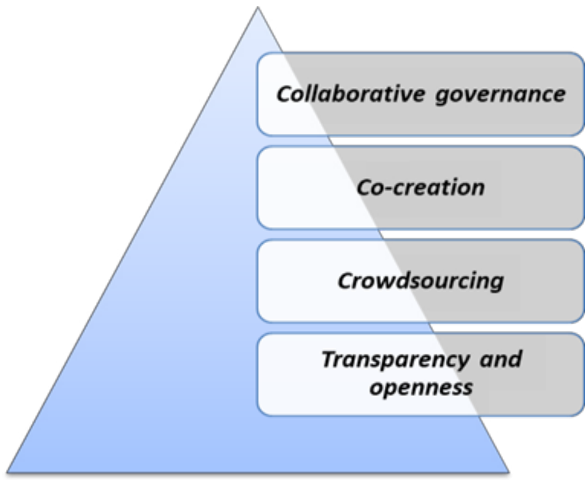


Figure 4. The fundamental factors for collaboration in Smart City

The interdisciplinary of smart cities is associated with the strong demand for the integration of innovative technologies in this field, manifold arising challenges of higher education. There is a need to strengthen collaboration between enterprises on how to work for innovation to design smart city projects, and higher education, which should provide companies with a workforce who have the skills and competencies needed in this sector. Therefore, there is a need to generate different forms of teaching and learning in education at both individual and social levels. All this requires a modern vision of performance based on competencies for learning throughout one's individual life.

7. Smart City needs Smart Education

Education is one of the basic conditions for human development and, consequently, for territorial development. The increase in the education level, in particular of the active population, is positively correlated with the income growth since it is possible to incorporate more knowledge, innovation, technology and differentiation, allowing integrate more added value into the production. More qualified individuals are also more flexible, better adapted to new procedures and activities, more resilient in case of unemployment and more proactive in setting up new businesses (Rego and Sánchez-hernández 2019). Promoting education is a fundamental condition for the development of Smart people skills to ensure efficient interactions among various actors and their smart living environment. In the context of smart cities, smart education is a key issue. It is considered one of the most important needs of every citizen, which plays a key role in achieving the well-being and well-knowing of every dweller. Also, it contributes to developing skills and innovative capacities for the benefit of citizens to be sustainable and intelligent actors for the smart city. The construction and development of the learning environments in a city provide important supports for realizing a learning city, and important contents of a smart city (R. Zhuang et al. 2017). Learning environments in smart cities include school learning spaces, home learning spaces, community learning spaces, working places, learning stadiums, and other virtual learning spaces. Smart learning environments aim to increase the public opportunities to learn, improve the scientific literacy and knowledge acquisition capacity of citizens, and upgrade the city's soft power (D. Liu et al. 2017). Recently, the focus of the smart city is not only on the construction of the environment but also on the human infrastructure in the city. In a Human Smart City, people rather than technology are the true actors of the urban "smartness". The creation of a participatory innovation ecosystem in which citizens and communities interact with public authorities and knowledge developers is key. Such collaborative interaction leads to co-designed user-centered innovation services and calls for new governance models (Oliveira and Campolargo 2015). Smart city projects are taking place through the collaboration of many public and private organizations everywhere in the world. Participation in these projects is a big challenge for the universities, where their strengths are multidisciplinary activity in education, research and societal cooperation, and the success can be ensured with the high quality and professional quality management of the universities (Anttila and Jussila 2018).

8. Conclusion and Futur work

Miscellaneous topics related to the definition, technology, design, planning and management of smart cities were presented to define the scope of this innovative project. The smart city project's shared goals are to improve people' quality of life via the use of smart architecture and technologies. Indeed, the collecting, processing, and analysis of data will make it easier to provide customized and efficient services for city people. Smart cities recognize the need to be a smart educational city for "producing graduates with modern knowledge, practical skills and collaborative attitudes" (Smart Education for Smart Cities: Visual, Collaborative & Interactive n.d.). At the heart of the smart city ecosystem, occurs the Smart University (SU) subsystem as an educational structure organized under a smart system architecture, designed to develop the appropriate learning environment and training of learners, who will have future jobs. In this way, a new ecology of learner-centered education can be constructed. The questions addressed in this thesis are the following: 1. What are the motivations for setting up an intelligent education system? 2. What are the characteristics of learning environments in smart cities? 3. Are there solutions for smartly sharing knowledge within the smart university? If so, what do such opportunities imply?

Chapter II : Universities and Smart City: new concepts, opportunities and challenges

1. Introduction

Digitization has pioneered tremendous strides in transforming the traditional university where human intelligence is primarily invoked in learning practices, towards an intelligent university that dematerializes learning using artificial intelligence (AI) and which integrates intelligent learning platforms, namely MOOCs. Certainly, the traditional university boasts of human intelligence that has served many learning techniques for many years. However, as the content and organization of higher education must evolve with the evolution and needs of society, these requirements have led to the emergence of a Smart University (SU) System (SUS) that is compliant with the Smart City standard in terms of infrastructure, interactions, reasoning, and visualization. The Smart University (SU) system relies on the deep integration of information technology into the educational process and enables the smart learning environment that encompasses a range of smart components, which involve the implementation of an adaptive educational model using intelligent information technologies. Thanks to these smart components, the Smart University (SU) establishes a process of interaction between academics and organizational structures, to promote modern methods of collaboration that increase the success and effectiveness of education. We can't talk intelligence without addressing the Service-Oriented Computing (SOC) paradigm; it allows for a smooth transition from a traditional university to a Smart University (SU) by leveraging the service characteristics of loose coupling, granularity and scalability.

The modern university, taking the student's profile, emphasis, and objectives into account in these decisions, can offer a training curriculum for future professions to generate students for society as they adapt with the demands and changes of society. As a result, the service-oriented paradigm is essential in guaranteeing a collaborative learning environment that is enhanced by intelligent layering, additional treatment, and data-driven.

2. From digital university to Smart University

2.1. Digital university: constraints

Today, with the introduction of the internet and the increased usage of digital technologies, the digitalization of higher education institutions has gradually taken hold. Faced with these new challenges, and cognizant of the necessity to give digital a place, universities must stay perpetually in pursuit of performance and modernity. Their objective is to satisfy a new generation of highly connected students who choose digital tools for learning, gathering knowledge, and keeping connected to the outside world. Of course, digitalization has profoundly transformed the university by introducing new communication channels, persistent connectivity, digital media, and the dematerialization of education through the creation of new concepts, such as e-learning. However, the digital experience has revealed its limitations. Consider e-learning: the learner is alone in front of his screen. There is no interaction, exchange of ideas, or sharing of experiences with other trainees. Another downside is that, because there is no physical trainer, motivation is not always present. Beginning with this point, we must implement a truly intelligent system that meets the expectations of today's university: a contemporary institution that adapts, predicts, recommends, and provides appropriate services to its students.

2.2. Smart University: the emergence of a smart system

Smart universities' core focus is teaching, but they also drive change in other areas such as management, safety, and environmental preservation.The availability of newer and newer technology reflects on how the relevant processes should be performed in the current fast-changing digital era. As a result, a number of smart solutions are being used in university settings to enhance the quality of education and the performance of both teachers and students. (Rutkauskiene, Gudoniene, and Maskeliunas 2016) , and to provide self-learning, self-motivated and personalized services that learners can attend courses at their own pace and can access the personalized learning content according to their difference (Kim, Cho, and Lee 2013).

Considering it the next level of university's evolution based on the integration of (1) Internet-of-Things technology, (2) cloud computing technology, (3) Radio Frequency Identification

(RFID) technology, (4) ambient intelligence technology,(5) smart agents technology, (6) augmented and virtual reality technology, (7) remote (Virtual) labs, (8) location and situation awareness technologies (indoor and outdoor), (9) Wireless Sensor Networking (WSN) technology, (10) sensor technology (motion, temperature, light, humidity, etc.), as well as many other types of emerging and advanced technologies (Uskov, Bakken, Howlett, et al. 2018). A Smart University (SU) could potentially be a proper place where all of these technologies could be examined and applied continuously as a sustainable evolution. Based on the performed analysis done of software systems and related aspects of this evolution (Uskov, Bakken, Howlett, et al. 2018). Smart University (SU) as a smart system can be characterized by its ability "to learn" about itself and, therefore, be able "to self-optimize" teaching and learning strategies to better operate and perform the main business and educational functions (Uskov et al. 2017), additionally, smart-university —as a smart system—should implement and demonstrate significant maturity at various "smartness" levels such as (1) adaptation, (2) sensing (awareness), (3) inferring (logical reasoning), (4) self-learning, (5) anticipation, and (6) self-organization and re-structuring (Heinemann and Uskov 2018).

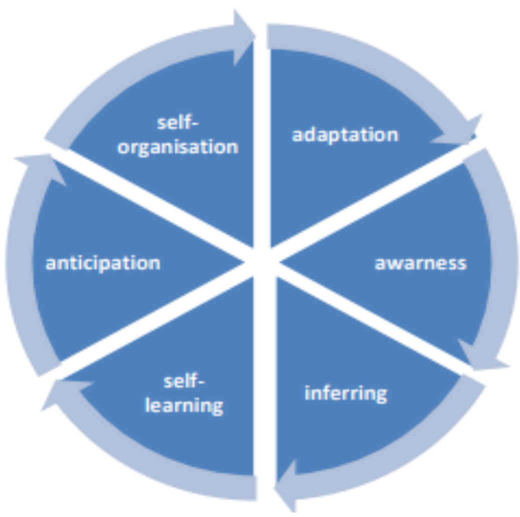


Figure 5. Smartness levels in the Smart system

Another definition of "smart" is provided by the Interactive Technology and Smart Education Information peer-reviewed journal. It states that "SMART" is used as an acronym referring to interactive technology that offers a more flexible and tailored approach to meet diverse individual requirements by being "Sensitive, Manageable, Adaptable, Responsive and Timely" to educators' pedagogical strategies and learners' educational and social needs' (Gomede et al. 2018).

3. Smartness university

The Smart University (SU) as a smart system should significantly emphasize, not only pioneering software/hardware features and innovative modern teaching/learning strategies, but also "smart" features of smart systems (Heinemann and Uskov 2018). Therefore, the designers of Smart University (SU) should pay more attention to the maturity of smart features of Smart University (SU) that may occur on various levels of Smart University's smartness (Uskov, Bakken, Karri, et al. 2018). Several examples of possible Smart University (SU) distinguishing functions for each proposed Smart University (SU) intelligence level are shown in Table 2 (Rutkauskiene, Gudoniene, and Maskeliunas 2016).

Table 2. Smartness levels in the Smart system

SU Smartness level	Details	Examples
Adaptation	To better run and fulfill its main business functions, SU's capacity to autonomously adjust its business functions, teaching/learning methodologies, administrative, safety, physical, behavioral, and other features, etc. (teaching, learning, safety, management, maintenance, control, etc.)	SU easy adaptation to a new style of learning and/or teaching (learning-by-doing, flipped classrooms, etc.) and/or courses (MOOCs, SPOCs, open education and/or life-long learning for retirees, etc.)
		SU easy adaptation to needs of students with disabilities (text-to-voice or voice-to-text systems, etc.)
		SU easy network adaptation to new technical platforms (mobile networking, tablets, mobile devices with iOS and Android operating systems, etc.)
Sensing (awareness)	SU's ability to use multiple sensors automatically to identify, recognize, understand, and/or become aware of various events, processes, objects, phenomena, and so on that may have an impact (positive or negative) on SmU's operation, infrastructure, or the well-being	Various sensors of a Local Action Services (LAS) system to get data regarding power use, lights, temperature, humidity, safety, security, etc.
		Smart card (or biometrics) readers to open doors to mediated lecture halls, computer labs, smart classrooms and activate features/software/hardware that are listed in the user's profile

	of its constituents—students, faculty, staff, resources, properties.	Face, voice, gesture recognition systems and corresponding devices to retrieve and process data about students' class attendance, class activities, etc.
Inferring (logical reasoning)	SU's capacity to develop logical conclusions (or inferences) based on raw facts, processed information, observations, evidence, assumptions, rules, and logical reasoning.	Student Analytics System (SAS) to create (update) a profile of each local or remote student based on his/her interaction, activities, technical skills, etc.
		Local Action Services (LAS) campus-wide system to analyze data from multiple sensors and make conclusions (for ex: activate actuators and close/lock doors in all campus buildings and/or labs, turn off lights, etc.)
		SAS can recommend administrators take certain pro-active measures regarding a student
Self-learning	SU's capacity to automatically gain, acquire, or formulate new information, experience, or behavior in order to enhance its operation, business functions, performance, effectiveness. (It should be noted that self-description, self-discovery, and self-optimization are all aspects of self-learning.)	Learning from active use of innovative software/hardware systems—Web-lecturing systems, class recording systems, flipped class systems, etc.
		Learning from anonymous Opinion Mining System (OMS)
		Learning from different types of classes—MOOCs, blended, online, SPOCs, etc.
Anticipation	SU capacity to think or reason automatically in order to predict what will happen, how to deal with that event, or what to do next	Campus-wide Safety System (CSS) to anticipate, recognize, and act accordingly in case of various events on campus
		Enrollment Management System to predict, anticipate, and control variations of student enrollment
		University-wide Risk Management System (snow days, tornado, electricity outage, etc.)
Self-organization and configuration, re-structuring, and recovery	Under certain conditions, SU has the power to modify its internal structure (components), self-regenerate, and self-sustain in a purposeful (non-random) way without the assistance of an external agent/entity. (It's	Automatic configuration of systems, performance parameters, sensors, actuators and features in a smart classroom under instructor's profile.
		Streaming server automatic closedown and recovery in case of temp electrical outage

	worth noting that self-protection, self-matchmaking, and self-healing are all aspects of self-organization.)	Automatic reconfiguration of the wireless sensor network (WSN) because nodes may join or leave spontaneously (i.e. evolving network typology), university-wide cloud computing (with multiple clients and services), etc.

In addition to the traditional university's components, the Smart University (SU) must contain additional components that characterize and apply the intelligence levels stated in the table. The section that follows describes these primary components or subsystems that we identified based on our review and prior research.

4. Smart University’s components

The emergence of the Smart Institution (SU) gave rise to new concepts that define the actors, environment, and business inside the new university. To do this, an intelligent university taxonomy must be developed in order to identify these key components. We shall define the four essential components of a Smart University in our research:

4.1. Smart learner

The smart learner is a model of an individual or community who interacts with the Smart University (SU) to obtain information for his learning and training certificates or to exchange knowledge in an academic context. The Smart University (SU), for its part, aims to deliver fair, sustainable, and individualized educational services that will be utilized to educate and study in a smart learning environment for the benefit of the smart student. Learner models represent a learner's characteristics such as learning styles, preferences, performance, and abilities.

How does a smart-university system characterize smart learners?

The changing educational environment requires a new learner description to be used for learning environment design and to meet Smart University (SU) requirements; in order to perform tailored and proactive services driven by smart technology integration, researcher results define a range of characteristics that define a smart learner within a smart university:

Profile: Identifying learning profiles assists students in their orientation and promotes educational opportunities for each student depending on his or her profile. To provide a recommendation system to help students, instructors, and administrators in increasing their chances of success in their specified educational goals (Gomede et al. 2018). Students' profiles give important information to educators in understanding the wide range of qualities and quantities that may be examined when grouping students into categories such as abilities, fitness, interests, and perceptions. A suitable profile might be the first step in recommending an appropriate path for students who want to attain specific goals. (C. Romero and Ventura 2007; Cristbal Romero and Ventura 2010). To better understand their knowledge profiles, students can be grouped according to their customized features, and personal characteristics, the creation of the learning profiles uses technologies such as adaptive learning, artificial intelligence, digital assessment, listening and sensing technology, predictive analytics, and hybrid integration platforms (Gomede et al. 2018).

Preferences: Learners' preferences are primarily referred to as learning styles, which help in the selection of appropriate learning objects. Preferences are often specified by the student or deduced by the accumulation of several student models in a group student mode, a Bayesian-Inference-based technique that may rapidly learn user preferences with a small number of user behavior observations (Cristbal Romero and Ventura 2010).

Behavior: Behaviors may be considered as social contexts that can subsequently be used to deliver appropriate reactions or interventions to learners based on their circumstances. Artificial intelligence (AI) in various forms is widely utilized in applications to comprehend and adjust user behaviour (Chin, Callaghan, and Lam 2017).

Performance: smarter leaner prediction Performance is mostly dependent on information about the learner's profile, behaviors, preferences, and activities, in order to provide an idea of the learner's skills and efficiency. This methodology also attempts to construct key performance indicators to measure the achievement of objectives within the context of personalized strategic planning assembled for each student. Is one of the oldest and most widely used data mining applications in education. To make such predictions, several approaches and models such as neural networks, Bayesian networks, rule-based systems, regression, and correlation analysis have been used (C. Romero and Ventura 2007; Cristbal Romero and Ventura 2010).

4.2. Smart knowledge

Smart University (SU) extracts common contents, knowledge and skills that an individual must have in multiple scientific areas (Heinemann and Uskov 2018), according to a smart knowledge model that determines teaching content, educational tools and presentation methods based on the outcomes of the student model (Hwang 2014), and characterized by adaptive representation understanding of various learning environments such as classroom, procedures, and simulation. Knowledge Sharing (KS) can be defined as a process by which people exchange their (tacit and explicit) knowledge to create new knowledge together (van den Hooff, Schouten, and Simonovski 2012). When, (Lin 2007) Identified KS as a process of social interaction by which people can exchange mutual knowledge, experiences, and skills within the organization.

4.3. Smart learning

The Smart Learning paradigm is a convergence concept of ubiquitous Learning (u-Learning) and Social Learning which is expected to improve the educational environment to the advanced level regarding device, network, and education programs using a strong IT infrastructure and advanced social technologies (Kim, Cho, and Lee 2013). It is an abstraction of a smart learning environment, through which the smart learner takes advantage of personalized services that provide access to learning anywhere, anytime and anyway. The smart learning environment is a concept that represents a digital environment designed for achieving self-learning and efficient sharing of knowledge.

4.4. Smart Interaction

This is a feature of a smart learning system that is manifested in the interaction with the environment and gives the system the ability to: i) immediately respond to the changes of external surroundings ii) adapt to the changes of environmental conditions; iii) improve self-development and self-control and effectively achieve goals. These are mechanisms of transmission and technological means through which the learner interacts with his environment promoting his participation, collaboration, and optimization of his capabilities. In an academic environment, it is an ability to adapt to and smooth accommodation of special students with various types of disabilities including physical, visual, hearing, speech, cognitive, and other types of impairments (Uskov, Bakken, Howlett, et al. 2018).

Figure 6 shows the main components of the intelligent university and the different relationships between them :

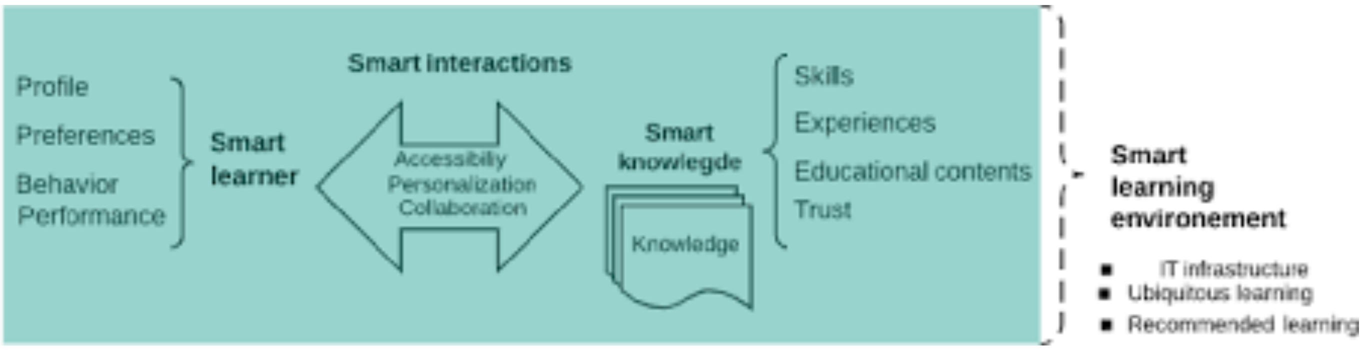


Figure 6. Smart University's components

Smart University's components meet the greatest challenges of any 21-century by providing services such as education, teaching, research and training with high performance to improve and modernize the quality of learning.

5. Smart University's opportunities

Smart University's challenges call for modern sophisticated smart devices, smart systems, and smart technologies to create unique and unprecedented opportunities for academic and training organizations in terms of new approaches to education, learning, and teaching strategies, services to on-campus and remote/online students, set-ups of modern classrooms and labs (Rutkauskiene, Gudoniene, and Maskeliunas 2016). To approach education and how we teach various types of students differently (Rutkauskiene, Gudoniene, and Maskeliunas 2016), we must develop smart universities based on Smart Education, Smart Pedagogy, Smart Classroom, Smart Learning, and Smart Campus. The table below describes, in a summary way, each concept and its opportunities.

Table 3. Smart University (SU) opportunities

Approach	Smart Education	Smart Campus	Smart Classroom	Smart Learning	Smart Pedagogy
Opportunities	Adaptive learning programs; Collaborative technologies and digital learning resources; Personalized services and seamless learning experience; (Heinemann and Uskov 2018)	Smart building management with automated security control and surveillance; Protective and preventative health care; Social networking and communications for work collaboration; Green and ICT sustainability with intelligent sensor management systems;	Learning Disabilities; Speech or language impairments; Visual Impairments; Hearing Impairments; Reading comprehension; Writing comprehension; (Rutkauskiene, Gudoniene, and Maskeliunas 2016)	Provides the necessary learning guidance, hints, supportive tools, or learning suggestions in the right place, at the right time and in the right form; (Gros 2016)	Learning-by-doing (including active use of virtual labs); Collaborative learning; E-books; Learning analytics; Adaptive teaching; Student-generated learning content; Serious games- and gamification-based learning; Flipped classroom; Project-based learning; Bring-Your-Own-Device; Smart robots (robotics) based learning; (Rutkauskiene, Gudoniene, and Maskeliunas 2016)

Smart University (SU) brings new opportunities for Smart learners to well-being and well knowing in the academic environment. The combination of these opportunities either improves existing learning techniques or identifies other concepts that provide fair training and skills acquisition in an environment designed to provide meaningful services for the benefit of students. Among these new concepts, we list Smart Collaborative Learning which is based on collaboration between students as a pedagogical method of sharing and acquiring knowledge. Recently, collaboration is considered as a lever for learning thanks to the advent of innovative technology and precisely the social networks which offred more interactions and communication between users. In fact, Mobile devices and social media provide excellent educational e-learning opportunities to the students for academic collaboration, accessing in course content, and tutors despite the physical boundary, and has a significant contribution to students' academic performance and satisfaction (Ansari and Khan 2020). For this reason, we studied how the Smart Collaborative Learning concept contributes to smartly sharing and acquiring knowledge in the Smart University (SU) ecosystem.

6. Smart University (SU) Challenges

6.1. Technical Challenges

The technical architecture of a smart system has illustrated the different technical challenges faced when setting up a smart university; indeed, the development of a Smart University (SU) that conforms to the model of the smart city is characterized by technical specifications that require the integration of innovative technologies, an ICT-based infrastructure and high quality of data and services.

Internet of things is an issue that allows the integration of a variety of heterogeneous sensors networks and devices in an intelligent system, thus several technical challenges are shown its importance as the effective management of the material resources (battery degradation, charge and communication efficiency), high-performance requirements (Frequency, Power Consumption, Temperature Stability, Long-term Stability, Start-up Time), Network Integration and Communication Protocol, Security Protocol and Synchronization of Communication Between IoT Nodes. The network of IoT generates a huge volume of data that forces a new way of storage, analysis, and visualization intended specifically for big data, which occurs technical challenges related to the quality of data, inference, scalability and error tolerance of data. For the applicative layer, Technical challenges of Smart services are mostly related to Quality of Service (QoS)(Time sensitivity, Location-based service, Number of end-users, Scalability..), and touch technical aspects of a smart service such as performance and availability.

6.2. Business Challenges

Recently, the university began attracting researchers to engage in order to improve its activities which are mainly based on four pillars: learning, pedagogy, education and resource management and infrastructure. With the emergence of smart city trends, in terms of interaction mode, ubiquitous services, digitalization of resources, and intelligence services that adapt and be personalized for a user (learner) in a situation parameterized by a time interval and by applying knowledge derived from prediction procedures, needs and preferences, in order to recommend services improving the learner performance.

Starting with learning which is called in smart-university environments “smart learning”, is primarily intended to provide a learning guide that helps learners to be involved with the user interface, hints and support tools. On the other hand, education also presents a major challenge. In "smart education," several levels to satisfy the learning needs of the student are involved, wherein; we find the concept of collaboration that is needed to involve learners in the learning process. Collaboration is an important motivational element, whether between learners or with their tutors. Smart education is also linked to adaptable services for learners. For the third pillar, smart pedagogy refers to instruments for improving the quality of learning that touches many aspects such as learning-by-doing with virtual labs. There are also intelligent services linked to experiential learning to meet gamification-based learning projects and the integration of the Bring-Your-Own-Device concept in an academic environment.

In the field of intelligent university, we pay attention to collaboration because it is an important criterion to achieve quality learning. Indeed, The keys of the smartness of a Smart collaborative learning environment, in the field of the Smart University (SU) axis, are:

- **Transparency:** ubiquitous learning provides a clear idea about accessing learning resources and services by learners to administration;
- **Collective intelligence:** refers to the ability of a group to emerge smartness based on heterogeneous knowledge driven by team-learners setting;

- **Democratized learning:** offers equitable learning services for all the stakeholders of a Smart University (SU) taking into account accessibility and learning capability of the learner;
- **Smartly share knowledge:** the smart role for sharing knowledge consists to offer pertinent content and personalized universalization of knowledge;
- **Engagement:** smart learning environment deploys engagement strategies to stimulate students attention and interest, in a collaborative environment the engagement can be determined by the integration of the student in collective works, and his participation in the management of his team according to organizational and decisional dimensions;
- **Openness:** Providing access to data that contains external knowledge and experiences of team members is a key for boosting collective intelligence and enhancing skills, through the creation of added value for the learner services.

7. Smart Collaborative Learning

7.1. Introduction

The Internet has led to the development of so-called Web 3.0 and distributed computing applications promoting collaboration between groups of users. The Smart University (SU) has aligned with these changes and adopted intelligent interactions to promote modern methods of collaboration between teams of smart learners. Indeed, the emergence of the Smart University (SU) concept enables smart learning process by encompassing a range of smart components, which involve the implementation of an adaptive educational model using informational smart technologies, in order to allow the transmission of knowledge and experience in new ways and with advanced modes of interactions.

These interactions do not remain transparent; they must be mastered by a layer that capitalizes on history, service research, forecasts, and recommendations. The opportunities offered by the Smart University (SU) and the integration of service-oriented intelligence have fostered the emergence of a new generation of collaborative learning and have led to a sustainable interface between universities and businesses, which remains essential for the development of the skills and competencies of learners.

To achieve that, it promotes the development of a suitable environment for training and research that supports learner-centered. While taking into account the didactics, pedagogical and logistics factors in order to smartly share knowledge. The servitization of the learning process and its application is a very important issue that allows the university to be proactive, scalable and smart. Thanks to the specificity of the service-oriented concept; the Smart University (SU) adopts technologies and innovates intelligent environments characterized by collaboration, adaptation, and personalization to imply innovative and pertinent strategies to improve academic environments.

7.2. Collaboration in Smart University

Smart University (SU) requires a collaborative vision to create innovative solutions that help to increase educational success. Universities and educational institutes can collaborate for offering a wide range of interdisciplinary expertise through a process of interaction between academics and the organizational structure (Verstegen et al. 2018). Smart Collaborative Learning promotes modern methods of collaboration between teams of the learner and allows a sustainable interface between universities and companies that remain crucial in developing the skills of learners. Thus, collaboration can potentially improve professional skills through which the trainee acquires coordination and co-management techniques by working within teams in various contexts and with other members with different expertise (interprofessional collaboration). It is also a way to develop collective intelligence to achieve common goals (O. Akhrif et al. 2019).

The collaborative web-based system allows the Smart University (SU) to develop its organizational practices and canal of communication throw mobile technologies. Indeed, the web 2.0 tool used by individuals, community groups, small teams, big businesses, government agencies, provides an avenue for any collaborative works and enables collaborative learning with forums, chats, file storage areas, and news services (Chatti et al. 2006).

The integration of collaboration in the Smart University (SU) system is the key to achieving success: it contributes in managing the flow of ideas, the organization of teammates' roles and ensuring the university’s partnerships. These opportunities lead to perform principal goals meetings the university strategies and resolve relevant constraints related to share knowledge between smart learners and emerge an intelligence collective based on heterogeneous ideas. in fact, Smart University’s opportunities have specific collaboration requirements to make intelligent learning applications, we can cite (Heinemann and Uskov 2018):

Smart Collaborative Classroom (SCC): A new type of learning activity and test, that performs experiments with joint works and the collaboration of in-classroom and remote/online students when they work on a joint course project.

Smart Collaborative Education (SCE): A multi-disciplinary student-centric education system— linked across schools, tertiary institutions and workforce training, using: (1) adaptive learning programs and learning portfolios for students, (2) collaborative technologies and digital learning resources for teachers and students, (3) computerized administration, monitoring and reporting to keep teachers in the classroom, (4) better information on our learners, (5) online learning resources for students everywhere.

Smart Collaborative Pedagogy (SCP): Smart Collaborative pedagogy guides educators and students as they strive to use technology for collaboration and navigate the potentially conflicting role of autonomous collaborative learning. It highlights the importance of students contributing personal meanings and using appropriate communication strategies as they work together using interactive technologies in innovative ways.

7.3. Main pillars

The objective of the collaboration is to improve the learning quality and student performance during their educational processes. The collaboration focuses on contextual, personalized, and transparent learning to encourage the emergence of learners’ intelligence, based on heterogeneous knowledge and driven by a team-learning setting. Thus, collaboration as a pedagogical model remains essential for the establishment of an intelligent university that allows sharing through the following three aspects:

7.3.1. Sharing resources

Experts in collaborative learning have developed and evaluated technology-enhanced learning resources in higher education and they maximize educational resource mutualization. In fact, scholars can exchange different ideas and share experiences, expertise, and resources (Kong and Xu 2018). Another benefit of a collaborative learning environment consists of a dynamic mix of many different types of resources and facilities, which teachers should be aware of, and adapt to, the learner in his/her current context to better coordinate and optimally use the educational resources (Fang and Sing 2009).

7.3.2. Interdisciplinary collaboration

Universities and educational institutes of the city can collaborate offering a wide range of interdisciplinary expertise (Psomadaki et al. 2018). The interdisciplinary and collaborative projects were rewarding experiences (based on informal student feedback on both campuses) that focused on enhancing interdisciplinary research, collaboration and shared leadership skills, along with improving critical thinking oral and written communication skills (J. B. Ray 2017). Interdisciplinary collaboration can bring about a greater sense of mutual respect and exchanges of ideas across disciplinary boundaries. This blurring of boundaries, necessary in the field of practice will not affect the development of the disciplines in the academic world. There will always be a need for compartmentalization of the fields of the various fields of knowledge. The maintenance of disciplinary boundaries helps to make the growth of human knowledge more systematic and manageable (Wan and Wan 2020). Therefore, Collaborative learning across borders gave a much broader perspective of knowledge, like every student, has a different background and skills.

7.3.3. Trust

Trust is a factor that mitigates the barriers to collaboration and reduces both orientation-related and transaction-related barriers (Bruneel, D'Este, and Salter 2010). This may be because trust relies on strong bonds of mutual understanding and adjustment. Therefore, trust helps firms to manage their differing expectations of research and to lower the considerable transaction costs of working with university partners (Bruneel, D'Este, and Salter 2010).

7.4. Building high-performing teams

7.4.1. Motivation

Collaborative Learning is an educational approach to teaching and learning that involves groups of students working together to solve a problem, complete a task, or create a product(Fang and Sing 2009; J. B. Ray 2017). It main features include: (a) active use of online tools and to instruct students:(b) student collaboration (interaction, communication) with those teachers and other students: and (c) team-working approach to problem-solving while maintaining individual accountability. Based on published reports, Collaborative Learning (a) develops social interaction skills: and (b) stimulates critical thinking and helps students clarify ideas through discussion and debate (Ramírez-Donoso, Pérez-Sanagustín, and Neyem 2018). The main aim of the collaborative learning module was to integrate interdisciplinary learning while engaging students and helping them to develop knowledge and problem-solving skills (J. B. Ray 2017). Thus, building a team of learners is crucial to group learners in an optimal way.

7.4.2. Team building

Teamwork is an effective way to improve learning outcomes. It is a means of learning used in most general or professional programs. Building an efficient and harmonious team is a major challenge for collaborative learning. Therefore, the success of this environment is often the result of a close collaboration between the different teammates, allowing a convergence of the knowledge of each of these members. The cohesion of teamwork is based on the quality of the relationships between its different members in order to achieve optimal objectives. To achieve that, there are three models of team composition:

i. Composition Based Learner

The smart learner plays a vital role in building a Smart University (SU) by participating in successful learning processes and problem-solving. This is why collaborative strategies consider smart learners as an important part of team building that is focused on the profile and abilities of the learner. In order to group it under knowledge, skills, attitudes, and values (Liao et al. 2020). The participation of each learner takes into account accessibility and capabilities and requires a deep understanding of profile members in order to create complimentary participation between each member of the team. Its main challenge is to integrate and stimulate the participation of a learner to: 1) share knowledge; 2) integrate each student in the learning process, and 3) develop communications and collaborations skills within a team of learners.

The composition of a team in an intelligent university has certain specificities because its main objective is to share and transmit knowledge between all learners in an optimal and personalized manner, as opposed to the collaboration in a professional field that aims to realize a project. More precisely, collaboration is a pedagogical means of developing team spirit among learners by pooling the material resources used in a course or an educational project, as well as exchanging ideas. Unfortunately, in a university system, the tutor is responsible for building the work team using data that remains restricted in accounting for the abilities of each member. This method fails to encourage effective sharing of knowledge and achievement of collaborative work goals.

In this regard, the prediction of the most appropriate collaborator is a solution that allows building an agile training team thanks to the self-organization team, which relies mainly on the following prediction-based parameters:

- User profile;
- Regrouping students by performance;
- Homogenous similar collaborator;
- Heterogeneous difference collaborator;
- Randomly groups.

ii. Composition-Based Problem-Solving

Problem-solving learning is a great opportunity to improve student collaboration, and these strategies can help to ensure true collaboration in the learning process. Problem-solving learning refers to grouping smart learners around a project according to their interests, which are presented as a description of the project. These aspects motivate students to work together toward a common goal, generating positive interdependence within the team and creating individual responsibilities for each student to benefit the group’s progress (Ramírez-Donoso, Pérez-Sanagustín, and Neyem 2018). The assignment of a learner to a project is based on its interest or deduced from its histories, such as learner evaluations, feedbacks, and performed interventions. Also, task assignment is based on the predefined role of the project that is part of its description. Participation in this project is a voluntary activity that motivates students to improve their academic performance. The team composition based problem-solving relies mainly on the following aspects:

- Educational relationships;
- Learners management ;
- Organizational role ;
- Project description.

iii. Composition based interdisciplinary completeness

In this case, building a team of learners is based on two fundamental parameters: Interdisciplinary collaboration and complementarity. A true team spirit in interdisciplinary collaboration requires the willingness to offer equal recognition, respect, and Interdisciplinary Collaboration responsibilities to all members rather than trying to dominate and dictate the direction of the teamwork (Wan and Wan 2020). Fostering interdisciplinary collaboration in higher education is the main goal of The Smart Collaborative Learning. It provides incentives for learners from diverse disciplines to participate in the collaborative sharing of knowledge like the problem-solving pedagogy.

Complementarity offers a unique basis for interpersonal attraction and group effectiveness. Its congruence occurs where “the characteristics of the individual serve to ‘make whole’ or complement the characteristics of an environment”. Thus, complementarity is defined in terms of mutual need (Tett and Murphy 2002). Building a complementary team is like putting learners into a puzzle. Not every part has the same function, nor every learner has the same role and each role required a different skill and a different profile. At the same time, co-workers may be most compatible when similar in some ways and complementary in others (Tett and Murphy 2002).

Prior approaches used techniques to deduce and calculate the best collaborators participating in a problem-solving team, namely: classification, binary trees, and fuzzy logic. This constitution of teams and communities remains limited compared to the stake of the collaboration, which consists of sharing, discussing and evaluating ideas. Indeed, the power of the collaboration helps with:

- Communication ;
- Time management ;
- Resource allocation ;
- Openness ;
- Completeness ;
- Organization;
- Preferences.

All of these factors require a novel and adaptive processing to promote participation and contribution in a team. Also, the prediction maintained the environment of collaborative work through a permanent suggestion for the benefit of learners to perform together as a group.

7.5. Smart collaborative learning approaches

A collaborative approach is a way in reaching a higher level of a coworking environment, this part discusses the fundamentals of collaborative approaches adopted by smart-university (SU) systems. We cannot define the most effective approach, because each one responds to different needs and specific constraints, in our analysis, we will establish a list of criteria that evaluate the level of intelligence reached by these approaches and evaluate them by the smartness criteria through the Table 4:

Table 4. Smartness criteria

Approaches	**Smartness criteria**					
	Transparency	**Collective intelligence**	**Democratized learning**	**Smartly share knowledge**	**Engagement**	**Openness**
(Chatti et al. 2006)	*	*		*		*
(J. B. Ray 2017)	*		*	*		*

(Fang and Sing 2009)	*			*		*
(Yin and Tabata 2009)	*			*		
(L. Huang, Liu, and Liu 2013)	*					
(Ning et al. 2018)	*			*	*	
(Paskaleva and Cooper 2020)				*	*	*
(Martín, Gómez-pablos, and Muñoz-repiso 2017)		*				
(Kathayat and Braek 2011)	*	*		*		*
(C. Huang et al. 2006)	*		*			
(Civitarese et al. 2019)	*	*		*		*

The next section will be devoted to examining the influence of each smartness criterion on the quality of collaboration and determining the most successful ones.

8. Synthesis

The provided literature research analyzes the terminology of "Smart University" and may resolve any ambiguities, enabling for improved research in this field. Before we begin in-depth study on our thesis topic, we have provided the foundations of the smart university, including these stakeholders, the learning methods, the environment in which the teaching occurs, and the tools. Of course, this intelligent system provides new prospects for modernization of the educational system, fills the gaps left by previous systems, and provides new features to assist learners. Various learning approaches were renewed during the development of the intelligent university, but it still confronts several challenges in maintaining its capacity to be intelligently sustainable in the completion of the learning process. The SU is continually searching for new methods to promote access and knowledge-sharing among many stakeholders. Collaboration is a significant notion among these solutions that creates opportunities for exchanging ideas in a virtual space with numerous contributors. The goals, tools, and duration of the collaboration must all be decided. Indeed, trust is essential for supporting these components and allowing for flexible resource sharing. This chapter investigates the application of evaluation-based smartness criteria as a guideline for identifying successful collaborations within a Smart University (SU) environment.

To establish the collaboration criteria in smart learning environments, a list of key features that respond to the requirements and restrictions of a learning environment is utilized. To begin, many approaches need the availability of more information and data than others in order to gain a thorough understanding of learning parameters and Smart University (SU) component features. On the other side, engagement is a key factor in guaranteeing long-term cooperation and developing effective participation and accountability skills among team members. Based to what has been stated, smartly sharing knowledge is an important issue in offering personalized and relevant content to learners in order to drive successful knowledge generation. Furthermore, collective intelligence is an important part of cooperation for obtaining higher performance learning for addressing complicated issues within a group of learners. All of these criteria have shown to be important, however the user assignment role aspect remains more restrictive. To close this gap, we must develop a collaborative user-centric model able to accurately predict the effective role assignment of learners to complementary teams, based on their profile, which includes personal information, results, performances, and activities.

9. Conclusion

We have emphasized the importance of collaboration in Smart City (SC) and, more particularly, Smart University (SU), which may be expressed as a notion of Collaborative Learning in this work (CL). In essence, collaborative learning is necessary for intelligently acquiring and sharing information through intelligent interactions amongst team's learners. Our research aims to highlight the learner participation in interdisciplinary collaborative work in order to effectively complete a task or reach a common goal. In this regard, semantic modeling is required for data processing in collaborative work in order to increase the ability to search for information. In other words, collaborative work may be represented by heterogeneous data that changes depending on the nature of the project, a range of institutions, the scope of disciplines, and the variety of participant profiles. According to our research, building collaborative teams may be achieved in two steps: The initial stage is to identify possible combinations of complementary teams based on the ontological inferences of the learner's profile and the collaborative project model, using a proposed heuristic completenss processing. The second stage is to integrate the first step's results into the EDM process in order to predict the most successful teams from the combinations already establishe

Chapter III : Ontologies in Educational Data Mining

1. Introduction

Currently, the rapid advancement of computer networks and information technology is ushering all areas of society into the era of big data. Education, being a vital element of all aspects of society, is necessarily influenced by developments in information technology and data. Different types of education systems at different levels of education have emerged successively, and a large number of educational databases and data have arisen. These databases constitute a vast amount of "knowledge" for investigating and comprehending educational issues in order to create a predictive learning environment. The fast rise of educational computerization is supported by educational data mining technologies. It can play an auxiliary role in reforming the learning process thanks to the deployment of artificial intelligence technologies in the sphere of education.

Common applications of data mining in education include improving course completion, guiding students in course selection, managing student profiles, detecting issues that lead to dropout, classifying students, developing personalized programs, and predicting student performance to assist in decision making during student enrollment (Zorić 2020). Machine learning, statistics, data mining, and data analysis are the foundations for the methods, algorithms, and techniques utilized in Educational Data Mining. These strategies use data acquired during the teaching and learning process to uncover hidden knowledge and recognize patterns in data. Educational data mining technology has evolved from rule-based knowledge representation and inference to deep learning-based natural language processing, speech recognition and image recognition, and the algorithm model has been significantly improved.

The use of artificial intelligence in education required the development of a data representation model that is especially built to handle interdisciplinary and cross-cutting applications directly associated with the implementation of smart education. In this case, EDM analyzes data obtained by any sort of information system supporting learning or education; when confronted with such heterogeneity of data, a semantic representation is required to enable relevant data mining.

2. What is Educational data mining?

Educational data mining or data mining for education is the study of using data from education systems in great benefits for learners. This is a new discipline based on the foundations of Data Mining (DM), precisely, all the predictions that characterize the behaviors and achievements of learners, the content of domain knowledge, assessments, educational features and applications (Pe 2016). The use of Educational Data Mining techniques in online learning systems could be of big interest in solving learning problems, such as: What can predict the success of learners? Which sequence of scenarios is most effective for a particular student? What are the student actions that indicate the progress of learning? What are the characteristics of a learning environment allowing better learning? (Bousbia and Belamri 2014). By achieving these goals, educational data mining can be performed to design relevant and smarter learning technology and to better serve learners and educators (R. S. Baker 2014).

Various definitions have been provided for the term “Educational Data Mining” or EDM. They provide insight into this emergings concept and its main related areas.

“An emerging discipline focused with developing methods for exploring the unique types of data generated by educational settings and applying those approaches to better understand students and the environments in which they study.’’ (R. S. J. Baker 2011).

“Due to the increasing availability of educational data, EDM is a learning science and a rich application field for data mining. It allows for data-driven decision making in order to improve present educational practices and learning materials.’’ (McGee 2015).

“The application of data mining (DM) techniques to a specific type of dataset obtained from educational settings in order to answer critical educational problems..’’ (Cristobal Romero and Ventura 2013)

“The gathering, analysis, and reporting of data on learners and their contexts in order to better understand and optimize learning and the environments in which it occurs.”(Chatti et al. 2012)

“Educational data mining refers to the process of extracting useful information out of a large collection of complex educational datasets, whereas learning analytics focuses on insights and responses to real-time learning processes based on educational information from digital learning environments, administrative systems, and social platforms..”(Ifenthaler and Gibson 2020)

These definitions highlight a global vision about the main areas involved in Education Data Mining. EDM can be designed on basis of three main areas (see Figure 7): computer science, education and statistics. The intersection of these three areas also forms other sub-areas closely related to EDM, such as, computer-based education, data mining and machine learning, and learning analytics (Gabriel and Sandoval 2018) .

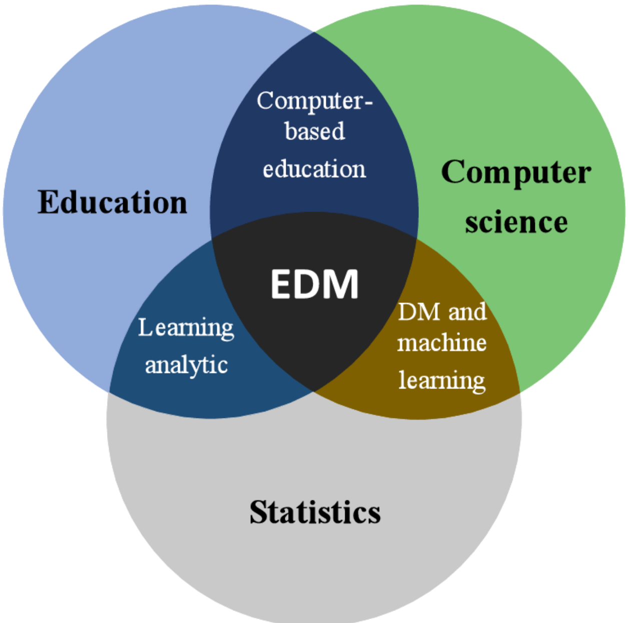


Figure 7. Main areas related to educational data mining

3. EDM process

Educational data mining is the process of transforming data from large educational databases into useful and meaningful information that can be used for a better understanding of students and their learning conditions, to improve learning support, teaching as well as decision-making in education systems (Zorić 2020). Typical parts of an EDM process include data acquisition, preprocessing, data mining, and validation of results (Scheuer and Mclaren 2011). The EDM process extracts useful information from the raw data of educational systems.

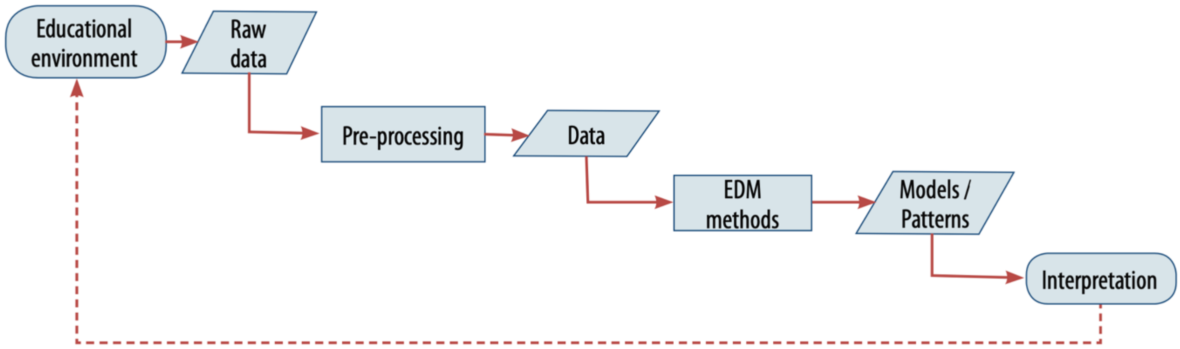


Figure 8. Educational Data Mining process(Liñán and Pérez 2015)

4. Data acquisition

Data is needed to better manage educational environments in order to make these systems more efficient and able to align with trends and be more innovative. Data collection can be from different types of student activities, such as solving homework, taking exams, online social interaction and participating in discussion forums. This data is used by Learning Analytics to extract valuable information, which could be useful for teachers to reflect on their instructional design and the management of their courses (Dyckhoff et al. 2014).

The input data of the educational data mining process can not only come from online learning systems or educational office software but also traditional learning classrooms or traditional test results. Data attributes can be either personal information or learning process information (J. Chen and Zhao 2018). A data warehouse for educational data mining should include student personal details, academic details, exam details, and accounting details (Agarwal 2012). (Bousbia and Belamri 2014) categorized type of collected data into:

- Qualitative or quantitative data;
- Personal, administrative and/or demographic data;
- Answers to psychological questionnaires for measuring users' satisfaction, motivation, skills, cognitive features;
- Answers to questions and/or test scores of the academic system;

- Individual interactions with the educational system: from fine-grained actions such as mouse click to high-level ones such as the number of attempts, the learner browsing pattern, etc;
- Social interactions;
- Visual and facial reactions.

Recently, the emphasis of scientific research has shifted from how to collect data toward how to organize and extract meaningful information from existing data to perform semantic analysis based on data interoperability for providing interdisciplinary learning. Simply put, the learning process is a process of acquiring knowledge, so smart education involves an environmental issue of dissemination and acquisition of knowledge. It can be said that the intelligent educational infrastructure is a knowledge platform, so the identification, formalization, organization and sustainable use of the components of knowledge and learning becomes very important.

In the learning process, miscellaneous scientific and literary disciplines within the university generate a whole field of vocabularies and nomenclatures which determine the specificities of each discipline, this varies greatly contributes to the acquisition of knowledge universally and allows a cross-collaboration between learners. However, this requires an adequate method of modeling, analyzing and interpreting information within the university. It is about, an intelligent system qualified to match the semantics of nomenclatures to define ambiguous terms and create a university thesaurus. Ontologies are adept means for knowledge representation, search, extraction, integration and sharing. Hence, ontologies are used very successfully to represent and manage knowledge within the university. Therefore, we presented an in-depth study of ontologies as well as how it introduces semantic intelligence into the knowledge of educational materials, pedagogy and learning processes.

4.1. knowledge representation and ontology

An ontology is a complex knowledge representation object, and its development requires the use of a certain methodology (Menolli et al. 2013). The knowledge representation languages are used to represent formal ontologies, and standard inference engines are used to make reasoning over ontologies. Similarly, the knowledge bases (KB) can be used to provide a better backbone for the ontology.

In computer science, the first definition of ontology is that proposed by (Gruber. n.d.) who asserts that: "An ontology is a formal, explicit specification of a shared conceptualization", in this sense, the ontology focuses on five relevant terms that have been defined in other research works:

Formal: An ontology is expressed in a knowledge representation language that provides formal semantics. This ensures that the specification of domain knowledge in an ontology is machine-processable and is interpreted in a well-defined way. Knowledge representation techniques help to achieve this aspect. (Grimm, Hitzler, and Abecker 2007)

Explicit: An ontology explicitly states knowledge to make it accessible to machines. Notions that are not explicitly included in the ontology are not part of the machine-interpretable conceptualization it captures, although humans may take them for granted by common sense. (Calvanese et al. 2005)

Conceptual: An ontology conceptually specifies knowledge in terms of symbols that represent concepts and their relations. The concepts and relations in an ontology can be intuitively grasped by humans, as they correspond to the elements in our mental model. In addition, an ontology describes a conceptualization in general terms and does not only capture a particular state of affairs. Instead of making statements about a specific situation involving particular individuals, an ontology tries to cover as many situations as possible, which can potentially occur (Guarino 1998).

Shared: An ontology reflects an agreement on a domain conceptualization among people in a community. The larger the community the more difficult it is to agree on sharing the same conceptualization. Thus, an ontology is always limited to a particular group of people in a community, and its construction is associated with a social process of reaching a consensus (Grimm, Hitzler, and Abecker 2007).

Domain specificity: The specifications of an ontology are limited to the knowledge of a particular area of interest. If the scope of the field of ontology is narrow, it will allow the ontology engineer to focus on axiomatizing details in this field rather than covering a wide range of related topics. In this way, the explicit specification of domain knowledge can be modularized and expressed using several different ontologies with separate domains of interest. (Grimm, Hitzler, and Abecker 2007)

From these definitions, we can gather the basic elements composing the ontology which are:

- Concepts/classes: represent entities or objects defining the existence of the domain represented by the ontology;
- Instances (or individuals): They are concrete objects of a class;
- Relationship: defines various links between classes;
- Properties: serve to define the semantics expressed by each defined concept.

In summary, researchers and developers use ontologies to define standardized and machine-readable definitions and concepts of a specific domain to be utilized for a broad range of applications. Reasons for using ontologies include sharing a common understanding of the structure of information among people or software agents, enabling the reuse of domain knowledge, making domain assumptions explicit, separating domain knowledge from operational knowledge, and analyzing domain knowledge (Noy and McGuinness 2001);

4.2. Ontology vs taxonomy

Defining the fundamental difference between "ontology" and "taxonomy" is a good way to gain clarity. An "ontology with a subclass-based taxonomic hierarchy" leaves less room for doubt than using just the term "ontology" (Rees 2003). A taxonomy containing axioms (additional constraints) can create an ontology. A well-formed ontology provides both the syntactic and the semantics representation of information that can be used in the software. The taxonomy provides the basic information and structure to be converted into an ontology, which is the logical machine-readable structure that can be implemented in software. It should be noted that the DataSet and Taxonomy are also both machine-readable, which allows information to be shared, but they do not contain the logical structure necessary to implement the software.

The logical structure contains the additional axioms (logical assertions) provided by the ontological language is a common form (structure)(Costin 2016).

Figure 9 displays the relation of an ontology with the taxonomy and DataSet:

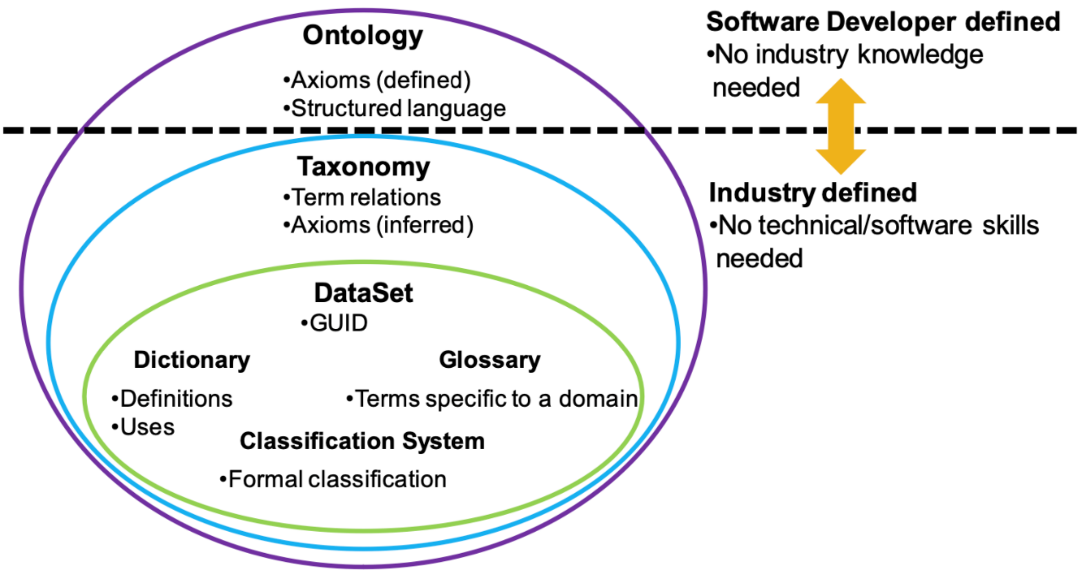


Figure 9. Ontology, Taxonomy and Dataset structure relationships

Building a taxonomy is an essential step before the development of the ontology because a well-defined taxonomy: (1 Reduces ambiguities in domain lingo, (2 Clarifies the semantics of terms, (3 Ensures consistency of terminology, and (4 Reduces the time and effort required to build an ontology (Costin 2016).

4.2.1. Building an ontology

Ontology is the result of the massive and interconnected processing of knowledge in the organization. Thus, the process of building an ontology is represented by the following steps:

Figure 10. Process of building an ontology

4.2.2. Acquisition

Knowledge acquisition, defined by (Fernández, M., Gómez-Pérez 2007), is the process of extracting, structuring and organizing knowledge, it includes the formulation of domain ontologies, the formulation of abstract inference rules and the validation of ontologies (Adetunji et al. 2020). In this phase, all the knowledge about the domain of ontology must be gathered. However, this process of acquiring knowledge is done incrementally, which facilitates the understanding of the subject. (Menolli et al. 2013)

4.2.3. Modeling

Knowledge modeling consists of representing concepts, properties, facts, logical axioms and rules describing an application domain (Ullah and Hossain 2019). The knowledge modeling follows the main steps:

- Define the main concepts: Based on a taxonomy that determines the domain of knowledge, we extract the hierarchy of key concepts and their data properties that will form the ontology of the domain;
- Define object properties: ObjectProperty is used to describe the relationships between concepts in an ontology;
- Data type restriction: This is a mechanism that allows defining new data types that can be constructed from the union or intersection of already existing data types;
- Property restriction: A property restriction is a special kind of class description. It describes an anonymous class, namely a class of all individuals that satisfy the restriction. OWL distinguishes two kinds of property restrictions: value constraints and cardinality constraints (S. Bechhofer F. van Harmelen and Stein 2004).
- Axiom: it is a term that relates entities in order to declare true affirmations. It concerns classes, properties of objects, properties of data and individuals, the axiom is also used to characterize properties;
- Logical rules: Defining logical rules for ontology has been usually considered a demanding task from the viewpoint of Knowledge Engineering (Lisi 2008).

4.2.4. Representation

It consists of providing a set of symbols and a set of rules making it possible to represent the modeling of knowledge in a unified way. The representation of knowledge can be translated by the triplet: subject/predicate/object , or by a graph of nodes connected by arcs.

4.2.5. Evaluation

This step validates the homogeneity of the ontology in terms of defined rules, declared axioms and the domain application, thanks to the reasoning mechanism we can detect the conflicts existing in the ontology. Other approaches validate ontology according to three aspects: Logical (Rule-based), Metric-based (Feature-based) and Evolution-Based.

4.2.6. Reuse

At the knowledge sharing stage, ontology is the knowledge container ensuring the interoperability of knowledge between the different information systems. For the dissemination of knowledge, the ontology represented by the OWL plays the role of a programming language allowing the automatic processing of knowledge.

4.3. Smart University (SU) ontologies

4.3.1. Learning object ontology

The learning object is any digital or web-based resource that can be used, reused, or referenced during technology-supported learning (Mavers 2020). At the learning object level, the educational content is broken down into small units that can be used and reused in various learning environments. For content to be a Learning Object, it must be educational and have intended learning outcomes.

Learning objects composition is one of the main challenges in e-learning management systems and can be improved by exploiting ontological reasoning (Neri and Colombetti 2009). Because, creating a course can be carried out in two phases, the first one is composing concept level entities to obtain an outline of the course, then filling such an outline with actual resources from the repository. Both phases can use ontology-based models to capture specific domain knowledge (Neri 2001).

Along with the increasing use of online and blended teaching/learning systems, learning objects become increasingly valuable and, at the same time, the management of learning objects repository becomes complicated (Wang 2008). Thus, The management of learning objects needs intelligent artificial agents (Neri and Colombetti 2009), serving to solve the constraint of semantic heterogeneity linked to the integration of these objects in a relevant learning context. Generally, the learning context depends on the needs and abilities of the learners; it becomes possible to create educational content that uses different types of added media to provide a very rich learning experience to the student and offer personalized teaching styles according to their specificities.

IEEE LOM (Learning Object Metadata), is a well-known standard used in the e-learning field for modeling Learning objects. It provides a metadata element set for the annotation of learning resources(Ate 2016). According to these standards,(Andrade Menolli, Reinehr, and Malucelli 2012) have established an ontology of Learning objects that we have reused to model our ontology in the contribution part.

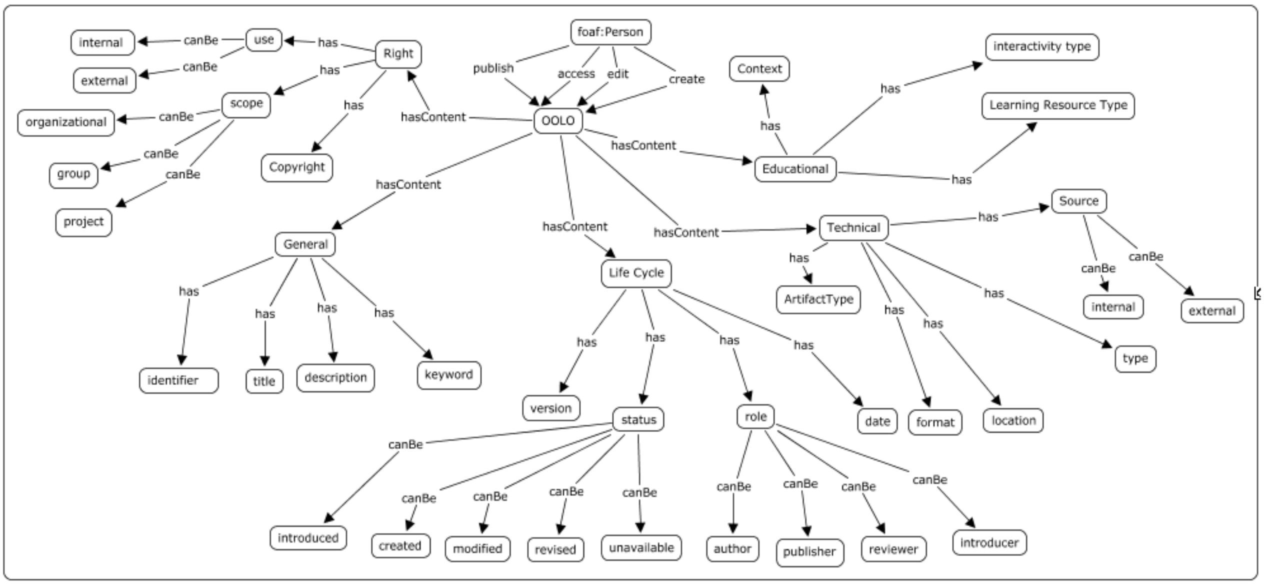


Figure 11. Learning object ontology

4.3.2. User modeling ontology

In the academic field, user modeling requires in-depth knowledge of all information concerning the user and these interactions with his learning environment. This information is useful in creating an expert system suitably designed to manage content, services, learning flows and personalized learning paths according to the specific needs of the students. Designing this decision support system needs resolving the heterogeneity of the data of the user while ensuring respect for their consistency and their semantics.

As pointed out by (Rezgui, Mhiri, and Ghédira 2014; Shrivastava, Mathur, and Joshi 2018; Zine, Derouich, and Talbi 2019), ontology offers a higher conceptual knowledge level to user modeling. The ontology-based user modeling is structured according to well-developed learner model specifications, namely, the IEEE PAPI and the IMS LIP (IMS Learner n.d.). It also reuses terms from well-developed Semantic Web vocabularies, such as FOAF (FOAF n.d.) which introduce a semantic learner model based on the FOAF ontology to support automation of the process of grouping students and preserve at the same time each learner's personal needs and interests (Zine, Derouich, and Talbi 2019).

Our study seeks to model the actors involved in collaborative learning work. We carried out a comparative study to select the most appropriate standards for user modeling which will be part of the ontology that we will present in "2.1 Ontology layer".

We are based on the study of (Zine, Derouich, and Talbi 2019), who presented the table that summarizes the differences between all of the learner models described above based on their proposed taxonomies and supported features.

Table 5. Dimensions of learner model

Supported features/aspects	**Reference Model**			
Personal	**IEEE/PAPI**	**IMS/LIP**	**FOAF**	**eduPerson**
Personal data	+	+	+	+
Competencies	-	+/-		
Affiliation		+	+	+
Accessibility Info				
Info portability	+	+	+	+

Personalization	+	+	+	
Recording Achievements	+	+		
Relations and Community building	+/-		+	+/-
Learning Styles	+	+		
Academic performance	+	+		
Preference	+/-	+/-		
Security	+	+		+
Goal		+		
Disability		+		+/-
Certification	+	+		
Portfolio	+	+		
Learning objective	+	-		

Providing close collaboration between learners is primarily based on standards that manage information about learners' profiles, skills and relationships to give a holistic specification of user modeling in collaborative work. As shown in the table below, the learner models offered by IEEE PAPI and FOAF represent a basic step towards structuring user data which is exploited in various collaborative learning systems.

4.3.3. Learning design ontology

In the educational field, ontologies can be used to create necessary semantics metadata allowing: (1) to structure the learning content of technical documents; (2) model the elements required for the design, analysis and evaluation of the interaction between learners in cooperative computer-assisted learning ; (3) describe the knowledge necessary to define new collaborative learning scenarios and to model groups formation and students' interaction considering their affective states (Reis et al. 2018); or (4) model the semantics of learning objects based on metadata standards (like LOM) (Lama et al. 2005).

Learning design is a formal description of a method enabling learners to achieve particular goals by performing learning scenarios in a certain order within the context of a learning environment (Leo, Pérez, and Dimitriadis 2004). The IMS Learning Design specification is a meta-language that describes all the elements of the design of a teaching-learning process (IMS n.d.). This specification is based on:

- A well-founded conceptual model that defines the vocabulary and the functional relations between the concepts of the LD;
- A model that describes in an informal (natural language) way the semantics of every concept and relation introduced in the conceptual model;
- A behavioral model that specifies the constraints imposed on the software system;

The IMS Learning Design (IMS n.d.) is an open standard that is used to code a wide variety of digital courses, known as units of learning, in a formal, semantic, interoperable and machine-readable fashion. The IMS Learning Design supports a wide range of modern pedagogical approaches such as active learning, collaborative learning, adaptive learning, and competency-based learning (Koper et al. 2004; Leo, Pérez, and Dimitriadis 2004). The learning design ontology based on the IMS Learning Design specification is presented in Figure 12:

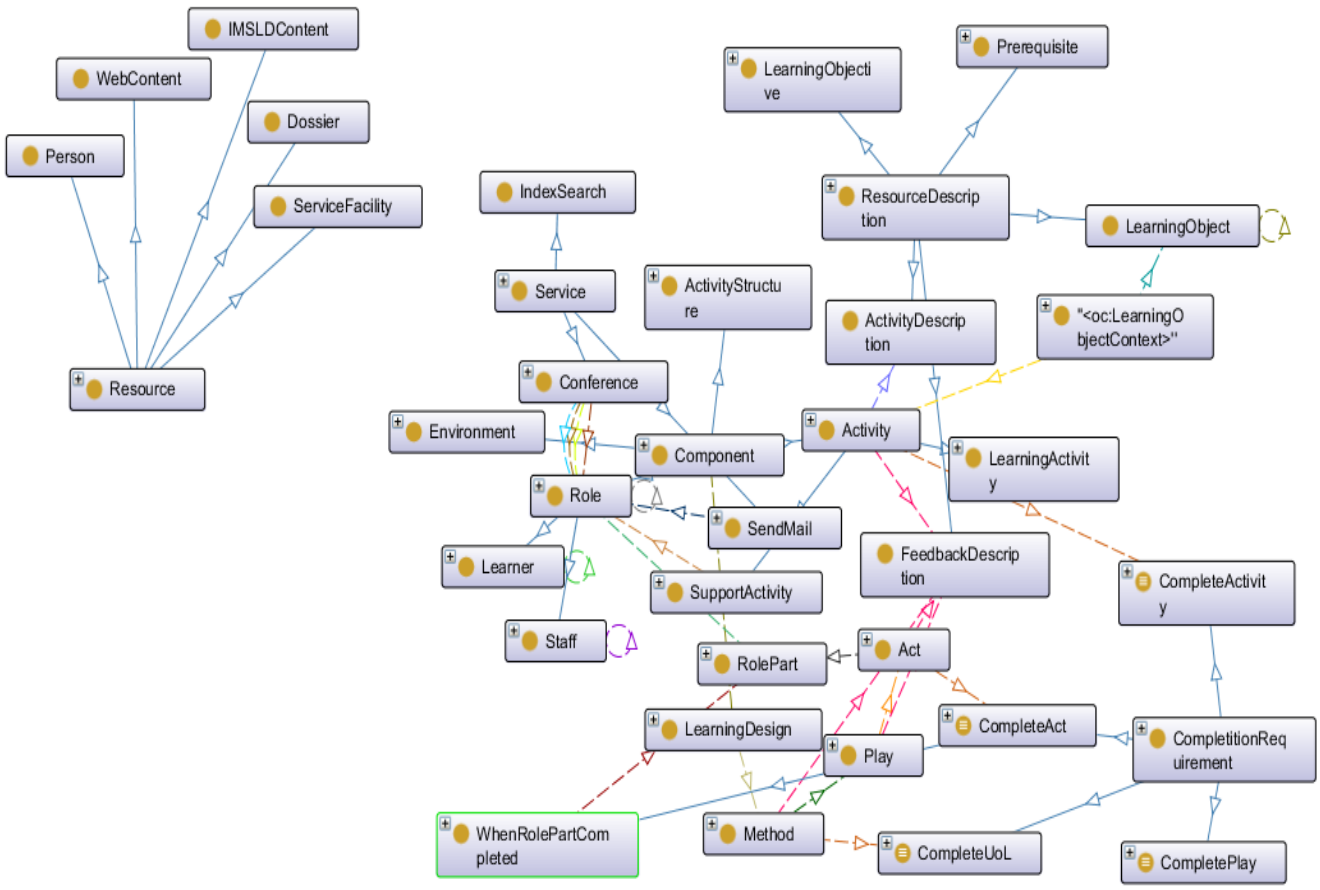


Figure 12. IMS-LD ontology

IMS-LD ontologies provide information models exchange and XML binding that facilitate conceptualization and formalization of a learning process, in order to standardize information modeling and offer more flexibility to represent the various educational process.

4.3.4. Motivation

Any subsystem in the smart city ecosystem has the mission of intelligently managing its activities; for its part, the Smart University (SU) is distinguished by its role of delivering relevant and equitable learning that is responsive to the specificities of learners. Since the learning process is based on the acquisition and transmission of knowledge, an intelligent university must encompass a very broad field of information that includes the nomenclatures and vocabularies of interdisciplinary science, as well as the description of new learning practices that emerged through the integration of technology; for this, a new conceptual paradigm of the intelligent university is required.

The core goal of smart universities is to provide universal and shared knowledge for the benefit of students. This information is acquired through scientific research and experiments in many disciplines, as well as rich educational arrangements in many areas of learning. Indeed, knowledge management is an impetus to smartly manage and transmit content in the smart university, as it can be defined as the process of organizing and harnessing collective knowledge for universities to achieve sustainability and enhanced innovation (Chergui, Chakir, and Mansouri 2020). This process begins with the acquisition and structuring of information, which is then translated to knowledge using formal methods known as "ontology", which allow to provide a formal semantics meaning for information in that domain. Simply expressed, ontologies that are part of the Semantic Web improve computing solutions within universities, by enabling relevant research, reuse, and interchange of knowledge across diverse educational systems. As a consequence, universities would be able to construct a new generation of "Smart" institutions by developing an ontology that connects the primary components, functionalities, through semantic descriptors of the relationships. Precisely, this technology is in high demand for the deployment of collaborative learning solutions; in fact, interdisciplinary collaboration is founded on the heterogeneity of knowledge acquired and the various learner profiles participating in a collaborative project. This is why we created a conceptual model of the intelligent university, including its main concepts and their semantic relationships; this step allows us to represent information clearly and efficiently, which is useful for developing decision-making support when mining educational data, and in order to design a workspace that supports smart collaborative learning.

4.4. Data preprocessing:

Data pre-processing is the first step in any data mining process, it transforms the available raw educational data into a proper format ready for use by a data mining algorithm to solve a specific educational problem (Cristobal Romero and Ventura 2013). In the educational field, it is necessary to acquire adequate data sets and find relevant sources to collect and prepare the data that include all potentially useful information (Vranic, Pintar, and Skocir 2007). The data pre-processing tasks can be reduced to two main techniques (Gibert et al. 2008): Detection techniques to identify imperfections in datasets and transforming techniques oriented to correct detected imperfections in datasets. In this regard, data processing is necessary to make the dataset suitable for various machine learning algorithms (Agarwal 2012).

4.5. Educational data mining:

EDM uses computational approaches to analyze educational data and generates knowledge and cognition to aid the decision-making process. This field exploits the use of statistical algorithms, machine learning and data mining for improving adaptation and personalization in educational environments and systems (Cristbal Romero and Ventura 2010; Cristobal Romero and Ventura 2013). EDM performs techniques and concepts from these different fields in the research, development and implementation of software tools to identify pertinent patterns in large collections of educational data (S. Ray and Saeed 2018). The main functionality of EDM techniques is applying various methods and algorithms in order to discover and extract patterns of stored data. These interesting patterns are presented to the user and may be stored as new knowledge in the knowledge base. All this indicates that EDM is a mature area that will be widely used not only by researchers but also by instructors, educational administrators, and related business from all over the world (Cristobal Romero and Ventura 2020).

4.6. Interpretation and evaluation:

This step acts in evaluating extracted information from the mining step. The evaluation of this kind of information is carried out by experts in the field of education. Usually, visualization of this information is essential, in order to help experts understand and interpret the results of the exploration. Indeed, the results of data mining can be voluminous and require sophisticated visualization techniques to be able to interpret them (Hassan 2017). It is important to emphasize that the educational data mining process is iterative, in other words, the process does not stop when a particular solution is deployed. but it can be an input for a new educational data mining process (Zorić 2020).

5. Predictive methods used in EDM

Educational data mining methods are algorithms and techniques used to extract or "mine" meaningful knowledge from huge amounts of educational data sets. This process is used to recognize patterns and relationships, which is useful for higher education organizations in making data-driven decisions (EDM n.d.). Common methods of educational data mining are categorized into techniques of statistics, machine learning, data mining, information retrieval, recommender systems, psycho-pedagogy, cognitive psychology, psychometrics.(Pe 2016; Venkatachalapathy, Vijayalakshmi, and Ohmprakash 2020). The selection of a suitable method to use depends on the addressed educational issue, the learning environment and the data gathered, either by comparing the experimental results of one of these methods.

5.1. Machine Learning

Machine learning is inspired by human learning which consists of acquiring new knowledge to correct, structure and improve the knowledge already acquired. At this stage, the system improves its behavior as it acquires knowledge. When reaching a maturity level, this smart system becomes an expert in its application area.

Machine learning is a set of methods that computers use to make and improve predictions or behaviors based on data, as a subset of artificial intelligence, it builds a mathematical model based on sample data, known as “training data,” to make predictions or decisions without being explicitly programmed to perform the task (Zhang 2020). Machine learning aims to establish a regressor or classifier through learning the training set and then to evaluate the performance of the regressor or classifier through the test set.Machine learning algorithms can be categorized according to these main learning models (Zhang 2020):

- Supervised learning;
- Unsupervised learning;
- Semi-supervised learning (van Engelen and Hoos 2020);
- Reinforcement learning (Yi, Fu, and Liang 2018);
- Transfer learning (F. Zhuang et al. 2021).

5.1.1. Supervised machine learning

Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs (Olson and Wu 2016) . It infers a function from labeled training data consisting of a set of training examples. The four main steps of supervised machine learning are as follows:

i. Data preprocessing

This step converts the raw data into structured data according to a model required by the machine learning algorithm. Imputation of missing values, selection of characteristics, and additional data transformations are the main steps in data preparation in the Classification and Regression Trees (CART) algorithms. In this case, the dataset is already prepared in the previous steps, but it remains the feature selection step to define the split criterion.

ii. Feature Selection

Feature selection is an important problem in data classification, to ensure that the most discriminative structures are used for classification. The goal of feature selection algorithms is to select the most informative features concerning the class label. During training, a selected quality measure is calculated for all candidate features to find out which one will produce the best split (Agarwal 2012). Our approach that is based on the classification trees requires a specific quality measure called the Gini index.

Gini Index is an impurity splitting method. It is suitable for binary, continuous numeric type values, etc. It was proposed by Breiman in 1984 and has widely been used in algorithms such as CART, achieving fairly good classification accuracy (Manek et al. 2017). The Gini index is based on Gini impurity. Gini impurity is defined as 1 minus the sum of the squares of the class probabilities in a dataset.

Equation 1. Gini Impurity

$$Gini\ Impurity = 1 - \sum_{i=1}^{N} pi^2$$

Where p is the whole dataset, N is the number of classes, and pi is the frequency of class i in the same dataset.

iii. Model training

Supervised classification algorithms use a training set to train the model and a testing set to evaluate the model quality. The input table should be split into two partitions train and test data. In our case, we adopted 80-20% partitioning, where 80% of samples are put into the training set, and the remaining 20% is reserved as the test set for the final model evaluation. We chose a random forest with 50 trees, all trained up to a depth of ten levels and with a maximum of three samples per node, using the information Gini Index as a quality measure for the split criterion.

5.1.2. Decision tree

The classification tree “decision tree” is a very common classification method. It's a kind of supervised learning. The so-called supervised learning receives a stack of samples each having a set of attributes and a category. These categories are predetermined so that through training a classifier can give the correct classification of emerging objects. A decision tree is a tree structure in which each internal node represents a test on an attribute, each branch represents test output, and each terminal node represents a category. This is a supervised learning algorithm based on if-then-else rules. These rules of the decision tree are obtained through training instead of manual formulation(Olson and Wu 2016).

The decision tree algorithm adopts a tree structure and layers of reasoning to obtain the final classification. The decision tree is composed of the following elements:

- Root node: contains the complete set of samples;
- Internal node: corresponding characteristic attribute test;
- Leaf node: represents the result of the decision;

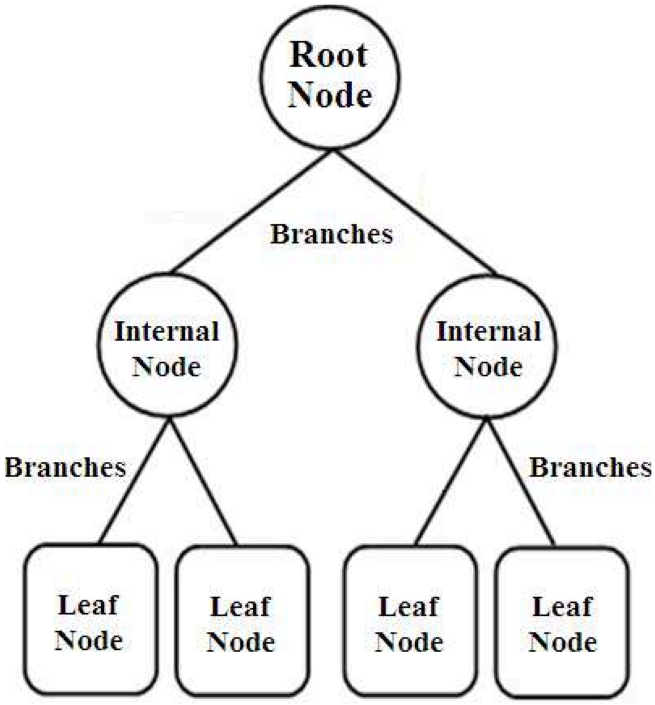


Figure 13. Components of a decision tree (Sá et al. 2016)

There are three stages for learning a decision tree:

- **Feature selection:** It determines which features are used to make judgments. In the training data set, there can be many attributes per sample, and different attributes have different effects. Therefore, the feature selection function is to select the most relevant features for the classification results, that is, features with strong classification ability. The common criterion used in the selection of characteristics is **Information gain**. **Entropy** or **Information gain** can be described as follows:

Equation 2. Entropy

$$H(Y) = H(p) = -\sum_{1}^{i=n} pi * \log(pi)$$

Suppose that we have a discrete random variable Y and the probability is:

$$P(Y = yi) = pi \quad i = 1,2, \ldots . n$$

- **Decision trees generation:** After selecting the feature, it is triggered from the root node, calculating the information gain of all the features of the node, selecting the feature with the greatest information gain as the characteristic of the node, establishing the child node according to the different values of the functionality and generating a new child node by the same way for each child node until choosing.
- **Decision tree pruning:** The main purpose of pruning is to combat "overfitting" and reduce its risk by actively removing certain branches.

5.1.3. Random Forest Algorithm

A random forest is a classifier consisting of a collection of trees structured classifiers {h(x, Θk), k=1, ...} where the {Θk} are independent identically distributed random vectors and each tree casts a unit vote for the most popular class at input x (Jin et al. 2020). The benefits of random forests are numerous. The individual decision trees tend to overfit the training data but the random forest can mitigate that issue by averaging the prediction results from different trees (Pal 2005). This gives random forests a higher predictive accuracy than a single decision tree. Random forest (RF) is an integrated algorithm that belongs to the bagging type. By combining multiple weak classifiers, the final result is obtained by voting, which gives the overall model result a higher precision and generalization ability (Breiman 1999).

The random forest algorithm performs in four steps:

- Step 1: select random samples from the incoming dataset;
- Step 2: create a decision tree for every sample. Then it will get a prediction result from each decision tree;
- Step 3: Establish voting for every predicted result;
- Step 4: Select the most voted prediction result as the final prediction.

i. Model evaluation

Evaluation metrics were used to evaluate the generalization ability of the trained classifier also to assess the model selection (M and M.N 2015). In our case, the evaluation metrics serve to identify the best classifier among the different types of trained classifiers. we will detail this part in the experimental results section.

5.2. EDM in collaborative Learning Environments

The educational data mining process converts data coming from different educational systems, such as traditional classrooms, e-learning and intelligent tutoring systems, into useful information that may be useful for researchers, professors, institutions and students on understanding and evaluating educational systems, aiming for improving the quality of the educational process (Cristobal Romero and Ventura 2020). Both (Abid, Kallel, and Ben Ayed 2016), (Cruz and Isotani 2014) and (Cristobal Romero and Ventura 2020) summarized in more than 100 papers, how Educational Data Mining has been applied to educational data. These surveys focus precisely on predicting students' performance and success, student modeling, courseware management, social network analysis and building teams of learners in the collaborative learning environment. Our research concerns the use of the data mining algorithm to solve the issues of group formation in collaborative learning environments. In the literature works, (Moreno, Ovalle, and Vicari 2012) proposed a method based on a genetic algorithm approach for achieving inter-homogeneous and intra-heterogeneous groups, its main goal is to obtain inter-homogeneous groups, which are as similar as possible to the general characteristics of the total sample of students, but also considering the heterogeneity inside each one. On the other hand, (Cen et al. 2016) defined a model for grouping learners in a collaborative learning environment based on their predicted academic performance. (Hernández-García et al. 2018) focused on a set of log data-based learning analytics indicators to facilitate group assessment in project-based learning courses, and identify relevant predictors of final project results.

(Cen et al. 2016) used Extreme Learning Machine (ELM) based feedforward Neural Networks (NN) and Classification and Regression Trees (CART) as representative instances of Machine Learning techniques applied to predict group performance following the features derived from group interaction data. Also, (Petkovic et al. 2016) used the Random Forest (RF) machine learning (ML) method to predict the effectiveness of software engineering teamwork learning based on data collected during student team project development. Finally, a Decision tree approach is proposed by (Agarwal 2012) which may be taken as an important basis of the selection of students during any course program.

6. Conclusion

Previous research has demonstrated the viability of applying machine learning to make a relevant decision, as well as their success in establishing the prediction model. Each of the above studies use a data acquisition and processing procedure to predict important information relevant to the issue at question. Our research seeks to reach the perspectives of the preceding approaches by combining the semantic aspect for modeling the student and project data, which provides relevant research of the student's profile who would engage in a collaborative project. In this situation, we broaden the area of information that reflects the students, and so the team composition benefits from a wide variety of skills. As well, we allow for interdisciplinary collaboration.

On the other hand, these approaches did not consider the completeness of the students' skills which is a key factor in the composition of the teams allowing the partnership and cooperation in the university environment. As stated in these approaches, the classification is based on algorithms that employ a unique prediction model in the test phase, however our solution provides a dynamic classification that changes with each project step or data update, owing to the layer that determines completeness. We also coordinated the use of Classification and Regression Trees (CART) algorithms and the semantic data source by using the KNIME (Konstanz Information Miner) Analytics Platform, which allows for data analysis, integration of heterogeneous sources, and flexible data manipulation. The next chapter aims to develop an intelligent service that predicts the completeness of teams based on the skills and preferences of learners according to the problem-solving designed in their educational program. In our research, we have designed a smart teaching method consisting to include all students in the teaching process ensuring democratized learning.

Chapter IV : Completeness based classification algorithm: A novel approach for educational semantic data completeness assessment

1. Introduction

According to UNESCO, on 1 April 2020, schools and higher education institutions were closed in 185 countries, impacting 1 542 412 000 students, or 89.4 percent of all enrolled students. Simultaneously, 60% of HEIs reported that COVID-19 has boosted virtual mobility and/or collaborative online learning as alternatives to physical student mobility (Marioni G. 2020). Therefore, universities face unprecedented challenges in functioning securely and correctly; these challenges have a substantial influence on the learning environment, which includes pedagogical approaches, educational resources, and means. Universities shifted emergency from in-person to remote learning in order to continue offering high-quality education through the use of digital learning and online collaborative technologies. The transition to online-only learning demands a large need for technology that enable virtual communication and interaction among learners, such as video conferencing, online collaboration tools, and online learning resources (Cruz and Isotani 2014). This technology remoted learning has supported successfully the learning and teaching process but has affected learners' participation and integration into virtual collaboration. To be more specific, the composition of teams and communities remains restricted in comparison to the stake of the collaboration, which consists of exchanging, discussing, and assessing ideas. (Herrera-Pavo 2021).

As a result, we intend to create a Smart Collaborative Learning Service (SCLS) to integrate all learners in the educational process and provide an equitable environment of involvement and communication. Smart Collaborative Learning's purpose is to serve as a pedagogical model based on problem-solving learning, which refers to grouping learners around a project based on their skills, which are provided as a project description. This features encourage students to collaborate toward a shared objective, generating positive interdependence within the team and creating individual responsibilities for each student to benefit the group's progress. It also assists the student in improving their competency by exchanging ideas, building cooperating abilities, and gaining interdisciplinary information related to academic disciplines. To do this, we proposed developing a system that allows for flexible involvement while also providing opportunities for interdisciplinary collaboration for all learners. Indeed, our approach strives to form teams of learners based on the complementarity of their skills, allowing each of these people's expertise to merge. It considers the heterogeneity of students' profiles and learning knowledge domains as critical information for forming relevant teams based on skill completeness. The Smart Collaborative Learning as a smart system implements significant maturity at various "smartness" levels as shown in " Figure 5. Smartness levels in the Smart system" (O. Akhrif, El Idrissi, and Hmina 2018).

This diagram provides an overview of the system requirements to smartly meet the need for collaboration within the university. This service uses a semantic representation of data to manage the heterogeneity of student profiles, also to structure the representation of the project disciplines. It is based on a heuristic that calculates the completeness between the learner's skills to define the model used during the prediction phase in building complementary teamwork. In this chapter, we present our approach to creating a workflow for acquiring, processing, and mining semantic educational data in order to create a Smart Collaborative Learning Service (SCLS). This workflow is separated into three parts: semantic data representation and inference, completeness processing, and classification data prediction using a machine learning algorithm. To provide further detail, we attempted to present an architecture that depicted the many tiers of service implementation.

This chapter is divided into five sections: Section 2 depicts the overview approach and Smart Collaborative Learning architecture. Section 3 describes the process of creating an ontology, as well as the suggested completeness method and machine learning techniques. Section 4 describes the suggested system, the experiments, and the outcomes. Finally, we closed with a discussion.

2. Approach overview

The Smart Collaborative Learning Service (SCLS) aims to integrate all the learners in the educational process and offers an equitable environment of participation and communication.

For this, we thought of developing a method allowing flexible participation and offering interdisciplinary collaboration opportunities for all the learners. Indeed, our approach aims to build teams of learners based on the complementarity of their skills, allowing a convergence of knowledge between each of these members. It considers the heterogeneous of students' profile and learning knowledge domain as key information to find a pertinent complementarity. The success of this environment is related to predicting efficient collaboration between the different teammates, to smartly sharing knowledge in the Smart University (SU) environment.

The central challenge in initiating a Smart Collaborative Learning Service (SCLS) is successfully creating a suitable and efficient bridge between learners using technological and decision-making tools. To achieve this goal, we design an architecture that responds to the needs and constraints of building complementary teams using semantic university data and machine learning algorithms.

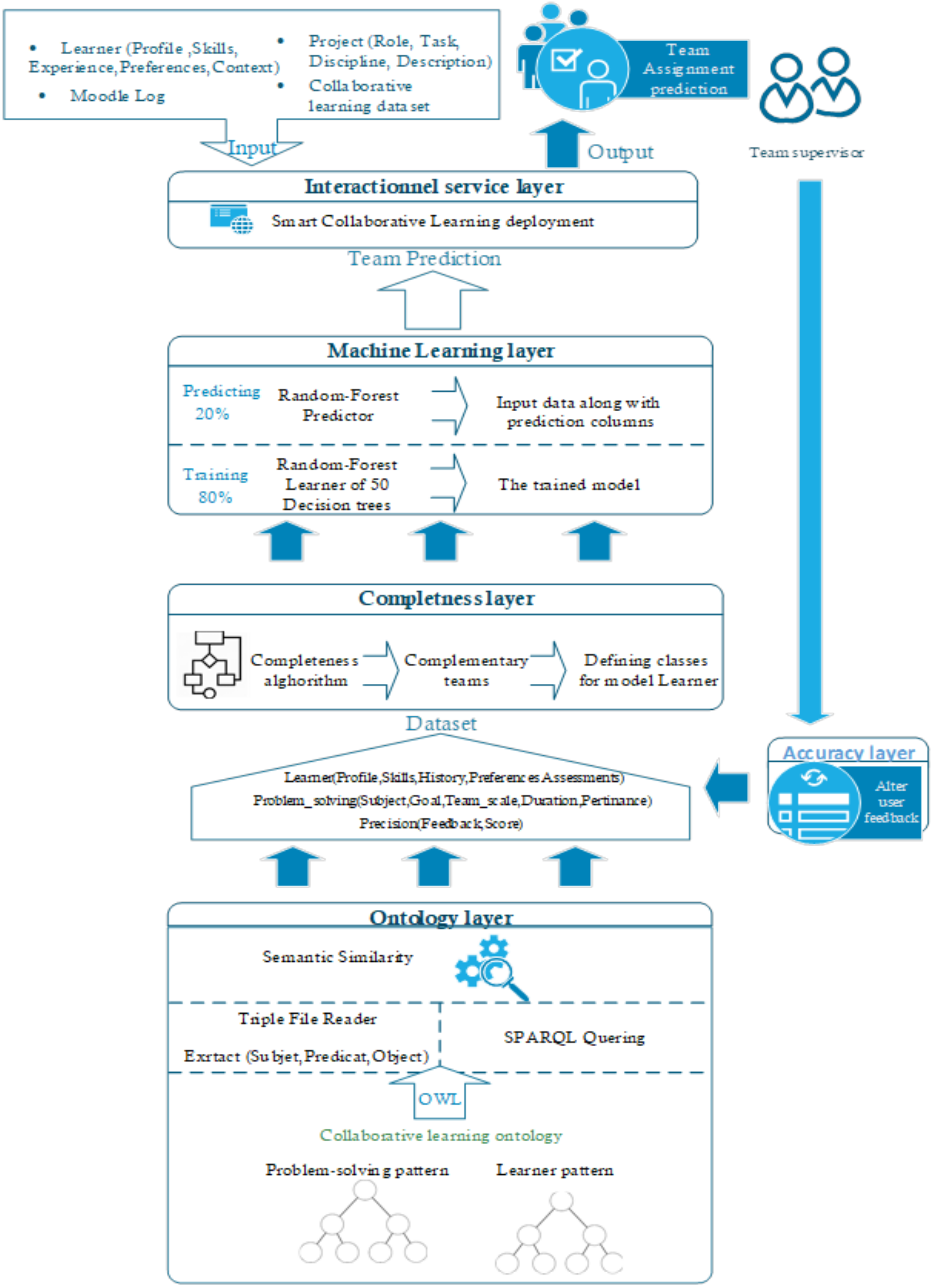


Figure 14. Smart Collaborative Learning Service (SCLS) architecture

According to the architecture of the Smart Collaborative Learning Service (SCLS) . Our approach focuses on the issues related to: (1 identify heterogeneous learners profiles that align with interdisciplinary problem-solving roles,(2 make complementary teams with regards to their respective roles, (3 create a classifier that predicts new complementary teams using a supervised machine learning algorithm.

2.1. Ontology layer

Fostering interdisciplinary collaboration in higher education is the main goal of The Smart Collaborative Learning Service (SCLS) . It provides incentives for learners from diverse disciplines to participate in the collaborative sharing of knowledge like the problem-solving pedagogy. The management of the various university disciplines generates a whole field of vocabularies and nomenclatures which define each of these disciplines, which makes it possible to acquire knowledge universally. However, this requires an adequate method to model, infer and analyze information within the university. Thus, the involvement of several disciplines and heterogeneous profiles of learners requires a semantic presentation of the data involved in the development of the Smart Collaborative Learning Service (SCLS) . Besides, building the collaborative learning ontology gives the possibility of extracting performant profiles and integrating more skills thanks to the semantic similarity.

At a practical level, we used Moodle log (Dalton,2017) and Wayuu dataset (Palma, 2020) to capture the data used in our approach.

2.2. Completeness Layer

Building a complementary team is like putting learners into a puzzle. Not every part has the same function, nor every learner has the same role and each role required a different skill and a different profile. This reflection has led us to perform a new Heuristic for Building Complemetray Teams (HBCT) that calculates the complementarity of learners and to make the different compositions of complementary teams according to learner's skills, this algorithm is a succession of sorting, search by criterion and Boolean algebra blocs.

The complementarity between the different students working in a project “P1” is calculated as follows:

Equation 3. Completeness formula

$$\sum_{i=1}^{N} d\mathrm{i} = 1$$

Where *di* is the index of student role, and *N* is the number of the project competency.

This formula is implemented by following the steps appearing in Figure 15:

Step 1: Initialization: feeding the skills_matrix with project data (Projet_role) and information on student skills (Learner_skills);

Step 2: Prerequisites for beginning the algorithm;

Step 3: Initialization of a counter “i” of the “Tamp” to start analyzing its complementarity with other students in the initial list of students.

Step 4: Update of the skills matrix which groups the skills of the students in the "Tamp" table;

Step 5: The boolean product BP calculates the complementarity of the skills, if BP=1, this means that all the skills of the "Tamp" table are present; thus the group of students stored in Tamp is definitively sent to the “Comp_table” table;

Step 6: If BP=0, then there is a skill required in the project that is not provided; A Miss_comp(Tamp) table is created. This table stores all of the missing skills;

Step7: Search for students (Learners) who have the missing skills;

Step8: Update the skills_matrix to recalculate the complementarities with the additional learners added to fill the gaps in the Miss_comp(Tamp) table;

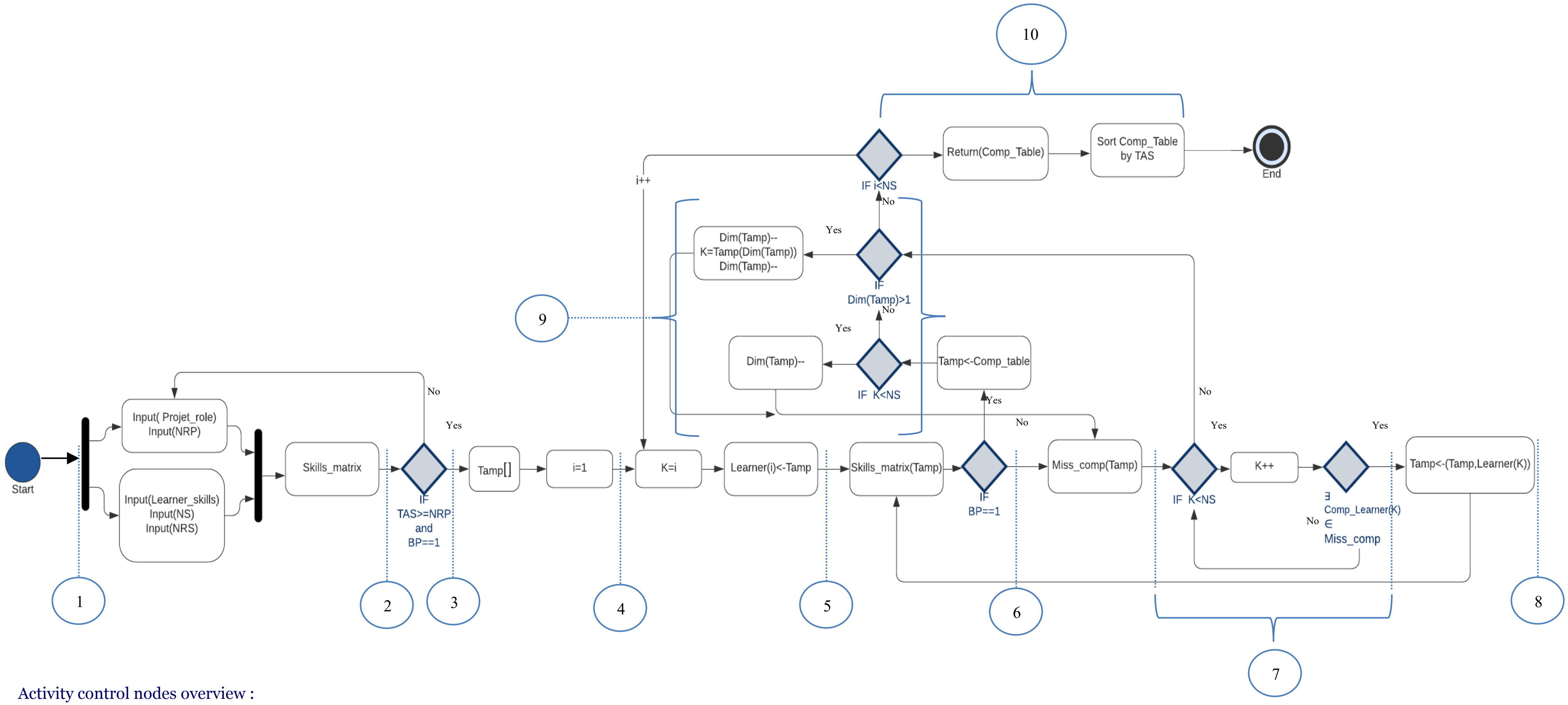
Step 9: Condition on the counter K for adding Learners as needed during the skills_matrix and table of missing skills creation stages. This loop provides an answer to the query, "Are there still learners to continue the algorithm and locate new complimentary learners?";

Step 10 : The algorithm's output once all of the learners have been tested and incorporated in the algorithm.

We defined the keywords used in this heuristic (HBCT) in Table 6:

Table 6. The heuristic keywords

Keyword	Designation
Project_data	Project data
Learner_data	Student data
NPR	Number of roles required by a project
Project_role	Roles required by a project
NS	Number of students
NRS	Number of roles per student
Learner_skills	The skills of a student
Skills_matrix	The confusion matrix between a project and the student's skills
TAS	Total algebraic sum
BP	Boolean product
Tamp	Buffer table
Miss_comp	Missing skills in a skills matrix
Comp_table	Table of students having complementary skills
Dim (Tamp)	The buffer size or the number of students in the Tampon table
Comp_learner	The skills of a student
Learner(i)	The student vector



Activity control nodes overview :

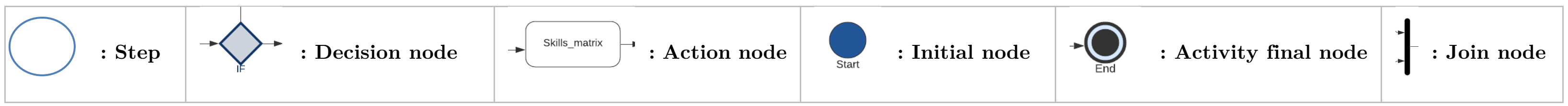


Figure 15. Completeness heursitic

2.3. Machine Learning Layer:

Therefore, we added an artificial intelligence (AI) layer to offer proactive and self-learning mechanisms. We introduced a supervised machine learning algorithm for the analysis and classification of learners in complementary teams.

Supervised learning can be divided into two categories: classification and regression. Our approach solved a classification problem where Team(i) is the corresponding class label of Leaner(i) among the {Team1,…TeamM} classes of targets (Zhang 2020). Therefore, we build a classifier as the following steps:

- Receive the pair: input data " the learner vectors", along with output data " the class that we calculated in the processing layer";
- Train the Random-Forest for classification using the class as a target column and Gini Index as a split criterion;
- Apply the trained model to predict the testing data;
- Evaluate the model performance.

The random forest classifier was performed accurately compared to other classification algorithms as we will show in the experimentation part. In the next section, we will discuss the proposed approach with experimentation and results.

3. Experiments and results

In this section, we will describe the experiments carried out by our approach, the results obtained and a comparative study examining how ontological modeling allows integrating more learners in the computation of complementary teams, as well as the contribution of the heuristic layer which calculates the complementarity of skills to improve the prediction of the assignment of students into complementary teams. Our approach focused on the application of machine learning techniques for classification and prediction in building complementary teams in an academic environment.

At the practical level, we build a workflow using the KNIME analysis platform, which presents the methodology of the proposed approach. The first part of this workflow attempts to extract knowledge from the Smart Collaborative Learning ontology, the second prepares the data that will be used in the completeness processing part, then the result data will be used in the classification and prediction part. Figure 16 details the established workflow:

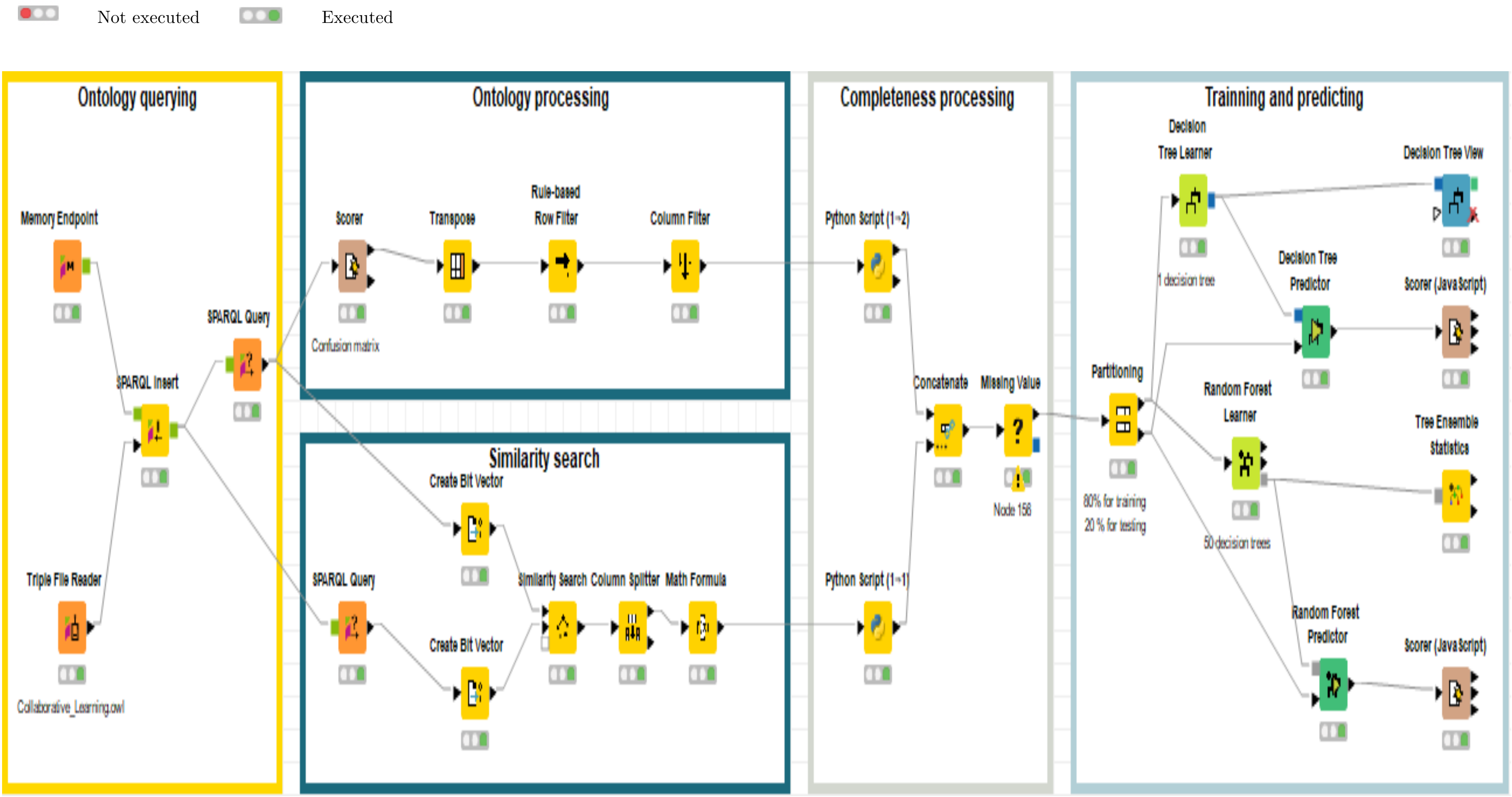


Figure 16. Smart Collaborative Learning workflow

The smart collaborative learning workflow is divided into three main steps:

1) Smart Collaborative learning ontology;
2) Completeness processing;
3) Classification and prediction.

These steps will be detailed in the next sections:

3.1. Smart Collaborative learning ontology

Building the SCL ontology follows the process depicted in the section : “ 4.2.1 Building an ontology”. As a first step, we drew the Smart University (SU) taxonomy, see: “Appendix 2 : Smart University (SU) taxonomy", which represents the main concepts of our ontology. In a second time, we associated these concepts by the semantic relationships as it is mentioned in “mm”.Then, we modeled the Smart University (SU) ontology using the Protégé tool (Tiwari and Abraham, 2020) in the Web Ontology Language (OWL). In our research, we are interested in Smart Collaborative Learning ontology as a part of the Smart University (SU) ontology, therefore, we extracted knowledge of collaborative learning from Moodle log data sources (Dalton, 2017) and Wayuu (Palma, 2020). Figure 17 depicts The Smart Collaborative Learning ontology:

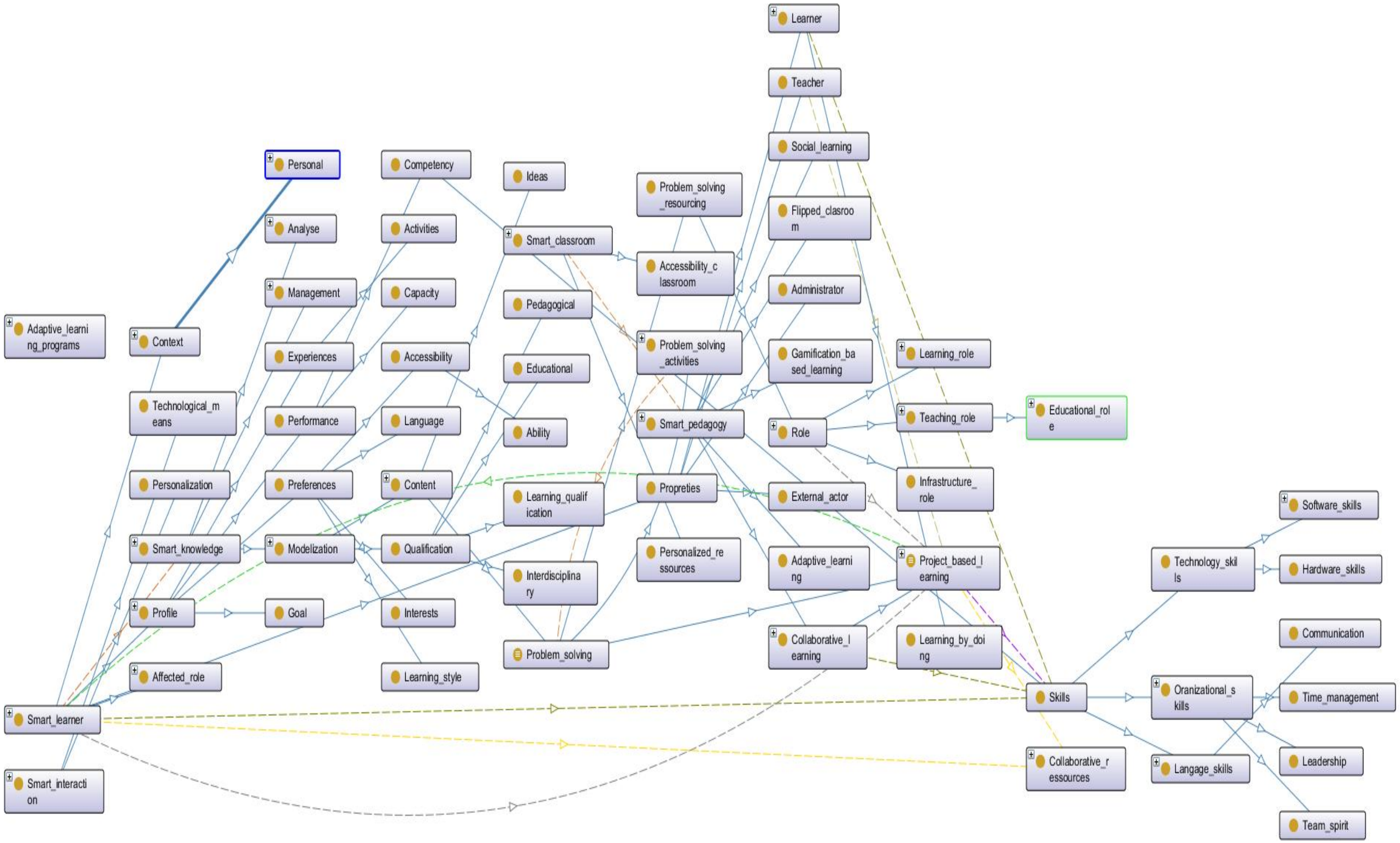


Figure 17. Smart Collaborative Learning ontology

Finally, we evaluated the homogeneity of the SCL ontology using the reasoner pellet. Figure 18 represents the result of reasoning :

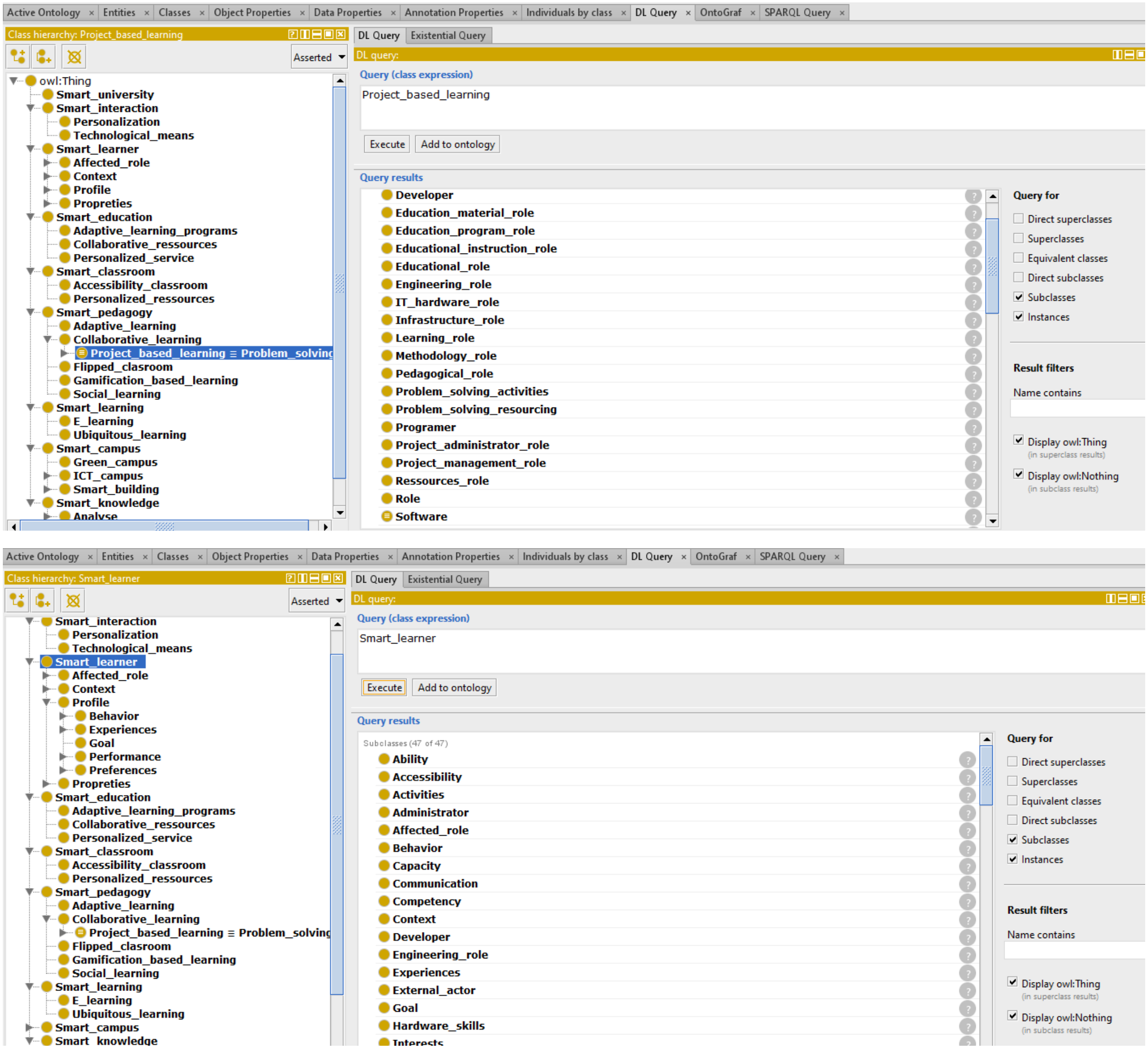


Figure 18. The SCL ontology reasoning

3.1.1. Ontology Processing

The processing of our ontology can be automated with the KNIME Analytics Platform. To achieve this, we created a KNIME workflow that uses a set of semantic web plugin nodes to extracts knowledge from an OWL file. This workflow implements all extracted knowledge and builds a group-based complementarity of learners. The following steps show the details of this workflow:

Step 1: Reading the OWL file

In the first step, the “Triple File Reader” node extracts the content of the “smart_university.owl” and inserts all the triples in a data table. Triples are a collection of three columns containing a subject, a predicate and an object (Figure 19).

Row ID		S pred	S obj
Row152	nticweb.org/hp/ontologies/2020/7/Smart-University#XML>	<http://www.w3.org/1999/0...	<http://www.w3.org/2002/07/owl#NamedIndividual>
Row153	nticweb.org/hp/ontologies/2020/7/Smart-University#Student9>	<http://www.semanticweb.or...	"N9"
Row154	nticweb.org/hp/ontologies/2020/7/Smart-University#Student9>	<http://www.semanticweb.or...	"ST9"
Row155	nticweb.org/hp/ontologies/2020/7/Smart-University#Student9>	<http://www.semanticweb.or...	<http://www.semanticweb.org/hp/ontologies/2020/7/Smart-University#UML>

Figure 19. Extracted triples

Step 2: Querying the OWL file

Once the Triple File Reader is executed, a SPARQL Endpoint can be created using the Memory Endpoint together with the SPARQL Insert node. This allows the execution of SPARQL queries. The following SPARQL query extracts the students who have skills required by the projects of our ontology (Figure 20):

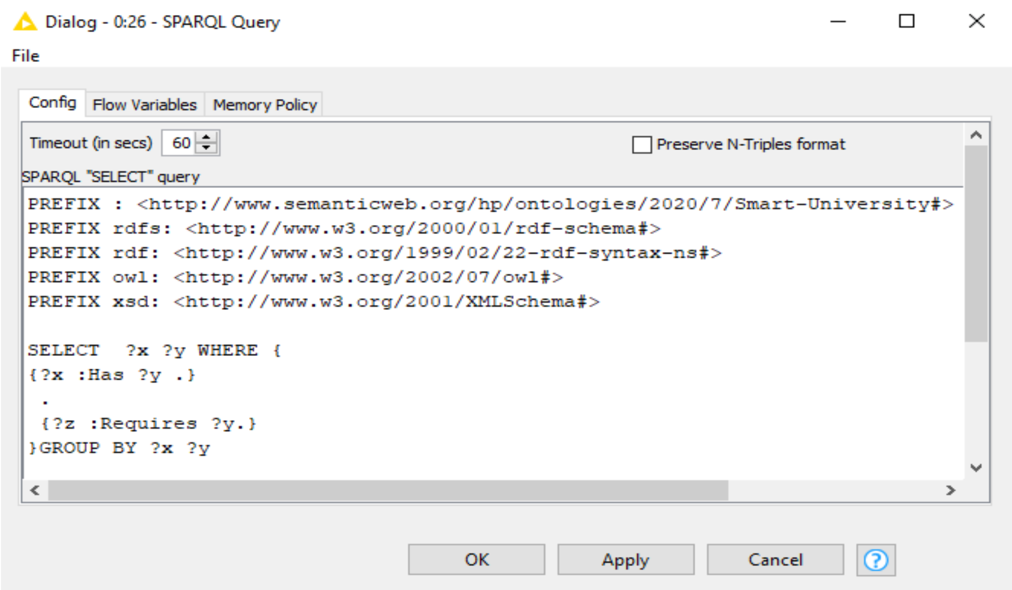


Figure 20. SPARQL query

To develop our skills matrix, we used the “confusion matrix” to obtain the structure requested for our calculation. We also used other manipulation tools namely “Transpose”, “Rule-Based Row Filter” and “Column Filter” to filter the involved columns. The result of this manipulation is a matrix which has for the column the project skills and the rows are students who have at least one or more skills concerning the project skills. Data 0 or 1 in the matrix show the matching result between the student roles and the skills required by the project (Figure 21).

Row ID	http://...	http://...	http://...	http://...	http://...	http://...
http://www.semanticweb.org/hp/ontologies/2020/7/Smart-University#Student6	1	0	0	0	0	0
http://www.semanticweb.org/hp/ontologies/2020/7/Smart-University#Student9	0	1	0	0	0	0
http://www.semanticweb.org/hp/ontologies/2020/7/Smart-University#Student8	0	0	1	0	0	0
http://www.semanticweb.org/hp/ontologies/2020/7/Smart-University#Student10	1	0	0	1	0	0
http://www.semanticweb.org/hp/ontologies/2020/7/Smart-University#Student5	0	1	0	0	0	0
http://www.semanticweb.org/hp/ontologies/2020/7/Smart-University#Student1	0	1	1	0	0	0
http://www.semanticweb.org/hp/ontologies/2020/7/Smart-University#Student7	0	0	0	1	0	0
[illegible]	0	1	0	0	0	0

Figure 21. Skills matrix

3.2. Completeness processing

In this paper, we integrated a scripting python node into the KNIME workflow to implement the heuristic that calculates learner completeness. We chose the python node because it offered flexible data manipulation thanks to the panda data frame compared to the Java Snippet node. The result of this python scripting is a data frame that contains the additional column named "Team", this column describes the assignment of learners to complementary teams calculated by this heuristic. This heuristic constitutes teams of size ranging from 1 to the number of skills required by the project, in our case, we have achieved the possible combinations of complementary teams which are represented by Teams 1-5. Figure 22 shows a snippet of the python scripting:

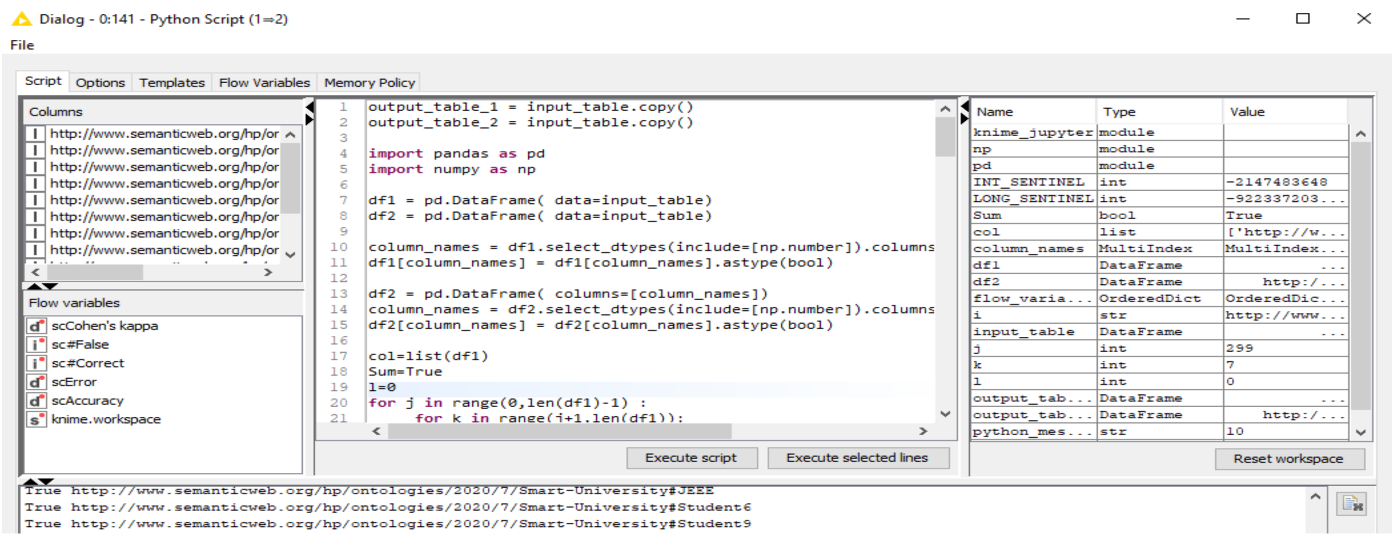


Figure 22. Completenss python script

The structure of the python data frame allows migration to the dataset that will be used in the classification and prediction step.

3.3. Classification and prediction

3.3.1. The dataset description

To build the prediction model, authors need to learn supervised learning algorithms on a dataset then apply this model to a new one. In this paper, we used the dataset that we have exploited in the completeness processing step to calculate the complementary team, this is a preliminary stage to define the output class that will be used in the classification and prediction process. The dataset contains three hundred and seventeen instances and twenty-one attributes that define the learner's sample.

Table 7 describes the selected attributes their type and value, which are relevant for calculating the completeness prediction for learner assignment in a problem-solving project.

Table 7. The trained dataset

Input	Type	Description
Gender	Numerical	[1, 2]

Problem-solving	Numerical	[0, 4]
Disciplines	Numerical	[0, 7]
Skill 1	Numerical	[0, 1]
Skill 2	Numerical	[0, 1]
Skill 3	Numerical	[0, 1]
Skill 4	Numerical	[0, 1]
Skill 5	Numerical	[0, 1]
Consider all possible alternates before doing an activity	Numerical	[0, 4]
Integrate the facts into coherent theories	Numerical	[0, 4]
Share leadership when working collaboratively	Numerical	[0, 4]
It is concerned with the performance of the members of the workgroup	Numerical	[0, 4]
Master the assigned topic	Numerical	[0, 4]
Responsibility is shared	Numerical	[0, 4]
Knowledge is equitable among people who work collaboratively	Numerical	[0, 4]
The objective of learning is achieved	Numerical	[0, 4]
Priority is to achieve a group goal	Numerical	[0, 4]
Worries about the opportunity to present the work	Numerical	[0, 4]
You are aware that you cannot depend on the work of others	Numerical	[0, 4]
Learning is an individual responsibility	Numerical	[0, 4]
Looking for new activities that generate knowledge	Numerical	[0, 4]
Practice learned knowledge	Numerical	[0, 4]
Class	Numerical	**Team 1** **Team 2** **Team 3** **Team 4** **Team 5**

Table 8. Designation of the data set inputs

Attribute	**Designation**	**Instance**
Gender	“Female”, “male”	2

Skill	“Have”, “have not”	2
All collaborative skills	“Rarely”, “sometimes”, “usually”, “always”, “never”	5

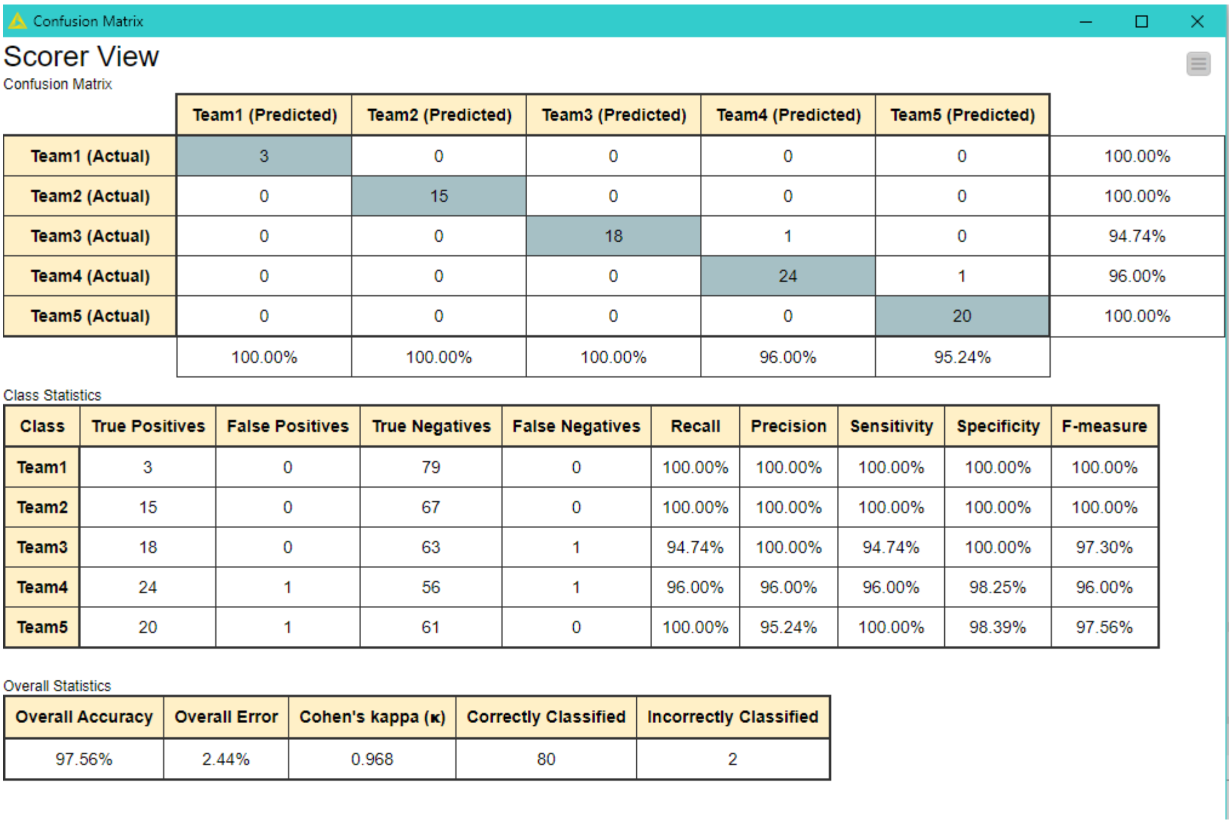
<u>Preprocessing data</u>

The data is ready for classification processing. by following a few steps

- Missing value: They can be replaced or columns with more than a particular number of missing data are suppressed.
- Data encoding: We used Label Encoding to convert each value in a column (category) to a number.
- Splitting and testing: After partitioning the data set in the trained samples (80%) and the tested samples (20%), a feature selection filter designates the subset of columns used in the training set.

3.3.2. Experimentation and validation

In this section, we build two classifiers using Random Forest and Decision tree algorithms on our dataset. The training set is used to build our classifiers, applying the Random Forest and the Decision tree algorithms result in the models confusion matrix as shown in Figure 23 :



Confusion Matrix

Scorer View

Confusion Matrix

	Team1 (Predicted)	Team2 (Predicted)	Team3 (Predicted)	Team4 (Predicted)	Team5 (Predicted)	
Team1 (Actual)	3	0	0	0	0	100.00%
Team2 (Actual)	0	15	0	0	0	100.00%
Team3 (Actual)	0	0	18	1	0	94.74%
Team4 (Actual)	0	0	0	24	1	96.00%
Team5 (Actual)	0	0	0	0	20	100.00%
	100.00%	100.00%	100.00%	96.00%	95.24%	

Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
Team1	3	0	79	0	100.00%	100.00%	100.00%	100.00%	100.00%
Team2	15	0	67	0	100.00%	100.00%	100.00%	100.00%	100.00%
Team3	18	0	63	1	94.74%	100.00%	94.74%	100.00%	97.30%
Team4	24	1	56	1	96.00%	96.00%	96.00%	98.25%	96.00%
Team5	20	1	61	0	100.00%	95.24%	100.00%	98.39%	97.56%

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
97.56%	2.44%	0.968	80	2

Figure 23. Confusion matrix and evaluation metrics of the Random Forest-based completeness model

Figure 23 presents the performance of a Random Forest-based completeness classifier model with a confusion matrix, and class and overall accuracy statistics.

Figure 24 presents the performance of a Decision Tree-based completeness classifier model with a confusion matrix, and class and overall accuracy statistics.

Confusion Matrix

Scorer View

Confusion Matrix

	Team1 (Predicted)	Team2 (Predicted)	Team3 (Predicted)	Team4 (Predicted)	Team5 (Predicted)	
Team1 (Actual)	3	0	4	0	1	37.50%
Team2 (Actual)	0	2	1	0	0	66.67%
Team3 (Actual)	3	1	40	2	2	83.33%
Team4 (Actual)	3	0	3	8	1	53.33%
Team5 (Actual)	1	1	4	0	2	25.00%
	30.00%	50.00%	76.92%	80.00%	33.33%	

Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
Team1	3	7	67	5	37.50%	30.00%	37.50%	90.54%	33.33%
Team2	2	2	77	1	66.67%	50.00%	66.67%	97.47%	57.14%
Team3	40	12	22	8	83.33%	76.92%	83.33%	64.71%	80.00%
Team4	8	2	65	7	53.33%	80.00%	53.33%	97.01%	64.00%
Team5	2	4	70	6	25.00%	33.33%	25.00%	94.59%	28.57%

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
67.07%	32.93%	0.438	55	27

Figure 24. Confusion matrix and evaluation metrics of the Decision Tree-based completeness model

3.3.3. Evaluating the classification model's performance

After partitioning the input data set into the training and testing set, the classification base trees models are built on the training set and the model performance are evaluated on the testing set using the Scorer (Javascript) node. This KINME node produces a set of classification accuracy metrics such as confusion matrix, class statistics and overall accuracy statistics.

The confusion matrix gives an overall idea about the quality of the classifier. Also, it is the basis for calculating other evaluation metrics that allows studying the performance of the classifier, its reports the count of:

Table 9 . Confusion matrix

	Predicted class positive	Predicted class negative
Actual class positive	True positive (**TP**)	False negative (**FN**)
Actual class negative	False positive (**FP**)	True negative (**TN**)

The confusion matrices that were obtained during the application of Random-Forest and the decision trees on our Dataset are presented in Figure 23 and Figure 24.

<u>Class statistics</u>

The last four counts in the confusion matrix allow calculating the class statistics measures to quantify the model performance.

Sensitivity and Recall quantify how many of the actual positive classes are correctly predicted as target classes:

Equation 4. Sensitivity

Sensitivity=Recall=TP/(TP+FN)

Specificity quantifies how many of the actual negative classes are correctly predicted:

Equation 5. Specificity

Specificity=TN/(TN+FP)

Precision measures the ability of the model to attribute positive events to the positive class:

Equation 6. Precision

Precision=TP/(TP+FP)

F-measure combines Recall and Precision:

Equation 7. F-measure

F-measure=2*(Recall*Precision)/(Recall+Precision)

The following diagram defines how well the Random Forest classifier predicts the class values compared to the Decision Tree classifier:

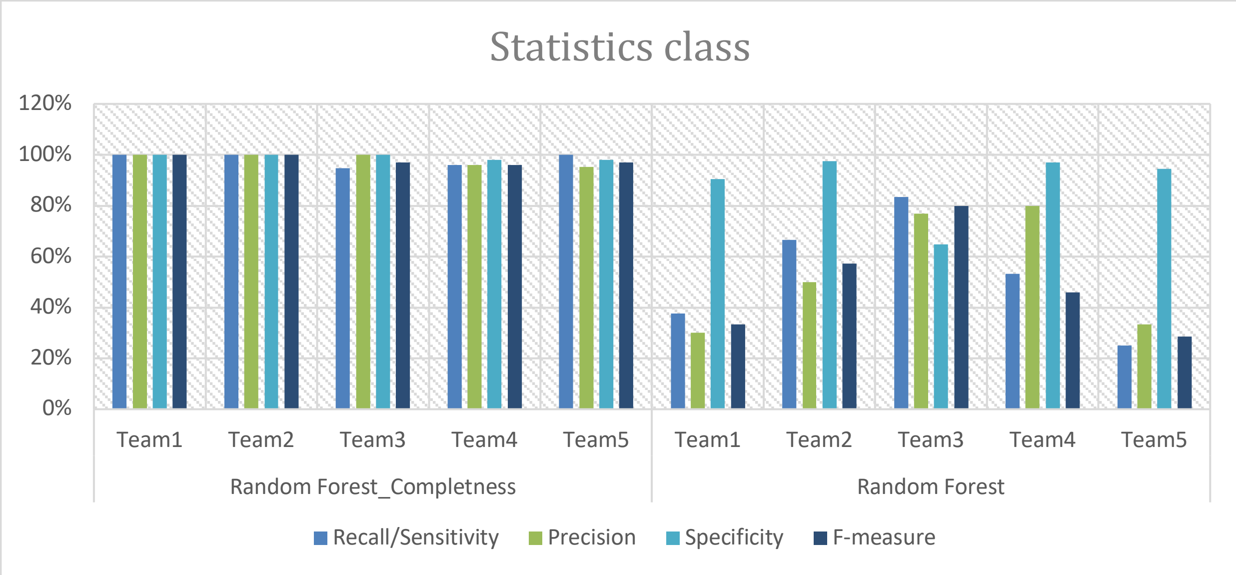


Figure 25. Statistics class comparing the Random Forest and the Decision Tree-based completeness models performance

Figure 25 evaluates the performance of the Random Forest-based completeness model compared to the Decision Tree-based completeness model with statistics class and highlights the most qualified classifier that is the Random Forest-based completeness model.

Overalls statistics

Cohen's Kappa (k) and overall accuracy give global results about the model performance and define the most suitable classifier for our case study. The Random Forest classifier gives more relevant predictions than the decision tree classifier, as it shows in the following diagram:

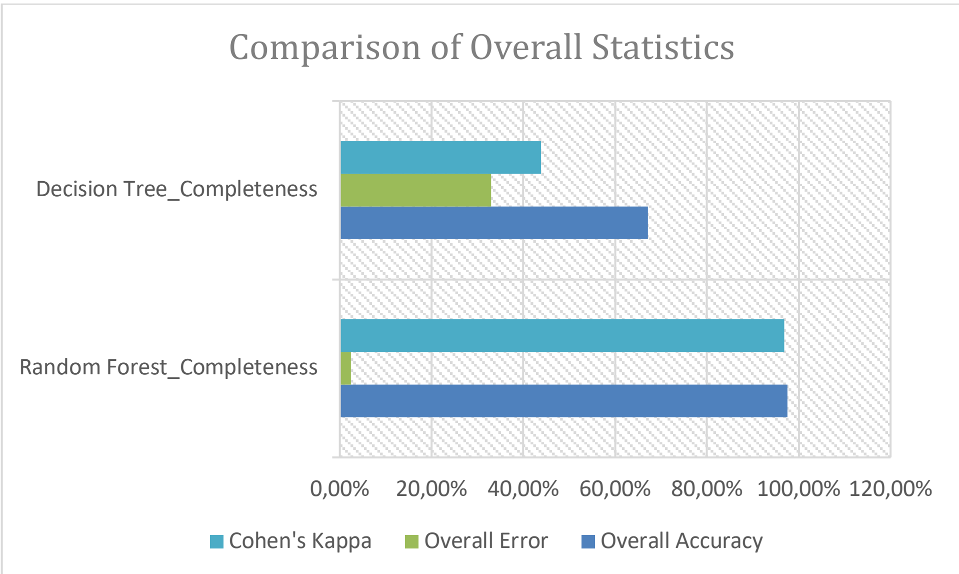


Figure 26. The Random Forest and the Decision Tree overall statistics

Figure 26 shows the overall accuracy metrics results of the classified model using the Random Forest-based completeness algorithm against the decision tree-based completeness algorithm, it also shows how the overall accuracy and Cohen's Kappa get higher value due to the better performance of the Random Forest-based Completeness model.

3.3.4. Experimental comparison

This section assesses the contribution of the ontology and the completeness layers for building an efficient Random Forest predictor of complementary teams. Thus, we compared the evaluation metrics of the classifier proposed by our approach to the results obtained by a Random Forest classifier where the definition of the target classes was defined randomly without taking into account the completeness of the teammates. Figure 27 shows the confusion matrix of the second classifier:

Confusion Matrix

Scorer View

Confusion Matrix

	Team1 (Predicted)	Team2 (Predicted)	Team3 (Predicted)	Team4 (Predicted)	Team5 (Predicted)	
Team1 (Actual)	7	0	0	0	0	100.00%
Team2 (Actual)	0	0	2	0	0	0.00%
Team3 (Actual)	0	0	33	0	0	100.00%
Team4 (Actual)	0	0	2	7	0	77.78%
Team5 (Actual)	1	0	5	0	0	0.00%
	87.50%	NaN%	78.57%	100.00%	NaN%	

Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
Team1	7	1	49	0	100.00%	87.50%	100.00%	98.00%	93.33%
Team2	0	0	55	2	0.00%	0.00%	0.00%	100.00%	0.00%
Team3	33	9	15	0	100.00%	78.57%	100.00%	62.50%	88.00%
Team4	7	0	48	2	77.78%	100.00%	77.78%	100.00%	87.50%
Team5	0	0	51	6	0.00%	0.00%	0.00%	100.00%	0.00%

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
82.46%	17.54%	0.673	47	10

Figure 27. Confusion matrix and evaluation metrics of the second random forest model

In Figure 27, the confusion matrix shows an unbalanced responses classification, the classifier gives no prediction for Team2 and Team5. Despite its overall Accuracy = 82.46%, this model will not perform too well because it gives the Recall=Precision=0% for the outputs Team2 and

Team5. The following figure shows the statistics class obtained by this classifier compared to the Random Forest classifier based on completeness processing:

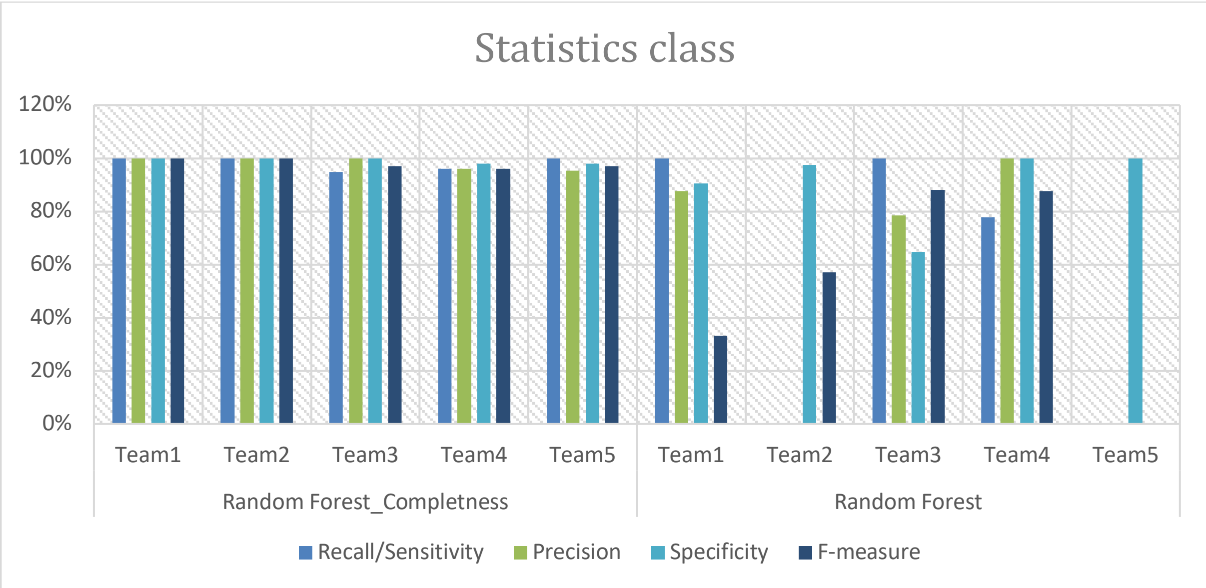


Figure 28. Statistics class of the Random Forest-based completeness and the second Random Forest models

At this stage, we established a comparative study based on the performance of each classifier used in our article to be able to select the most qualified classifier answering our issues. Figure 28 assumes that the Random Forest-based completeness classifier meets our expectations and gives accurate classification results.

4. Discussion

In this contribution, we tried to solve an issue that was expressed within the Ibntofail University-Morocco, it is about the composition of the teams of learners who will participate in the end-of-semester projects. According to the pedagogical description of the module, students must participate at the end of the module project to assess their skills acquired during the semester. It is a common module between three disciplines so the management of interdisciplinary collaboration is necessary. Usually, the composition of the teams is done randomly by the assignment of the teacher, or the students form teams according to their preferred partners. Faced with this problem, we proposed an intelligent method to help form complementary teams that stimulate student participation and performance, as well as to propose a new combination of students in teams allowing democratized learning.

In our approach, we leveraged Educational Data Mining (EDM) to manage this kind of collaboration, in this case, we used two data sources (Moodle log, and Wayuu dataset) to make an intelligent service prototype that will be integrated into the educational platform of the Ibntofail university. Before starting our data mining, we build an ontology to manage the heterogeneity of profiles of learners and disciplines of projects, after we integrated a heuristic layer to ensure the complementarity between the learners who will participate in the projects defined by the teacher, this complementarity is represented below:

A learner(i) may be a complementary element among the completeness_table (compl_tab(j)) of any of a given project(j). This statement can be expressed as follows:

$$\forall\ j \in \{1..m\}$$

$$\text{if } compl\_table(j) \neq \emptyset$$

$$\text{Then } \exists\ i \in \{1..n\}: \quad Learner(i) \quad \in \quad compl\_table\_(1 \leq j \leq m)\ (j)$$

Where:

n: is the number of learners

m: is the number of projects

compl_table(j): each row of this table is a combination of complementary learners depending to project(j) description.

Finally, we have exploited the outputs of these layers as being classes used in predicting new combinations of complementary teams.

Our study evaluates the integration of semantic representation of data related to the learner and the project, as well as the completeness layer to define the most qualified classifier that we will be integrated into our Smart Collaborative Learning Service (SCLS) , for this, we compared various results obtained during our work. Figure 29 presents a comparison of the overall statistics of the three established classifiers.

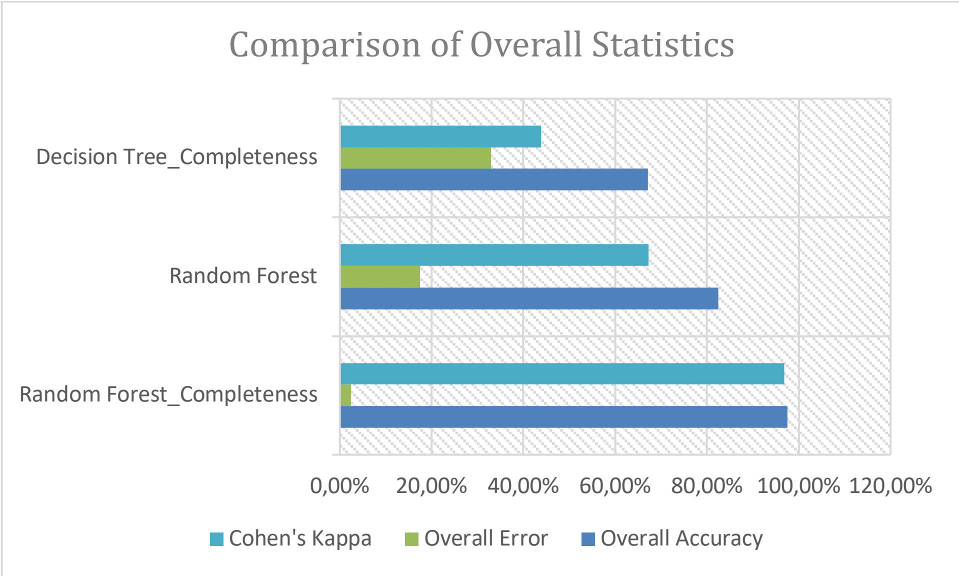


Figure 29. Comparison of overall statistics

Based on the result of Figure 29 the classifier named "Random Forest-based completeness" gives an enhanced Overall accuracy rate by approximately 67.07% to 97.56%, also, it reduced the Overall Error by approximately 32.93% to 2.44%, and improved Cohen's Kappa rate by 43.80% to 96.80% as compared to the other different classifiers. In this regard, we were able to affirm that the "Random Forest-based completeness" classifier is the most appropriate predictive model that will be used to solve the expressed issue of our research, also the integration of semantic data and completeness processing improved the classifier performance and gave the desired results.

Fortunately, the results of "Figure 29" validate the architecture proposed in "Figure 14," which integrates modeling, processing and decision support tools, to develop an intelligent system that predicts the completeness of teams based on the skills and preferences of learners according to the problem-solving designed in their educational program. Also, they help us to propose a smart teaching method consisting to include all students in the teaching process ensuring democratized learning.

5. Conclusion

We proposed a novel approach for the development of a Smart Collaborative Learning Service (SCLS). In general, an intelligent system need proper data representation and processing in order to provide proactive services. In fact, the integration of the ontology and the computation of completeness improved the classification quality and provided more relevant predictions of complementary teams. These experimental results allow for the incorporation of this smart service into collaborative learning platforms in order to develop successful teams and give equitable opportunities for knowledge exchange and acquisition in a Smart University (SU) environment. In the future, we will investigate the implementation of the Smart Collaborative Service in cloud computing, as well as how to use it in a scaled collaborative environment via its REST API.

Chapter V : An architecture for continuous deployment of the Smart Collaborative Learning Service (SCLS) based on a predictive model to build complementary teams

1. Introduction

The Coronavirus pandemic forced the governments of the world to close educational institutions caused 89% (more than 1.5 billion learners) from 188 countries to be forbidden access to educational institutions to receive face-to-face education as the UNESCO report in 2020 (Abusaada and Elshater 2020). At the same time, 60% of HEIs also reported that COVID-19 has increased virtual mobility and/or collaborative online learning as alternatives to physical student mobility (Marioni G. 2020). This situation raises an essential debate about implementing the insight of integrating smart technology in the fields of the university. In fact, there is a range of modern tools available to face the challenge of distance learning imposed by the COVID-19 pandemic (Gonzalez et al. 2020).

Now, it is clear that the outbreak situation due to covid 19 hurried the transition from the traditional university to a digital university to ensure sustainable learning and overcome the constraints of changes in the educational environment that require distance learning. Indeed, the expansion of the university in a virtual environment is a major challenge to achieve, it consists of digitizing the business, interactions, content, and information flow within the university. However, this digitization is insufficient to produce an environment that supports the learner during his educational curriculum, which leads to considering an evolutionary process based on the integration of new technologies and decision tools enabling a smooth migration to a university that offers personalized and proactive services to students: it’s about the emerging term “Smart University”. Therefore, the Smart University (SU) is defined as an enhanced model of the classical university able to deliver services like education, teaching, research and training with high performance to improve and modernize the quality of learning. The term "Smart University" is divided into two keywords that specify the technological, methodological and functional scopes that achieve smartness in the academic environment. Starting by the term "SMART", it is used as an acronym referring to interactive technology that offers a more flexible and tailored approach to meet diverse individual requirements by being "Sensitive, Manageable, Adaptable, Responsive and Timely" to educators' pedagogical strategies and learners' educational and social needs (Gomede et al. 2018). Also, a smart system implements significant maturity at various "smartness" levels as shown in " Figure 5. Smartness levels in the Smart system" (O. Akhrif, El Idrissi, and Hmina 2018).

On the other hand, the term "University" is related to a social institution and an educational process for higher education is a space in which both students and academics could freely pursue their intellectual interests. There are many functions of a university: the knowledge-producing university; the entrepreneurial university; the commodified university; the university as a place of learning; the moral university; the critical university; the philosophical university; the university of wisdom; the university of dissensus; the eudaemonic university; the metaphysical university; the concerned university, and the translucent university (Burwood 2020).

The Smart University (SU) is the involvement of innovative technologies such as Artificial Intelligence (AI) and cloud computing to optimize and improve the learning process of the traditional university. The use of technology in the university gives rise to certain concepts that accompany the emergence of the Smart University, namely: Smart learner, Smart learning, Smart Knowledge, and Smart Interactions.

Our approach focuses on the study and implementation of the concept of Smart Interactions and more precisely Smart Collaboration considering it the pedagogical method to intelligently share and acquire knowledge within the Smart University (SU) environment. Smart Collaboration consists of involving the learners in tutored projects, it is called upon to achieve the maximum success of a project, the success of a project relies mainly on a homogeneous and complementary team. Indeed, a project requires a set of criteria to be respected, namely: the skills of the learners required (profiles), the number of people per profile and the expected objectives. On the other hand, assigning learners to these projects present a complex task, as it is difficult to compare students whose profile is represented by several different qualities and values. To accomplish this task, we have developed an ontology-based approach to be exploited by a machine learning algorithm that will optimize team completeness and prediction.

In this chapter, we propose to deploy this approach as a service. The added value is synchronizing the composition or updating of teams with the changes that will occur in the future (Proactivity). In other words, each time a learner has an assessment or completes a project, this profile is updated, which will lead to a modification of the teams. This vision will allow much more accuracy when starting a continuous completeness building team classifier.

This study is divided into six parts: In section 2, we present the overview approach and Smart Collaborative Learning architecture, In Section 3, we describe the methods allowing the service deployment in Amazon Web Service, In Section 4, we present in detail the proposed system, the experimentation and the results obtained. Finally, we closed with a discussion.

2. Smart Collaborative Learning Service (SCLS) requirements

This approach is an extension of the work already published (Ouidad Akhrif, El, and El 2021), It intends to deploy the Smart Collaborative Learning Service (SCLS) , which may be used as a decision support layer for collaborative platforms in order to manage interdisciplinary collaboration inside the Smart University. It enables an optimal management of student assignments to complementary teams, ensuring equitable participation and knowledge exchange for all students. Furthermore, this service addresses the limits associated with team composition, which is carried out randomly by the teacher or freely chosen by the student who selects his preferred teammates. Indeed, embedding this service into a collaborative platform is seen as pedagogical help for the teachers, allowing him to properly assess students' skills and apply collaborative learning practices. On the other side, it enables students to collaborate with other profiles in order to learn new aibilities and to integrate into collaborative work.

The Smart Collaborative Learning Service (SCLS) is a representation, analysis, and data processing process that is specifically designed to develop a predictive model that builds complementary teams of learners relating to the collaborative project. This process can be divided into two parts: the semantic modeling and representation of data related to learners and the project to manage the heterogeneity of learner profiles and project disciplines in order to achieve an optimal selection of students who will be assigned to the collaborative project; The completeness was calculated by implementing a heuristic based on statistical and sorting studies to select complementary students related to the collaborative project; this heuristic layer is the entry point of the prediction layer, through which we defined prediction classes that will be used by a predicting model based on Classification and Regression Trees (CART) as a machine learning method instance; Ensuring continuous deployment of the Smart Collaborative Learning Service (SCLS) is linked to choosing the highest qualified predictor at all times and after each data layer modification. Of course, these changes can have an impact on the quality of the classifier. To achieve this goal, we integrated three types of classifiers at the predictive layer level of this smart service, then automated the selection of the classifier that will predict complementary teams based on its highest accuracy. Using this mechanism, we were able to obtain the best prediction during the smart service's deployment.

The Smart Collaborative Learning Service (SCLS) paradigm comprises data modeling and processing layers, as shown in the architecture below:

Figure 30 present the overall architecture designed to develop intelligent collaborative learning:

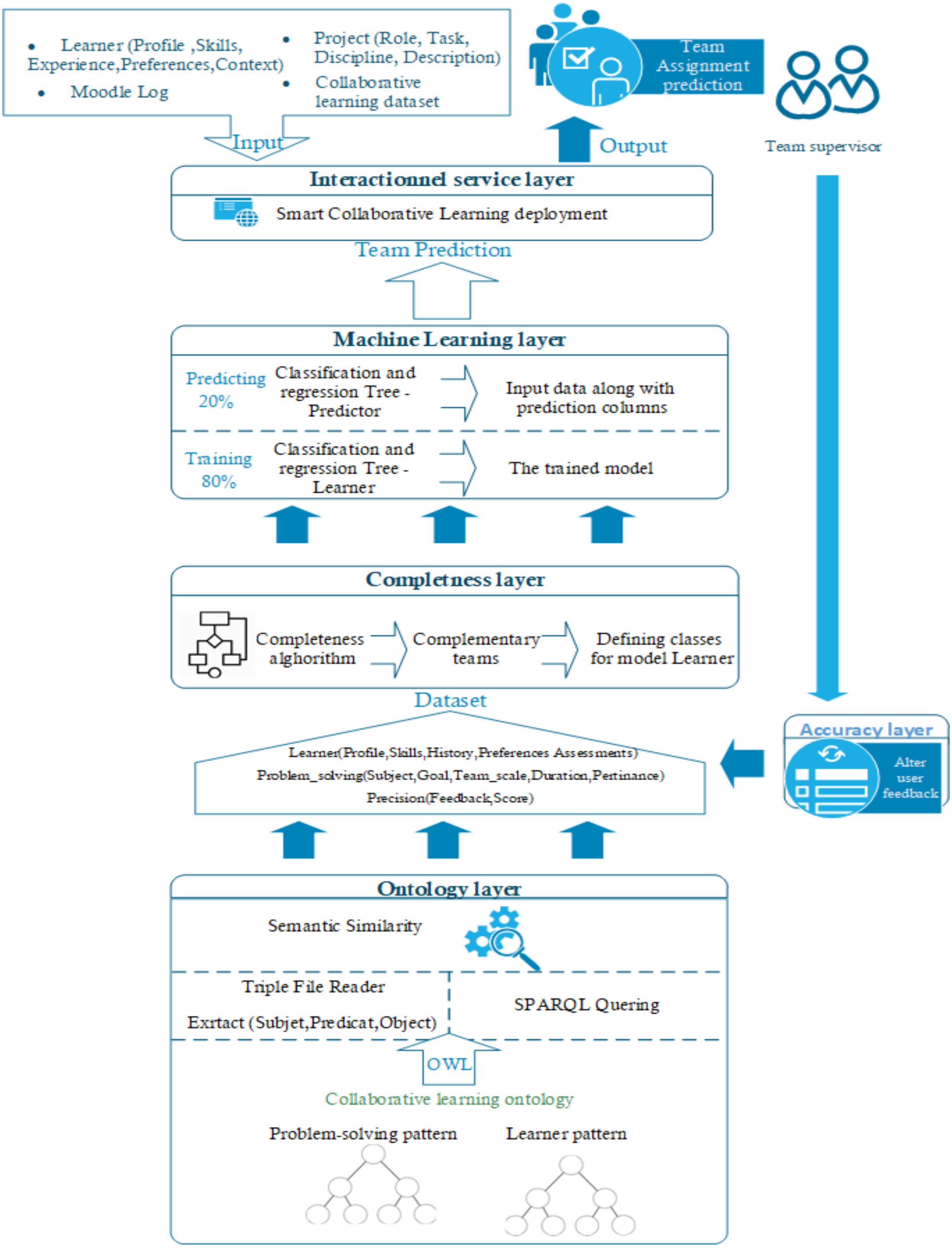


Figure 30. Smart Collaborative Learning Architecture (Based on (Ouidad Akhrif, El, and El 2021))

The Smart Collaborative Learning Service (SCLS) architecture depicts the method we took in developing this smart service, beginning with (1) developing an ontology to manage heterogeneous learner profiles and interdisciplinary collaborations; (2) compute complementarity for each student in comparison to all students who will be working on problem-solving to build complementary teams; (3) develop CART classifiers that predicts new complementary teams using a supervised machine learning technique; (4) Integrate this classifier into the Smart Collaborative Learning Service (SCLS); and (5) deploy the Smart service in the cloud computing environment using KNIME Server and Amazon API Gateway to assure this reuse via a REST API.

We studied the qualities of an intelligent service that can be incorporated into the university environment when deploying and delivering the Smart Collaborative Learning Service (SCLS), thus we conformed to the intelligence criteria listed below. (Akhri, El Bouzekri El Idrissi, and Hmina 2019):

- **Proactive:** By deploying a continuous team-building predictor model that listens to the student profile updates and predicts complementary teams, the service assists the teacher proactively in building complementary teams of learners while working on a collaborative project.
- **Ubiquitous:** This model attempts to give the smart service in ubiquitous accessibility in order to make it available to everyone and everywhere, as well as to ensure its on-demand availability across several platforms.
- **Sustainable:** We attached great importance to guaranteeing continual interactions between the smart service and its users, along with environmental elements.
- **Self-learned:** The smart service listens to all semantic data layer updates and adapts to data changes in a synchronized manner.
- **Secured:** The service is deployed in a secure architecture that controls the smart service's confidentiality and accessibility.

The criteria established for developing the smart service lead us to design a deployment architecture that supports the on-demand execution of the REST API, which is dependent on data updating.

3. Prerequisites for deployment

In practice, the continuous deployment of the Smart Collaborative Learning Service (SCLS) requires data analysis technology that: (1) allows the integration of an ontology as a data sources; (2) provides a custom snippets to add layers for heuristic completeness computation (3) builds classification models using machine learning algorithms for each data layer change; and (4) Supervises for changes in the data layer and automates the selection of the best qualified classifier that predicts complementary teams. The KNIME Analytics Platform addresses these issues by providing workflow systems as a platform for connecting tools that focus on data transfer and flow control integration, allowing third-party developers to easily embed their tools and make them interoperable with each other, independent of their respective domain. (Fillbrunn et al. 2017).

3.1. KNIME Server

KNIME Server complements the KNIME Analytics Platform and is the enterprise software for team-based collaboration, automation, management, and deployment of data science workflows as analytical applications and services. KNIME Server enables all stakeholders in the realm of data practice to work together on a single platform: from data engineers and data scientists to business users and domain experts, as well as models, IT, practice leaders and management. KNIME Server deploys KNIME Analytics Platform workflows and produces their data science applications and services as shown in Figure 31:

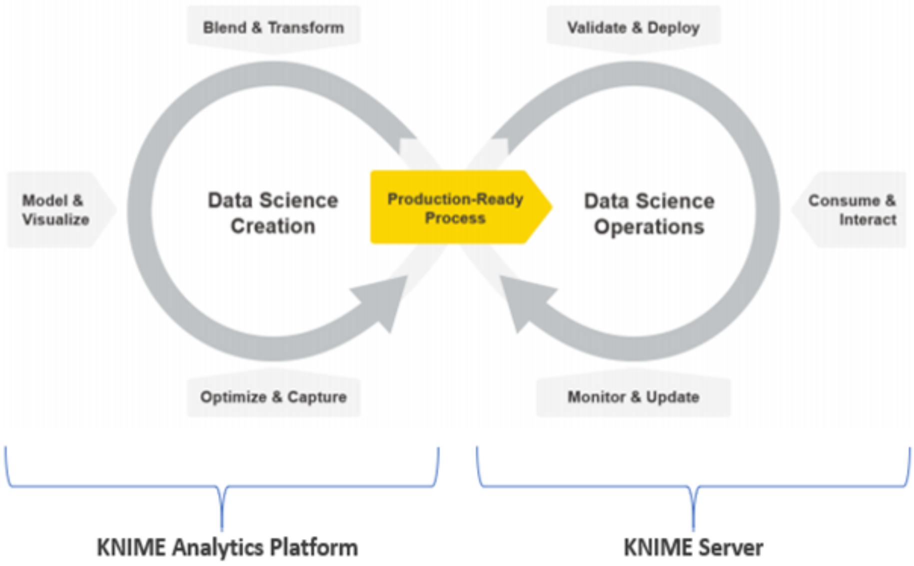


Figure 31. KNIME Server functionalities blocs

KNIME Server consists of two components: a server-side component and a client interface. The server-side component is installed on an application server (Tomcat) together with a KNIME Analytics Platform installation which will be responsible for executing workflows. The latter is referred to as "Executor". The client interface allows interaction with the server repository and is integrated into the KNIME Analytics Platform (via the KNIME ServerSpace extension), and KNIME WebPortal, or accessible via the REST API.

3.2. KNIME Integrated Deployment

Integrated Deployment eliminates the gap between the creation and production of data science by creating and deploying production workflows automatically - without manual intervention. It's possible to identify what's necessary for production and not just the model. All nodes and settings are captured to ensure the workflow always remains in sync.

Deploying a production workflow to KNIME Server is available on KNIME Server, Microsoft Azure and Amazon web service marketplace. To choose the most suitable environment for the deployment of the Smart Collaborative Learning Service (SCLS) , we have drawn up a comparative table to designate the appropriate environment for deploying the smart service.

Table 10 . Comparison of compatible cloud environment

Features	Microsoft Azure	Amazon Web Service	Knime Server
KNIME Server Small - 4.12.2		*	*
Knime Analytics platform executor - 4.3.1		*	*
OntoPortal		*	
Environmental prerequisites	*	*	
Trial version			*

The KNIME server offers a 3-month trial edition but has rigorous environmental requirements. We were unable to launch our smart service on Microsoft Azure since its environment was incompatible with the process we had built. All of these factors lead us to conclude that the smart service deployment environment is a cloud computing infrastructure, specifically the Amazon Web Service (AWS). Amazon Web Service (AWS) provides appropriate capabilities for implementing the Smart Collaborative Learning Service (SCLS) and allows for the environmental requirements to be built for KNIME Server as a service. KNIME Server Small for AWS requires a premium **EC2** instance type (**r5.2xlarge**) to run the server.

3.3. Cloud Computing

Cloud computing has become a great solution for providing a flexible, on-demand, and dynamically scalable computing infrastructure for many applications. IT is the IT foundation for cloud services and it consists of technologies that enable cloud services (Furht 2010). The term service in cloud computing is the concept of being able to use reusable, fine-grained components across a vendor's network. This is widely known as "as a service" (Velte, Velte, and Elsenpeter 2010). Cloud computing implements services at four distinct levels: SaaS, platform as a service (PaaS), and infrastructure as a service (IaaS) and Network as a service (NaaS) (Gavrilović and Mishra 2021).

Cloud service platform supports online collaboration, file storing, virtualization and flexible visit. These technologies support to use of "seamless connection", which is the basic guarantee to implement "learning at any time, anywhere, in any way and at any place (4A)" (Liu and

Huang 2017). In recent years, the Amazon Web Services (AWS) cloud platform has released several services that support machine learning (ML) and artificial intelligence (AI) capabilities to enable developers and data scientists to quickly and easily build, train, and deploy machine learning models at any scale (Varia and Mathew 2014).

3.4. Amazon Web Service (AWS)

In 2006, Amazon Web Services (AWS) began offering IT infrastructure services to businesses in the form of web services. Amazon Web Services (AWS) is a provider of cloud services, meaning on-demand access to IT resources via the Internet (Varia and Mathew 2014). In general, a deployment as per the KNIME Server Small for AWS guidelines looks something like the architecture in Figure 32 :

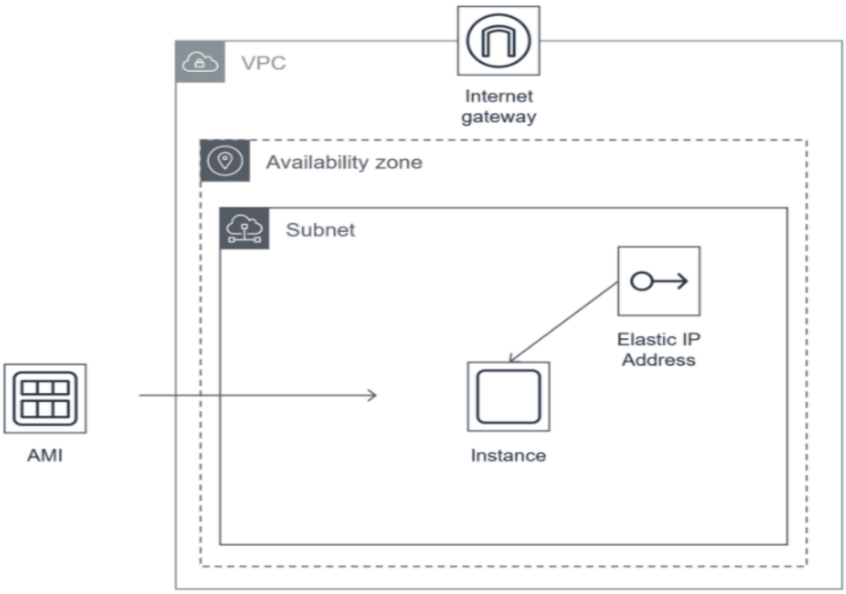


Figure 32. KNIME Server Small deployment on AWS

3.4.1. Amazon Machine Image (AMI)

An encrypted machine image that contains all information necessary to boot instances of your software. Using simple web service interfaces, users can launch, run, monitor and terminate their instances. Moreover, they can, on the fly, add any of the features to their configuration as they desire (Furht 2010).

3.4.2. Amazon Elastic Compute Cloud (EC2):

Amazon Elastic Compute Cloud (Amazon EC2) provides scalable computing capacity in the Amazon Web Services (AWS) Cloud. Using Amazon EC2 eliminates your need to invest in hardware upfront, so you can develop and deploy applications faster (Varia and Mathew 2014). You can use Amazon EC2 to launch as many or as few virtual servers as you need, configure security and networking, and manage storage. (Furht 2010).

3.4.3. Amazon Elastic Block Store (EBS):

Amazon Elastic Block Store (Amazon EBS) provides block-level storage volumes for use with EC2 instances. EBS volumes behave like raw, unformatted block devices. You can mount these volumes as devices on your instances. EBS volumes that are attached to an instance are exposed as storage volumes that persist independently from the life of the instance (Velte, and Elsenpeter 2010).

3.4.4. Amazon Virtual Private Cloud (VPC):

Amazon Virtual Private Cloud (VPC) is a virtual network that we create so that we can segregate certain things from the entire user domain. It acts just like a normal cloud, but instead of having separate infrastructures, there is only one cloud infrastructure but multiple virtual clouds made over it (Varia and Mathew 2014).

3.4.5. Amazon API Gateway:

Amazon API Gateway can be used to create, publish, maintain, monitor, and secure different kinds of APIs such as REST, HTTP, or the WebSocket API. These APIs can be made not only to have their applications but can also access and use different AWS services (Singh 2021).

At this point we quoted the prerequisites for installing the KNIME server on AWS, we need to define the methods used to capture the predictive model data and synchronize the model prediction when modifying the ontology layer.

4. Proposed architecture/ Implementation

In this section, we discussed the processes of the experiment that led to the continuous deployment of the Smart Collaborative Learning Service (SCLS). Moving a continuous intelligent service based on data science to production remains a significant issue; as a result, we designed a deployment architecture on an AWS environment to provide a predictive model as a service for building complementing teams. Indeed, the AWS environment provides the necessary cloud computing services for deploying the various layers that comprise the smart service. It also includes a service that supports synchronized interactions between the ontology layer and the machine learning layer, as well as a service for building a REST API for integrating the smart service into third-party platforms. The smart Service deployment follows the architecture presented in Figure 33:

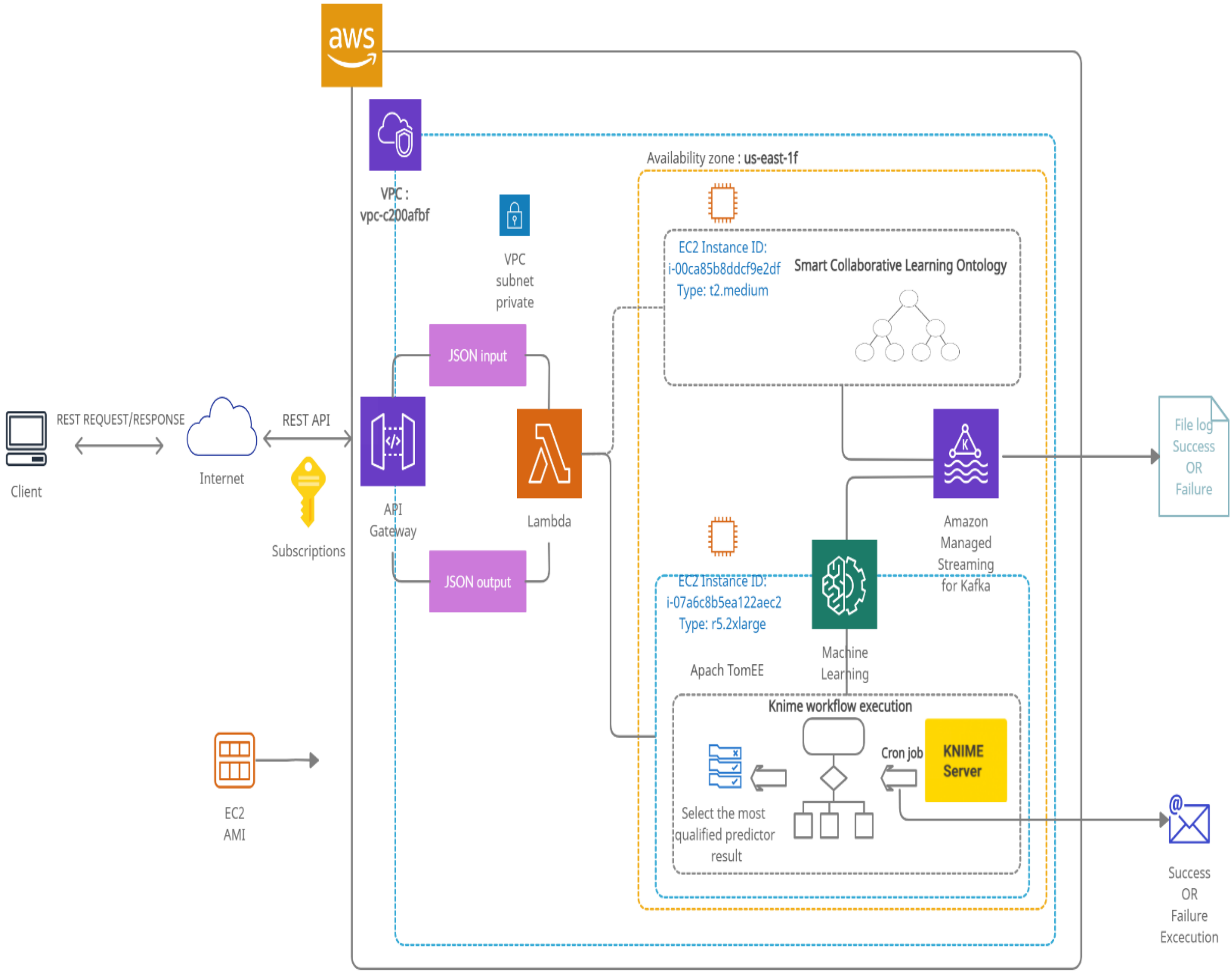


Figure 33. Smart Collaborative Learning Service (SCLS) deployment on AWS

The complete system can be managed and modified from the modeling workflow,thanks to Integrated Deployment.

4.1. Smart Collaborative learning ontology deployment

On the Amazon Web Service cloud, an Amazon Machine Instance (AMI) is available to deploy ontologies on OntoPortal Virtual appliance, is a copy of the BioPortal ontology repository software that allows managing properly formatted semantic content in a web service, and performing various analytic tasks with that content. Figure 34 shows the details of deploying the Smart Collaborative Learning ontology in the AWS environment:

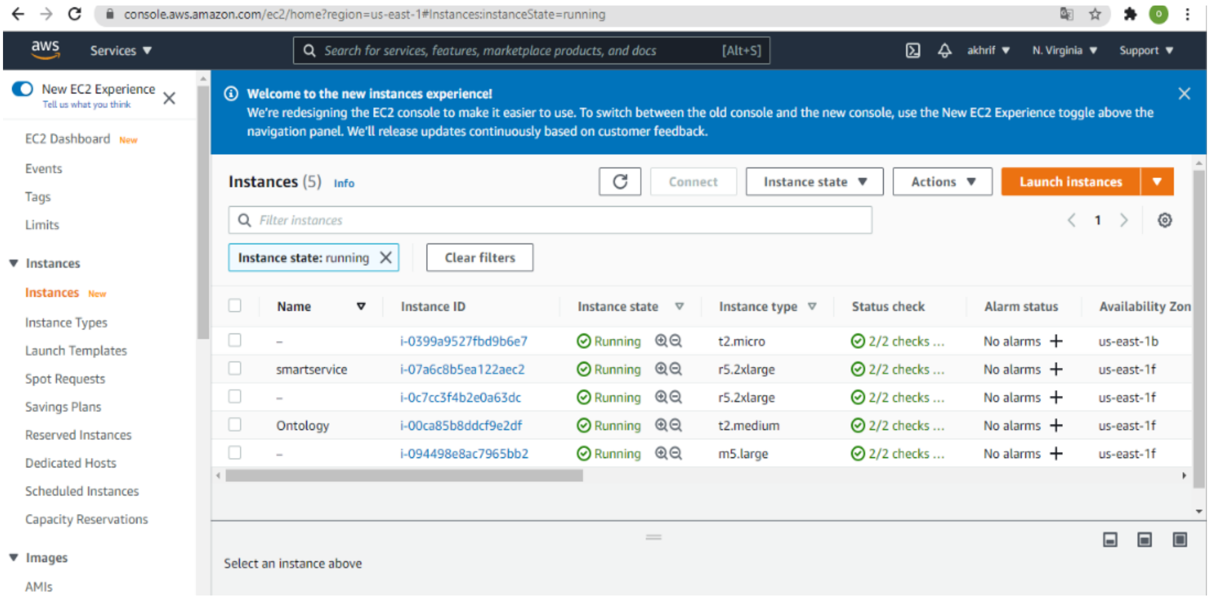


Figure 34. Running OntoPortal instance on AWS

Figure **34** gives current software and infrastructure informations, including the instance name **(Ontology)** and the ID (**i-00ca85b8ddcf9e2df**), instance type (**t2.medium**) , and Availability Zone (**us-east-1f**).

The OntoPortal Virtual appliance is a web-based application and API for accessing and sharing ontologies, we can access the deployed ontology via the Ontoportal client as shown in Figure 35:

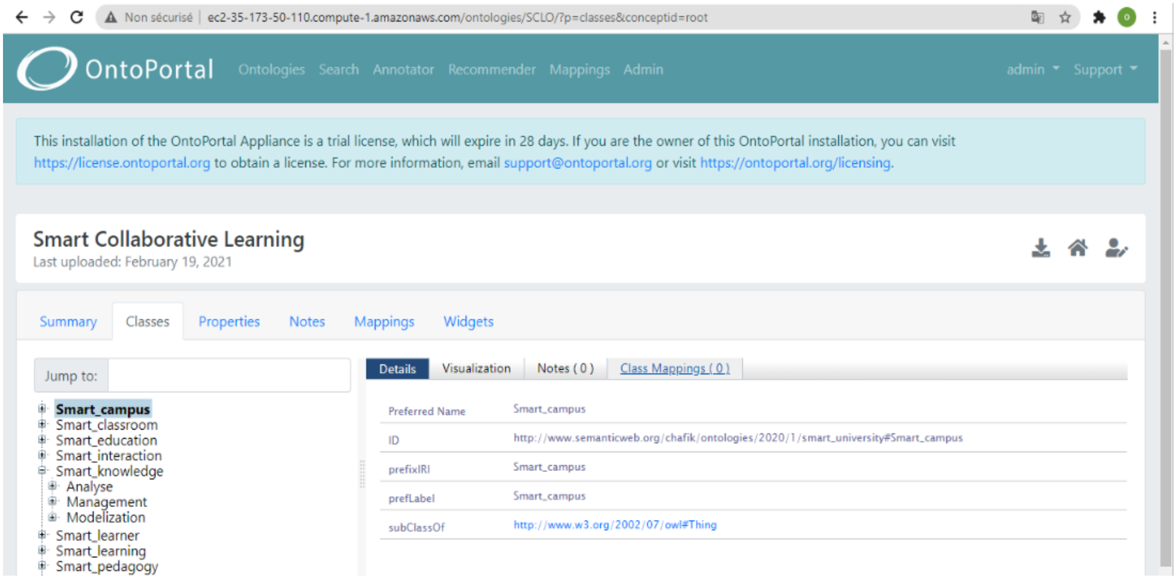


Figure 35. Smart Collaborative Learning Ontology

Figure 35 shows the different classes that constitute our ontology at the level of Ontoportal Appliance, at this stage our ontology is ready to be exploited by other components of the architecture of Figure 33.

4.2. Amazon Managed Streaming for Kafka (Amazon MSK)

The Apache Kafka platform is used by Amazon MSK, a data streaming service. Apache Kafka is a free and open-source platform for developing pipelines and serving real-time data applications. You may utilize Amazon MSK to populate data lakes, continually push database updates, and power analytics and machine learning applications by using native Apache Kafka APIs. In our study, we used Amazon MSK to listen for changes in the ontology layer, which triggered the execution of machine learning layer processes and generated predictive models that were synchronized with changes in the ontology layer. This phase enabled us to automate the Smart Collaborative Learning Service (SCLS) invocation and continually adapt it to data changes.

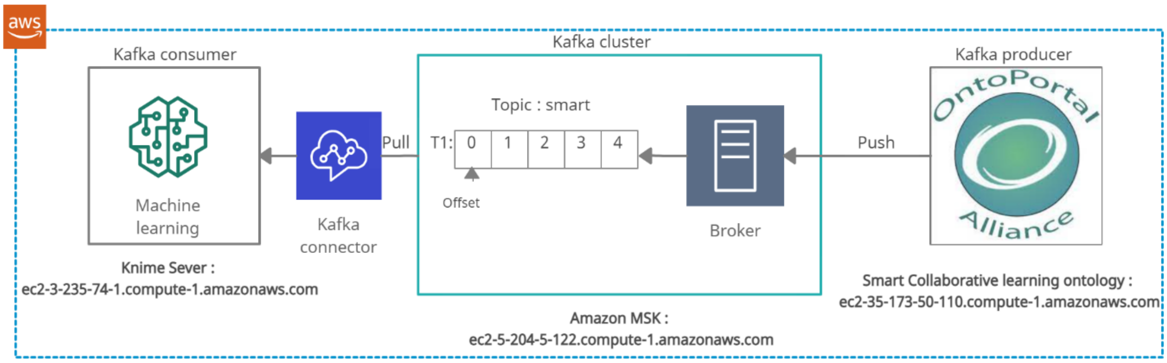
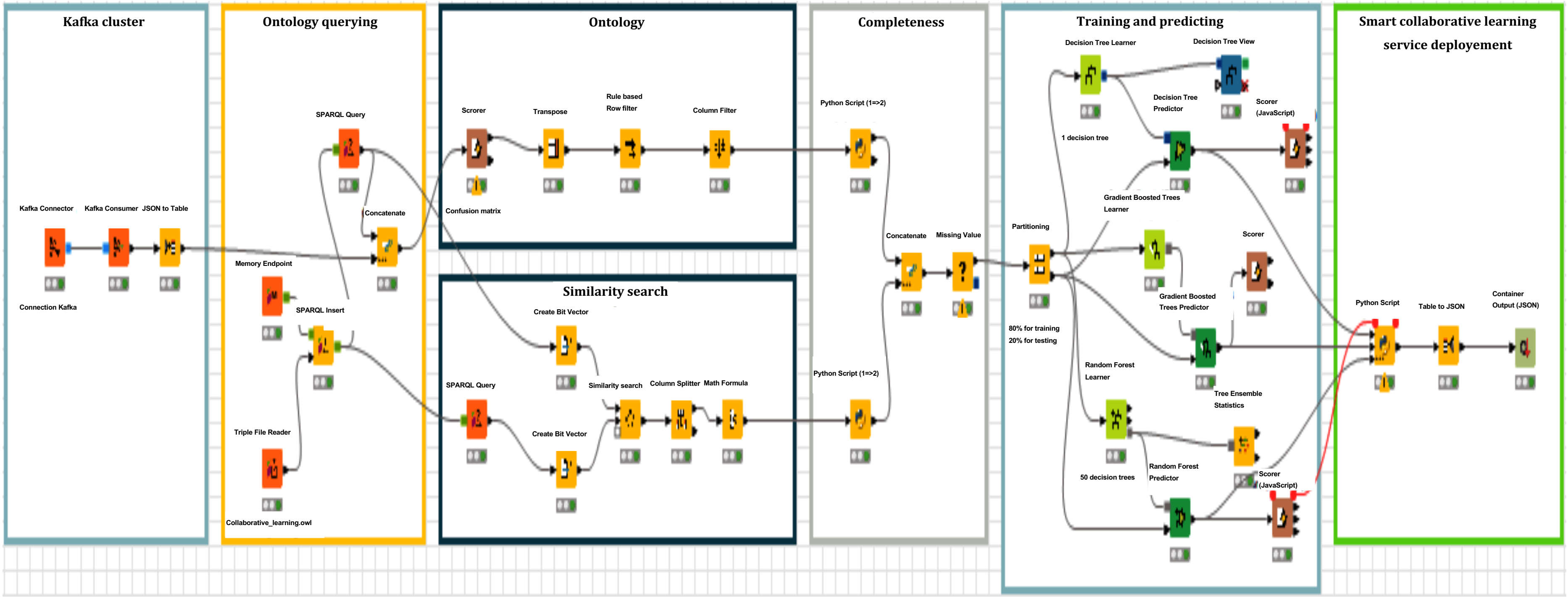


Figure 36. Apache Kafka architecture and components on AWS

We connected Amazon MSK between the ontology layer (Kafka producer) and the machine learning layer (Kafka consumer) to push the ontology updates in order to establish a stream listening to ontology layer changes. The machine learning layer connects to the Kafka cluster through a Kafka connector and pulls the needed data from the topic (smart). We built the machine learning predictors using the KNIME Analytics platform, which linked to the Kafka cluster through the KNIME Kafka connector and consumed the topic (smart) in JSON format.

4.3. Predictive model deployment

In this part, we looked at how to deploy a predictive model within a KNIME workflow that supports REST API calls, as well as how to publish workflows to the KNIME server. The KNIME Analytics Platform provides sophisticated prediction models that enable the use of machine learning methods. At the practical level, we build a workflow using the KNIME analysis platform: The first part of this workflow connects and reads ontology changes from the Kafka cluster, the second attempts to extract knowledge from the Smart Collaborative Learning ontology, the third prepares the data that will be used in the completeness processing part, then the result data will be used in the classification and prediction part, the results of the last steps is transmitted to the container (output) to enable JSON output for the REST service. The KNIME workflow in Figure 37 demonstrates a full lifecycle of the Smart Collaborative Learning Service (SCLS) that focuses on the application of machine learning techniques for classification and prediction in building complementary teams in an academic environment. Also, it shows the steps to create a REST-based web service that can be deployed via the KNIME Server.



Not executed Executed

Figure 37. Smart Collaborative Learning Service (SCLS) workflow

This workflow demonstrates the main steps in deploying the Smart Collaborative Learning Service (SCLS) , including querying and preprocessing the data, building the predictive model, selecting the most qualified predictor based on its highest accuracy, and producing result data via a REST API which can be deployed via the KNIME Server.

The smart collaborative learning workflow is divided into three main steps:

- Access and read data from the Kafka cluster: this step receives real-time students information updates from the Kafka cluster to ensure optimal students affectation to complementary teams. Figure 38 presents the Kafka cluster connection via the KNIME analytic platform.

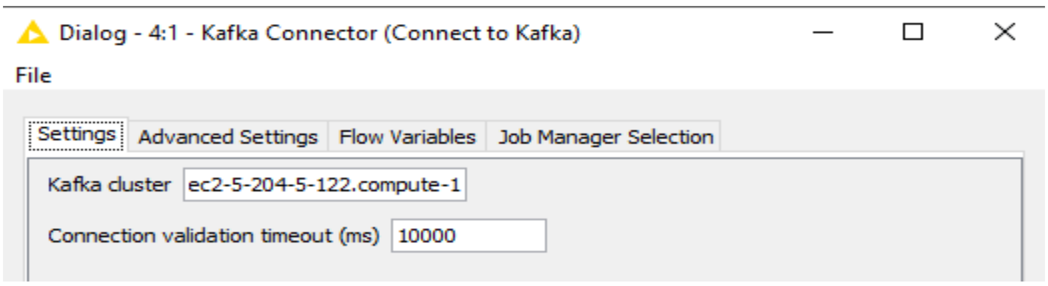


Figure 38. Kafka cluster connection

- Smart Collaborative learning ontology: this part is the data source that we operated to produce combinations of complementary teams, we have used a semantic data source because we managed interdisciplinary collaboration and heterogeneous student's profile, for that, we integrated the semantic similarity node to this workflow for expanding the research field of similar student's profile, after begging the completeness processing, we used the data received from the Kafka cluster to update learners profile that we have in the collaborative-learning.owl file.
- Completeness processing: this part is based on the implementation of a heuristic that we presented in previous work, allowing completeness processing of student skills as per the required skills by the problem-solving.

- Classification and prediction: The Smart Collaborative Leaning service must encapsulate a continuous and performant classifier. To achieve that, we train and test three classification and regression tree (CART) models, then, we select the most qualified model based on these accuracy values and automatically deploy it. This workflow trains the Random Forest, the Decision Tree and the Gradient Boosted trees models then selects the model that gives a higher Overall Accuracy using the Python Script node.

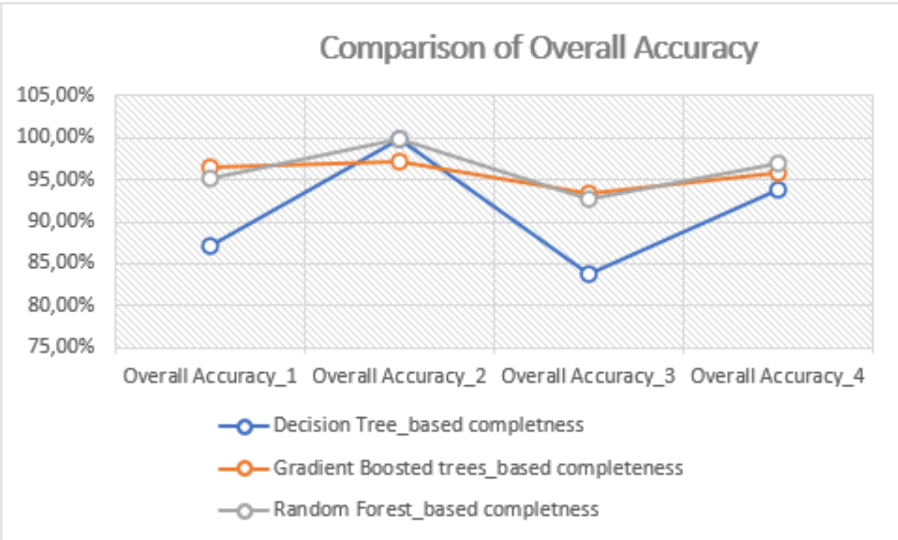


Figure 39. Comparison of Overall Accuracy

Figure 39 displays the result of the Overall Accuracy tests after four changes to the data source. Indeed, the change of data can affect the quality of the predictor which is related to the learner skills and can therefore modify the number of classes used by the training phase of the model.

- Deploy the workflow on KNIME Server: our research is dedicated to de workflow deployment as we will detail it in the next section.

4.3.1. Testing the deployment

Running KNIME Server Small on Amazon Web Services (AWS) requires a single AMI instance and launching this instance requires a VPC and subnet as well as another prerequisite AWS service as EBS volume and data transfer in/ out. Figure 40 shows the KNIME server Small instance information called “**smart service**”.

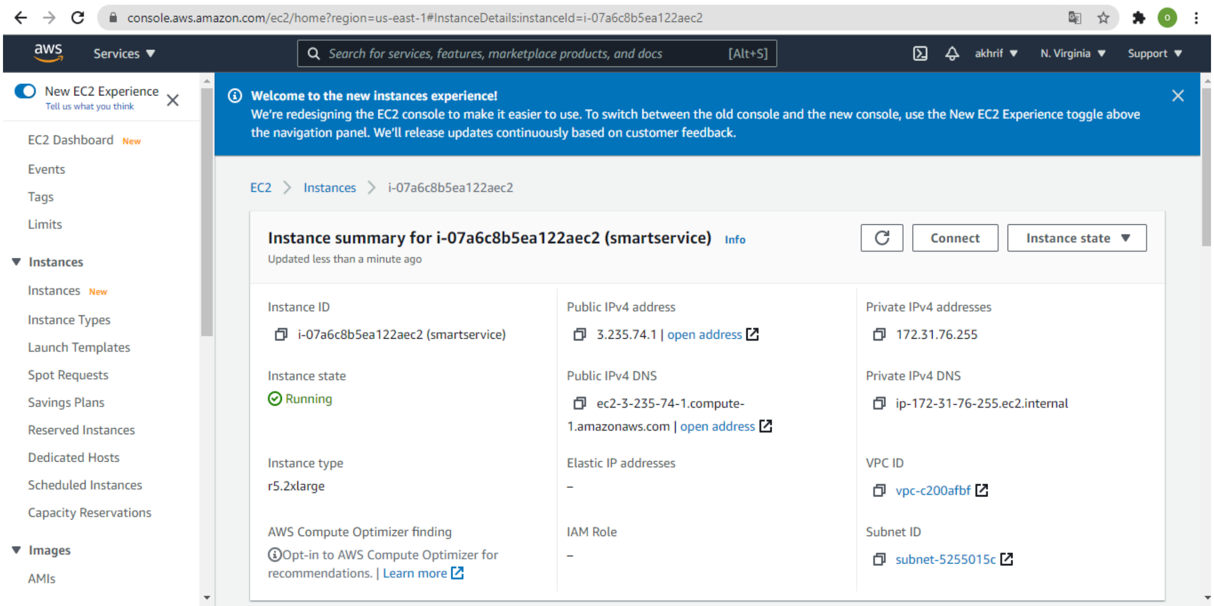


Figure 40. KNIME Server Small instance

Once the server is ready, we can deploy the workflow through the KNIME Analytics platform. we need to create the mount point address: **http://ec2-3-235-74-1.compute-1.amazonaws.com / knime** at the KNIME analytics platform to deploy the workflow, Figure 41 shows the deployment step to the KNIME server:

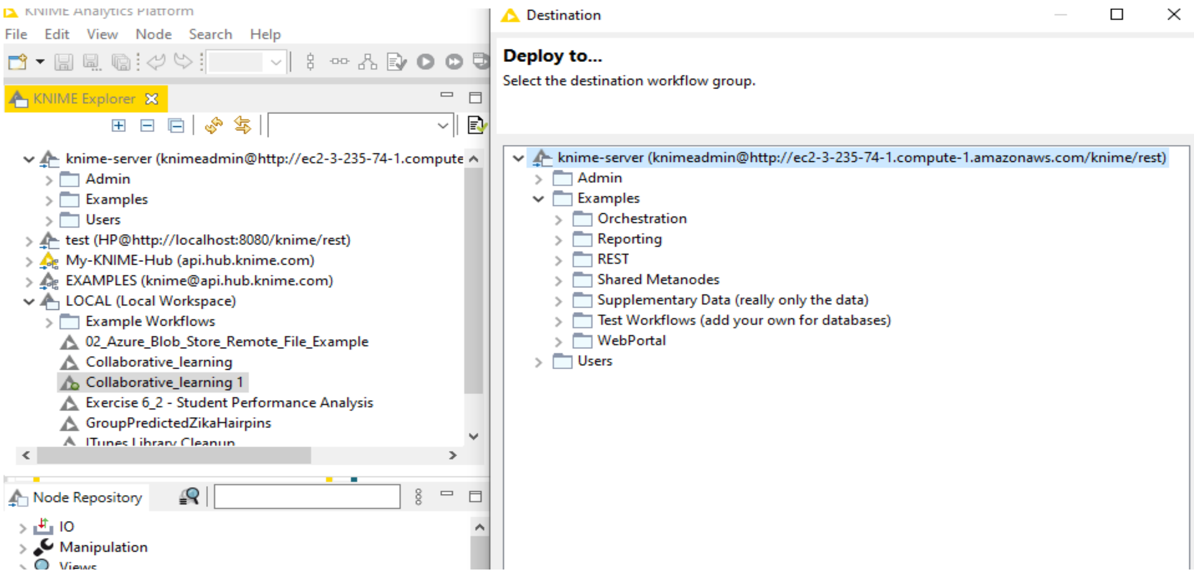


Figure 41. the workflow deployment

Testing the deployment can be performed by logging into KNIME Server WebPortal through the web browser. Once we run the KNIME server AMI, the resulting instance will automatically start the KNIME server. In our case, the KNIME server WebPortal is available in the browser at: **http://ec2-3-235-74-1.compute-1.amazonaws.com / knime**.

Logging to the KNIME server WebPortal request:

User: Knime .

Password: the MAC address of the AWS virtual machine.

The access to the KNIME Server WebPortal allows to view and run the workflow deployed on the KNIME Server, Figure 42 shows the workflow we have deployed called: **Collaborative_learning 1**.

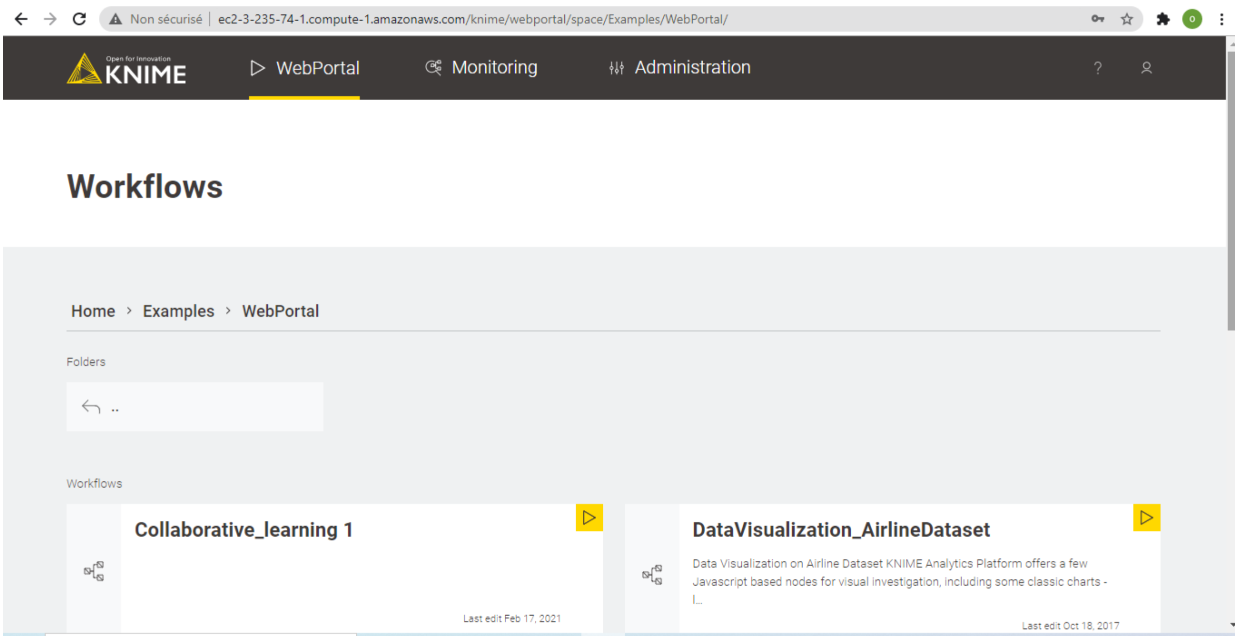


Figure 42. Smart Collaborative Learning Service (SCLS) workflows on WebPortal

The access to the KNIME Server from the KNIME Analytics Platform is via the KNIME Explorer, by accessing the mount point address: [http://<public-hostname>/knime]

4.4. Generating and testing the REST API

The Smart Collaborative Learning Service (SCLS) deployment generates a REST API that may be used by external applications. Knowing that the adoption of REST can lead to a simple, scalable, effective, safe and reliable architecture (X. Chen et al. 2017). KNIME Server provides a Swagger interface to generate the REST Endpoints, which makes finding and using REST services simple. Access to the Swagger definition of the Smart Collaborative Learning Service (SCLS) in the KNIME Server shows the REST API as bellow:

http://ec2-3-235-74-1.compute-1.amazonaws.com/knime/rest/v4/repository/Examples/WebPortal/Collaborative_learning%201:openapi

We used the KNIME Analytics platform to invoke this REST endpoint to evaluate the Smart Collaborative Learning Service (SCLS) call and functionalities. Wherefore, we build a KNIME workflow as is illustrated in Figure 43:

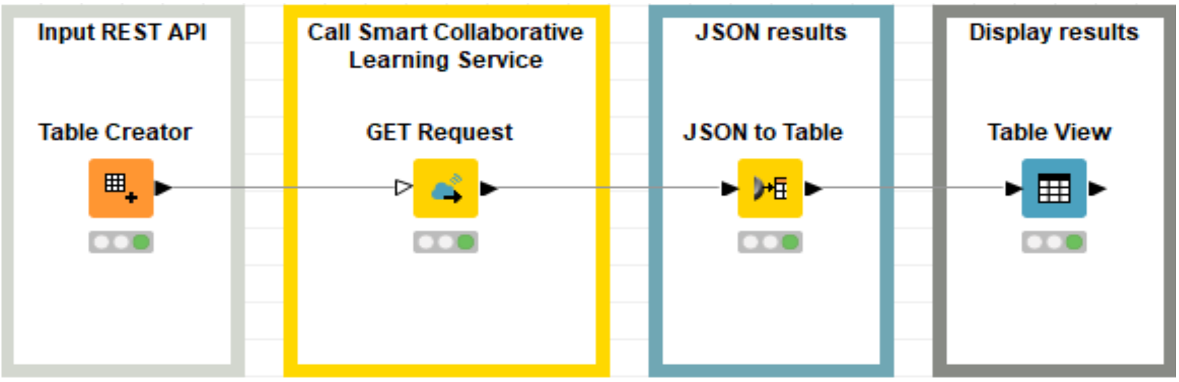


Figure 43. Calling the Restful smart service

This workflow accepts the REST API URL as input. Then the GET request node retrieves data from the smart service and outputs the results in a JSON format. We can execute the JSON Output node and open its view in order to see what the JSON results as Figure 44:

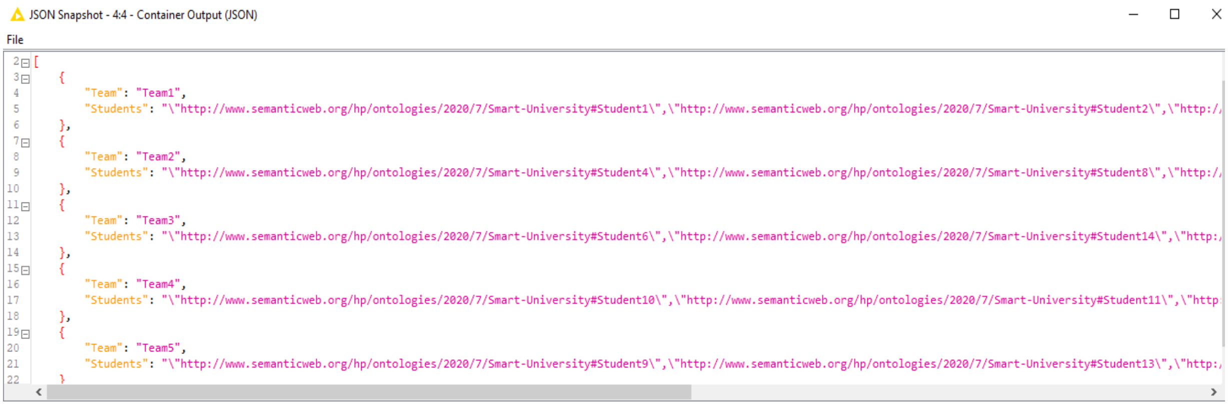


Figure 44. JSON results

In this section, we have developed a workflow for accessing the Smart Collaborative Learning Service (SCLS) from within the KNIME Analytics platform. Then, we detailed this workflow and showed how to call this service and fetch data in a JSON format. This phase demonstrates that we have accomplished the production of the smart service and made it available for third-party integration.

5. Synthesis

In this chapter, we covered a continuous Smart Collaborative Learning Service (SCLS) deployment, which includes a set of layers to enable smartly constructing complementing teams of learners who participate in collaborative learning projects. This smart service's functional role is to handle multidisciplinary collaboration amongst complementary teams based on a project. Solving this constraint, led us to plan for several parameters in the design of the smart service architecture. Starting with the data layer, this layer requires a semantic presentation of the learner profile and the project disciplines, which incites us to integrate the ontology layer into the smart service architecture to perform relevant research thanks to the semantic similarity. As for the processing layer, we integrated a post-processing layer called the “heuristic layer”, to calculate the complementarity between the students in terms of skills required by the project, this layer entrusts the students' assignment to complementary teams. At the next step, we used these assignments-based completeness outputs to define the classes used in predicting new combinations of complementary teams. This prediction was performed using machine learning algorithms namely the Random Forest, the Decision Tree and the Gradient Boosted Trees. To achieve that, we established a KNIME (Konstanz Information

Miner) workflow that integrates the main steps of the Smart Collaborative Learning Service (SCLS) architecture (Figure 30) to build dynamic classification that changes after each step of the project or data updating, this classification is based on the synchronous selection of the most qualified predictive model which is evaluated by its higher accuracy (Figure 39). Once we built the predictive model, we moved on to the process of deploying the smart services on the KNIME Small Server. As we mentioned in (Table 10), the AWS architecture is well suitable for deploying this smart service because it contains the prerequisites to mount the requested version of the KNIME server as well as to deploy the collaborative learning ontology.

On AWS, we want to move the predictive model into production and ensure its continuous deployment as a smart service, for that we tried to respect the criteria of smartness that we quoted in section "Smart Collaborative Learning Service (SCLS) requirements", Table 11 shows the smartness criteria and the experiments that we realize to achieve each of them:

Table 11. Smartness evaluation

Smartness Criteria	**Experiments**
Proactive	The service encapsulates Classification and Regression tree classifiers making it possible to predict the complementary teams and to propose the combinations of the students adapted to problem-solving required skills.
Ubiquitous	The smart service is deployed via a REST API interface to guarantee its portability and on-demand in different platforms.
Sustainable	By using the Amazon Web Services (AWS) Cloud architecture, we ensure continuous deployment of the smart service that is related to use the most qualified predictor at all times and after each change of data layer. We have integrated three types of classifiers at the predictive layer level of the smart service then we automated the selection of the classifier which will predict complementary teams according to its highest accuracy, thanks to

	this mechanism, we were able to obtain the best prediction during the deployment of this smart service.
Self_learner	We used the Submit/Publish concept to be on the lookout for a change of the ontology layer. We integrated Amazon MSK between the ontology layer (Kafka producer) and the machine learning layer (Kafka consumer) to push the ontology changes.
Secured	The Amazon API Gateway takes care of all the tasks required such as traffic management, security, monitoring, and version and environment changes.

The Smart Collaborative Learning Service (SCLS) overcomes the problem of randomized team formation; this type of team formation is insufficient to foster good collaboration among colleagues, which is primarily dependent on information exchange. This smart service serves as a catalyst for the adoption of effective collaboration practices in a complementary university setting.

6. Conclusion

In this paper, we investigated the implementation of the Smart Collaborative Learning Service (SCLS) in cloud architecture, notably Amazon Web Service (AWS), in order to make it available on demand to third-party collaborative platforms via its API REST interface. We presented an architecture for delivering this service that meets the criteria for determining the level of smartness of the service in an academic context. These criteria led us to merge novel data representation (ontology) and processing (machine learning) technologies, as well as its synchronous visualization (Kafka), via the smart service. By combining the KNIME Analytics platform with KNIME Server, significant benefits in terms of automation and connectivity may be gained. We intend to test the integration of this smart service into collaborative platforms in the future to reveal the limitations and benefits of each of its functions in order to use it as predictive support for complementing teams.

Overall Conclusion and Prospects

Our study emphasizes the concept of "Smart City," which blends technology into people's lifestyles through highly sophisticated systems and infrastructures, and highlights how smart cities prepare the young generation for a bright future.To achieve this objective, the smart city needs human development, which contributes to its good governance, by training new profiles suitably designed to build and sustainably protect their smart urban environment. Citizens in smart cities can learn new skills via the use of better educational systems, thanks to the new education reform. This potential permits intelligent information exchange inside the new learning environment known as "Smart University." The contribution of our research is to examine the salient features of the "Smart University" that enables smart knowledge sharing between learners. Moreover, We have designed a smart teaching method of including all students in the teaching process ensuring democratized learning. The following contributions detail the proposed approach:

- The first part of realizing this thesis consists in making in-depth research concerning the global theme that is the " smart city", this preliminary study is very important because it gave us a background about this smart system and its behavior in affording the smart urban services to these citizens.
- As stated in Chapter 2, educational reform must align with the implementation of the "Smart City" project in order to train new profiles who will engage with their smart environment. As a result, it is clear to create an intelligent educational system in accordance with the standard of the smart city, which is the "Smart University." At this stage, we have developed a Smart University (SU) taxonomy that will assist in identifying the essential components involved in providing opportunities for solving smart learning system challenges.
- Smart University (SU) is always looking for innovative solutions to improve access and knowledge-sharing between different stakeholders. Among these solutions, collaboration is

a powerful concept that creates opportunities for sharing ideas in a virtual space with many collaborators. Since the development of competency involves a combination of knowledge, attitudes and skills, modeling interdisciplinary collaboration is a promising lever in designing smart learning systems. The management of interdisciplinary collaboration is closely related to: (1- searching in heterogeneous fields that represents students profiles including competencies required by the collaborative project,(2 ensuring optimal coherence based on the integration of the students in complementary teams,and (3 how to improve the selection of the complementary team that will participate in the collaborative project?

- This thesis represents compelling reasons to introduce Semantic Web technology as an artificial intelligence platform to impart smartness to universities. The integration of an open architecture that connects heterogeneous data sources within the university can effectively manage educational data and further provides opportunities to achieve the goals of managing interdisciplinary collaboration. Mostly, the Smart University (SU) encompasses proactive applications that meet the expectations of their learners based on decision support systems. A fair decision is closely linked to the quality and representation of the data used to solve a targeted problem, yet the strength of semantic technology allows it to be essential in the implementation of interoperable and sustainable solutions. Thus, according to a standardized process of building ontologies, we have modeled an ontology of the Smart University (SU) that represents these main concepts, the semantic relationships and the different axioms.

- We proposed a new approach for developing a smart system that predicts complementary teams involving students who have skills and preferences required by the problem solving designed in their educational program. To achieve this, we have proposed a new heuristic that calculates the complementarity between learners according to their skills in order to make the different compositions of complementary teams, this algorithm is a succession of blocks of sorting, search by criterion and Boolean algebra, which we applied to the Smart

University (SU) ontology. Once the result of this heuristic has been reached, we moved on to the classification stage, which takes the assignment of students to their complementary teams as "class" for building the predictive model, we carried out a comparative study between three classifiers in order to select the most qualified in predicting the composition of complementary teams.

- In this work, we encapsulated our predictive model in the Smart Collaborative Learning Service (SCLS) and deployed it in cloud architecture, more precisely Amazon Web Service (AWS) to make it on demand for third-party collaborative platforms via the interface REST-API. We have proposed a deployment architecture for this service that meets the requirements of the criteria defining the level of intelligence of the service in an academic environment. These criteria have led us to integrate innovative technologies for the representation (ontology) and processing (machine learning) of data and their synchronous updates (Kafka) through the smart service.

However, the proposed approach has a drawback in representing the student's skills, which are measured on a binary scale. As a result, the following enhancements are on the way:

- Integrating a scale of knowledge of each skill for each learner, the replacement of the binary scale [0,1] by a graduated scale [0,1,2,...10] that expresses the degree of expertise of the skill per student;
- Using fuzzy logic if there are other qualitative criteria for evaluating skills;
- Studying the complexity of the HBCT algorithm in computing time (in number of operations);
- Performing the analytical study of the algorithm;

- Analysis in complexity in expectation in the case of the fuzzification of the HBCT algorithm;

On the other hand, using the Smart Collaborative Learning Service (SCLS) as a decision support system enabling effective collaboration within the Smart University (SU) ecology should be guided by key standards such as data, processing and reuse. Therefore, we introduce future areas of our research as follows:

- Standardize the SCLS according to an intelligent service referent architecture;
- Normalize each of SCLS's layers;
- Standardize the logical relationship and interaction between layers;
- Assess the smart service maturation and capacity;

References

Abaker, Ibrahim et al. 2016. “International Journal of Information Management The Role of Big Data in Smart City.” *International Journal of Information Management* 36(5): 748–58. http://dx.doi.org/10.1016/j.ijinfomgt.2016.05.002.

Abid, Abir, Ilhem Kallel, and Mounir Ben Ayed. 2016. “Teamwork Construction in E-Learning System: A Systematic Literature Review.” *2016 15th International Conference on Information Technology Based Higher Education and Training, ITHET 2016*.

Abusaada, Hisham, and Abeer Elshater. 2020. “COVID-19 Challenge, Information Technologies, and Smart Cities: Considerations for Well-Being.” *International Journal of Community Well-Being* 3(3): 417–24.

Adetunji, Tajudeen et al. 2020. “An Ontology-Based Knowledge Acquisition Model for Software Anomalies Systems.” *2020 International Conference in Mathematics, Computer Engineering and Computer Science, ICMCECS 2020*.

ae Chun, Soon et al. 2012. “Collaborative E-Government.” *Transforming Government: People, Process and Policy* 6(1): 5–12.

Agarwal, Sonali. 2012. “Data Mining in Education: Data Classification and Decision Tree Approach.” *International Journal of e-Education, e-Business, e-Management and e-Learning* 2(2): 140–44.

Akhri, O., Y. El Bouzekri El Idrissi, and N. Hmina. 2019. “Enabling Smart Collaboration with Smart University (SU) Services.” In *ACM International Conference Proceeding Series*,.

Akhrif, O., C. Benfares, Y. El Bouzekri El Idrissi, and N. Hmina. 2019. “Smart Collaborative Learning :” 6(2): 52–66.

Akhrif, O., Y.E. El Idrissi, and N. Hmina. 2018. “Smart University: SOC-Based Study.” In *ACM International Conference Proceeding Series*,.

Akhrif, Ouidad, Chaymae Benfares, Younès El Bouzekri El Idrissi, and Nabil Hmina. 2020. “Collaborative Approaches in Smart Learning Environment: A Case Study.” *Procedia Computer Science* 175: 710–15. https://doi.org/10.1016/j.procs.2020.07.105.

Akhrif, Ouidad, and Younès El. “Smart University : SOC-Based Study.”

Akhrif, Ouidad, Youness El, and Bouzekri El. 2021. “Completeness Based Classi Fi Cation Algorithm : A Novel Approach for Educational Semantic Data Completeness Assessment.”

Alawadhi, Suha et al. 2012. “Building Understanding of Smart City Initiatives.” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 7443 LNCS: 40–53.

Albino, Vito, Umberto Berardi, and Rosa Maria Dangelico. 2015. “Smart Cities: Definitions, Dimensions, Performance, and Initiatives.” *Journal of Urban Technology* 22(1): 3–21.

Aljoumaa, Kadan, and Carine Souveyet. “Publishing Intentional Services Using Extended Semantic Annotation.”

Alshawish, Raja A., Salma A.M. Alfagih, and Mohamed S. Musbah. 2016. “Big Data Applications in

Smart Cities." *Proceedings - 2016 International Conference on Engineering and MIS, ICEMIS 2016*: 1–7.

Andrade Menolli, André Luís, Sheila Reinehr, and Andreia Malucelli. 2012. "Ontology for Organizational Learning Objects Based on LOM Standard." *38th Latin America Conference on Informatics, CLEI 2012 - Conference Proceedings* (October).

Ansari, Jamal Abdul Nasir, and Nawab Ali Khan. 2020. "Exploring the Role of Social Media in Collaborative Learning the New Domain of Learning." *Smart Learning Environments* 7(1).

Anthony Jnr, Bokolo. 2020. "Smart City Data Architecture for Energy Prosumption in Municipalities: Concepts, Requirements, and Future Directions." *International Journal of Green Energy* 17(13): 827–45. https://doi.org/10.1080/15435075.2020.1791878.

Anthopoulos, Leonidas, Marijn Janssen, and Vishanth Weerakkody. 2016. "Smart Service Portfolios." : 357–62.

Anttila, Juhani, and Kari Jussila. 2018. "Universities and Smart Cities: The Challenges to High Quality." *Total Quality Management and Business Excellence* 29(9–10): 1058–73.

Appio, Francesco Paolo, Marcos Lima, and Sotirios Paroutis. 2019. "Understanding Smart Cities: Innovation Ecosystems, Technological Advancements, and Societal Challenges." *Technological Forecasting and Social Change* 142(xxxx): 1–14. https://doi.org/10.1016/j.techfore.2018.12.018.

Arena, Marika et al. 2013. "Smart Mobility for Sustainability." *AEIT Annual Conference 2013: Innovation and Scientific and Technical Culture for Development, AEIT 2013 - Selected Proceedings Papers*.

Ate, Alev. 2016. *ATEŞ ÇOBANOĞLU ALEV , UZUNBOYLAR OKŞAN , KOÇ MELTEM , EĞİN FİGEN (2016). Bir Bilgisayar Yazılım ve Algoritma Açık De ...*

Baker, Ryan S. 2014. "Educational Data Mining: An Advance for Intelligent Systems in Education." *IEEE Intelligent Systems* 29(3): 78–82.

Baker, Ryan S J. 2011. "Encyclopedia of Data Warehousing and Mining." *Encyclopedia of Data Warehousing and Mining*.

Bakıcı, Tuba, Esteve Almirall, and Jonathan Wareham. 2013. "A Smart City Initiative: The Case of Barcelona." *Journal of the Knowledge Economy* 4(2): 135–48.

Benevolo, Clara, Renata Paola Dameri, and Beatrice D Auria. 2016. "Empowering Organizations: Enabling Platforms and Artefacts." 11: 315. http://link.springer.com/10.1007/978-3-319-23784-8.

Bhabad, Mayuri A, and Sudhir T. Bagade. 2015. "Internet of Things : Architecture , Security Issues and Countermeasures." *International Journal of Computer Applications* 125(14): 1–5.

Bousbia, Nabila, and Idriss Belamri. 2014. "Which Contribution Does EDM Provide to Computer-Based Learning Environments?" In *Educational Data Mining: Applications and Trends*, ed. Alejandro Peña-Ayala. Cham: Springer International Publishing, 3–28. https://doi.org/10.1007/978-3-319-02738-8_1.

Breiman, Leo. 1999. "Randon Forests." *Machinelearning202.Pbworks.Com*: 1–35. http://machinelearning202.pbworks.com/w/file/fetch/60606349/breiman_randomforests.pdf.

Bruneel, Johan, Pablo D'Este, and Ammon Salter. 2010. "Investigating the Factors That Diminish the Barriers to University-Industry Collaboration." *Research Policy* 39(7): 858–68. http://dx.doi.org/10.1016/j.respol.2010.03.006.

Bullinger, Hans-jörg, Jens Neuhüttler, Rainer Nägele, and Inka Woyke. 2017. "Collaborative Development of Business Models in Smart Service Ecosystems."

Burwood, Stephen. 2020. *Karl Jaspers (1883–1969): Truth, Academic Freedom and Student Autonomy*.

Calvanese, Diego et al. 2005. "DL-Lite: Tractable Description Logics for Ontologies." *Proceedings of the National Conference on Artificial Intelligence* 2: 602–7.

Cen, Ling et al. 2016. "Quantitative Approach to Collaborative Learning: Performance Prediction, Individual Assessment, and Group Composition." *International Journal of Computer-Supported Collaborative Learning* 11(2): 187–225. http://dx.doi.org/10.1007/s11412-016-9234-6.

Cha, Jeonghun, Sushil Kumar Singh, Tae Woo Kim, and Jong Hyuk Park. 2021. "Blockchain-Empowered Cloud Architecture Based on Secret Sharing for Smart City." *Journal of Information Security and Applications* 57(January): 102686. https://doi.org/10.1016/j.jisa.2020.102686.

Chamoso, Pablo et al. 2020. "Smart City as a Distributed Platform: Toward a System for Citizen-Oriented Management." *Computer Communications* 152(December 2019): 323–32. https://doi.org/10.1016/j.comcom.2020.01.059.

Chatti, Mohamed Amine, Anna Lea Dyckhoff, Ulrik Schroeder, and Hendrik Thüs. 2012. "A Reference Model for Learning Analytics." *International Journal of Technology Enhanced Learning* 4(5–6): 318–31.

Chatti, Mohamed Amine, Satish Srirama, David Kensche, and Yiwei Cao. 2006. "Mobile Web Services for Collaborative Learning."

Chen, Jianhui, and Jing Zhao. 2018. "An Educational Data Mining Model for Supervision of Network Learning Process." 13(11): 67–77.

Chen, Xianjun, Zhoupeng Ji, Yu Fan, and Yongsong Zhan. 2017. "Restful API Architecture Based on Laravel Framework." *Journal of Physics: Conference Series* 910(1).

Cheng, Bin et al. 2015. "Building a Big Data Platform for Smart Cities: Experience and Lessons from Santander." *Proceedings - 2015 IEEE International Congress on Big Data, BigData Congress 2015* (July 2015): 592–99.

Chergui, Meriyem, Aziza Chakir, and Hajar Mansouri. 2020. "Smart Pedagogical Knowledge Management System." *Universal Journal of Educational Research* 8(12): 6585–97.

Chin, Jeannette, Vic Callaghan, and Ivan Lam. 2017. "Understanding and Personalising Smart City Services Using Machine Learning, the Internet-of-Things and Big Data." *IEEE International Symposium on Industrial Electronics*: 2050–55.

Civitarese, Gabriele, Claudio Bettini, Timo Sztyler, and Daniele Riboni. 2019. "NewNECTAR : Collaborative Active Learning for Knowledge-Based Probabilistic Activity Recognition ☆." *Pervasive and Mobile Computing* 56: 88–105. https://doi.org/10.1016/j.pmcj.2019.04.006.

Costa, Carlos, and Maribel Yasmina Santos. 2016. "BASIS: A Big Data Architecture for Smart Cities." *Proceedings of 2016 SAI Computing Conference, SAI 2016*: 1247–56.

Costin, Aaron. 2016. "A New Methodology for Interoperability of Heterogeneous Bridge Information Models Copyright © 2016 By Aaron Costin a New Methodology for Interoperability Of." (May).

Cruz, Wilmax Marreiro, and Seiji Isotani. 2014. "Group Formation Algorithms in Collaborative Learning Contexts: A Systematic Mapping of the Literature." *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 8658 LNCS: 199–214.

D. Liu et al. 2017. "Contexts of Smart Learning Environments Smarter City." : 15–29.

Dalton, E. (2017), "Learn moodle august 2016", anonymized data set.

Dobre, C, and F Xhafa. 2014. "Intelligent Services for Big Data Science." *Future Generation Computer Systems* 37: 267–81. http://dx.doi.org/10.1016/j.future.2013.07.014.

Dyckhoff, Anna Lea, Dennis Zielke, Mareike Bültmann, and Mohamed Amine Chatti. 2014. "Design and Implementation of a Learning Analytics Toolkit for Teachers." (November).

EDM. "Educational Data Mining - Javatpoint." https://www.javatpoint.com/educational-data-mining (November 29, 2021).

van Engelen, Jesper E., and Holger H. Hoos. 2020. "A Survey on Semi-Supervised Learning." *Machine Learning* 109(2): 373–440. https://doi.org/10.1007/s10994-019-05855-6.

Fang, Chua Fang, and Lee Chien Sing. 2009. "Knowledge-Based Systems Collaborative Learning Using Service-Oriented Architecture : A Framework Design." *Knowledge-Based Systems* 22(4): 271–74. http://dx.doi.org/10.1016/j.knosys.2009.01.003.

Fernández, M., Gómez-Pérez, M. and Juristo. 2007. "Methontology: From Ontological Art Towards Ontological Engineering." *Proceedings - 3rd International Conference on Automated Production of Cross Media Content for Multi-Channel Distribution, AXMEDIS 2007*: 115–22.

Fillbrunn, Alexander et al. 2017. "KNIME for Reproducible Cross-Domain Analysis of Life Science Data." *Journal of Biotechnology* 261(February): 149–56.

FOAF. "FOAF Vocabulary Specification." *FOAF*. http://xmlns.com/foaf/spec/#term_Group (November 17, 2021).

Furht, Borko. 2010. "Handbook of Cloud Computing." *Handbook of Cloud Computing*: 3–19.

Gaur, Aditya, Bryan Scotney, Gerard Parr, and Sally McClean. 2015. "Smart City Architecture and Its Applications Based on IoT." *Procedia Computer Science* 52(1): 1089–94. http://dx.doi.org/10.1016/j.procs.2015.05.122.

Gavrilović, Nebojša, and Alok Mishra. 2021. "Software Architecture of the Internet of Things (IoT) for Smart City, Healthcare and Agriculture: Analysis and Improvement Directions." *Journal of Ambient Intelligence and Humanized Computing* 12(1): 1315–36. https://doi.org/10.1007/s12652-020-02197-3.

Gibert, Karina, Joaquín Izquierdo, Geoff Holmes, and Ioannis Athanasiadis. 2008. "On the Role of Pre and Post-Processing in Environmental Data Mining." : 1937–58.

Giffinger, R., & Pichler-Milanović, N. (2007). Smart cities: Ranking of European medium-sized cities. Centre of Regional Science, Vienna University of Technology.

Gomede, Everton et al. 2018. "Application of Computational Intelligence to Improve Education in Smart Cities." *Sensors (Switzerland)* 18(1): 1–26.

Gonzalez, T. et al. 2020. "Influence of COVID-19 Confinement on Students' Performance in Higher Education." *PLoS ONE* 15(10 October): 1–23. http://dx.doi.org/10.1371/journal.pone.0239490.

Grimm, Stephan, Pascal Hitzler, and Andreas Abecker. 2007. "Knowledge Representation and Ontologies Logic, Ontologies and Semantic Web Languages." *Semantic Web Services*: 51–105. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.71.8581%5Cnhttp://www.springerlink.com/content/p462153xn066j90m/.

Gros, Begoña. 2016. "The Design of Smart Educational Environments." *Smart Learning Environments* 3(1). http://dx.doi.org/10.1186/s40561-016-0039-x.

Gruber., T. R. "A Translation Approach to Portable Ontologies."

Gu, Tao, Hung Keng Pung, and Da Qing Zhang. 2005. "A Service-Oriented Middleware for Building Context-Aware Services." *Journal of Network and Computer Applications* 28(1): 1–18.

Guarino, Nicola. 1998. "Formal Ontology and Information Systems." *Formal Ontology in Information Systems: Proceedings of the 1st International Conference* 46(June): 3–15.

Guo, Yongan, Hongbo Zhu, and Longxiang Yang. 2017. "Smart Service System (SSS): A Novel Architecture Enabling Coordination of Heterogeneous Networking Technologies and Devices for Internet of Things." : 130–44.

Harrison, C. et al. 2010. "Foundations for Smarter Cities." *IBM Journal of Research and Development* 54(4): 1–16.

Harrison, Teresa M. et al. 2012. "Open Government and E-Government: Democratic Challenges from a Public Value Perspective." *Information Polity* 17(2): 83–97.

Hassan, Mohsen. 2017. "And Ontologies : Application to Rare Diseases Mohsen Hassan To Cite This Version : HAL Id : Tel-01678860 Soutenance et Mis à Disposition de l ' Ensemble de La Contact : Ddoc-Theses-Contact@univ-Lorraine.Fr."

Heinemann, Colleen, and Vladimir L. Uskov. 2018. 70 Smart Innovation, Systems and Technologies *Smart University: Literature Review and Creative Analysis*.

Hernández-García, Ángel, Emiliano Acquila-Natale, Julián Chaparro-Peláez, and Miguel Conde. 2018. "Predicting Teamwork Group Assessment Using Log Data-Based Learning Analytics." *Computers in Human Behavior* 89: 373–84.

Herrera-Pavo, Miguel Ángel. 2021. "Collaborative Learning for Virtual Higher Education." *Learning, Culture and Social Interaction* 28(April 2020): 100437. https://doi.org/10.1016/j.lcsi.2020.100437.

van den Hooff, Bart, Alexander P. Schouten, and Stojan Simonovski. 2012. "What One Feels and What One Knows: The Influence of Emotions on Attitudes and Intentions towards Knowledge Sharing." *Journal of Knowledge Management* 16(1): 148–58.

Huang, Changqin, Fuyin Xu, Xianghua Xu, and Xiaolin Zheng. 2006. "Towards an Agent-Based Robust Collaborative Virtual Environment for E-Learning in the Service Grid." : 702–7.

Huang, Linna, Fenghua Liu, and Chunli Liu. 2013. "Design and Research on Collaborative Learning Program Based on Cloud-Services." 759: 1199–1203.

Hwang, Gwo Jen. 2014. "Definition, Framework and Research Issues of Smart Learning Environments - a Context-Aware Ubiquitous Learning Perspective." *Smart Learning Environments* 1(1): 1–14.

Ifenthaler, Dirk, and David Gibson. 2020. *Adoption of Data Analytics in Higher Education Learning and Teaching*. http://link.springer.com/10.1007/978-3-030-47392-1.

IMS. "IMS Learning Design Best Practice and Implementation Guide Version 1.0 Final | IMS Global Learning Consortium." http://www.imsglobal.org/learningdesign/ldv1p0/imsld_bestv1p0.html (November 16, 2021).

IMS Learner. "IMS Learner Information Package Accessibility for LIP Information Model | IMS Global Learning Consortium." https://www.imsglobal.org/accessibility/acclipv1p0/imsacclip_infov1p0.html#1537632 (November 17, 2021).

Ishida, Toru. 2002. "Digital City Kyoto." *Communications of the ACM* 45(7): 76–81.

Jin, Ziwei et al. 2020. "RFRSF: Employee Turnover Prediction Based on Random Forests and Survival Analysis." *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 12343 LNCS: 503–15.

Kathayat, Surya Bahadur, and Rolv Braek. 2011. "Modeling Collaborative Learning Services - A Case Study." In *Proceedings of the 2011 International Conference on Collaboration Technologies and Systems, CTS 2011*, , 326–33.

Khan, Zaheer, Ashiq Anjum, Kamran Soomro, and Muhammad Atif Tahir. 2015. "Towards Cloud Based Big Data Analytics for Smart Future Cities."

Khan, Zaheer, and Saad Liaquat Kiani. 2012. "A Cloud-Based Architecture for Citizen Services in Smart Cities." *Proceedings - 2012 IEEE/ACM 5th International Conference on Utility and Cloud Computing, UCC 2012*: 315–20

Khanfir, Emna, and Raoudha Ben Djmeaa. 2016. "Automated Publish , Discovery and Composition of Intentional Web Services Adaptable to Both Quality and Context."

Kim, Taisiya, Ji Yeon Cho, and Bong Gyou Lee. 2013. "Evolution to Smart Learning in Public Education: A Case Study of Korean Public Education." *IFIP Advances in Information and Communication Technology* 395: 170–78.

Kitchin, Rob. 2016. "The Ethics of Smart Cities and Urban Science." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374(2083).

Kong, Xiangjie, and Bo Xu. 2018. "TNERec : Topic-Aware Network Embedding for Scientific Collaborator Recommendation." : 1007–14.

Koper, Rob, Bill Olivier, Rob Koper, and Bill Olivier. 2004. "International Forum of Educational Technology & Society Representing the Learning Design of Units of Learning Published by :

International Forum of Educational Technology & Society Stable URL : Https://Www.Jstor.Org/Stable/10.2307/Jeductechsoci.7.3.97 Lin." 7(3): 97–111.

Lama, Manuel, Eduardo Sánchez, Ricardo R Amorim, and Xosé A Vila. 2005. "Semantic Description of the IMS Learning. Design Specification." *SW-EL'05: Applications of Semantic Web Technologies for E-Learning in conjunction with 12th International Conference on Artificial Intelligence in Education (AIED 2005)*: 37–46.

Leo, Davinia Hernández, Juan I.Asensio Pérez, and Yannis A. Dimitriadis. 2004. "IMS Learning Design Support for the Formalization of Collaborative Learning Patterns." *Proceedings - IEEE International Conference on Advanced Learning Technologies, ICALT 2004* (May 2014): 350–54.

Liao, Zhifang et al. 2020. "Core-Reviewer Recommendation Based on Pull Request Topic Model and Collaborator Social Network." *Soft Computing* 24(8): 5683–93. https://doi.org/10.1007/s00500-019-04217-7.

Lin, Hsiu Fen. 2007. "Knowledge Sharing and Firm Innovation Capability: An Empirical Study." *International Journal of Manpower* 28(3–4): 315–32.

Liñán, Laura Calvet, and Ángel Alejandro Juan Pérez. 2015. "Educational Data Mining and Learning Analytics: Differences, Similarities, and Time Evolution." *RUSC Universities and Knowledge Society Journal* 12(3): 98–112.

Lisi, Francesca A. 2008. "Building Rules on Top of Ontologies for the Semantic Web with Inductive Logic Programming." *Theory and Practice of Logic Programming* 8(3): 271–300.

Liu, Ming-Chi, and Yueh-Min Huang. 2017. "The Use of Data Science for Education: The Case of Social-Emotional Learning." *Smart Learning Environments* 4(1): 1–13. http://dx.doi.org/10.1186/s40561-016-0040-4.

M, Hossin, and Sulaiman M.N. 2015. "A Review on Evaluation Metrics for Data Classification Evaluations." *International Journal of Data Mining & Knowledge Management Process* 5(2): 01–11.

Manjarres, A. V., Sandoval, L. G. M., & Suárez, M. S. (2018). Data mining techniques applied in educational environments: Literature review. Digital Education Review, (33), 235-266.

Manek, Asha S., P. Deepa Shenoy, M. Chandra Mohan, and K. R. Venugopal. 2017. "Aspect Term Extraction for Sentiment Analysis in Large Movie Reviews Using Gini Index Feature Selection Method and SVM Classifier." *World Wide Web* 20(2): 135–54. http://dx.doi.org/10.1007/s11280-015-0381-x.

Marioni G., Land Hilligje van't & Jensen Trine. 2020. *The Impact of Covid-19 on Higher Education around the World*. https://www.iau-aiu.net/IMG/pdf/iau_covid19_and_he_survey_report_final_may_2020.pdf.

Martín, Marta, Verónica Basilotta Gómez-pablos, and Ana García-valcárcel Muñoz-repiso. 2017. "A Quantitative Approach to Pre-Service Primary School Teachers ' Attitudes towards Collaborative Learning with Video Games : Previous Experience with Video Games Can Make the Difference."

Mavers, Scott. 2020. "Running Head : A REVIEW OF LEARNING OBJECTS A Review of Learning

Objects Scott W . Mavers University of North Texas.” (November).

McGee, Clive. 2015. “Introduction to the Special Section on Curriculum.” *Waikato Journal of Education* 14(1): 3–6.

McGill, Tanya J., and Jane E. Klobas. 2009. “A Task-Technology Fit View of Learning Management System Impact.” *Computers and Education* 52(2): 496–508. http://dx.doi.org/10.1016/j.compedu.2008.10.002.

Menolli, Andre, H. Sofia Pinto, Sheila Reinehr, and Andreia Malucelli. 2013. “An Incremental and Iterative Process for Ontology Building.” *CEUR Workshop Proceedings* 1041: 215–20.

Moreno, Julián, Demetrio A. Ovalle, and Rosa M. Vicari. 2012. “A Genetic Algorithm Approach for Group Formation in Collaborative Learning Considering Multiple Student Characteristics.” *Computers and Education* 58(1): 560–69. http://dx.doi.org/10.1016/j.compedu.2011.09.011.

Nam, Taewoo, and Theresa A. Pardo. 2011. “Smart City as Urban Innovation: Focusing on Management, Policy, and Context.” *ACM International Conference Proceeding Series*: 185–94.

Nasution, A. A., K. Erwin, and Risanty. 2020. “Smart City: Is Your City Ready?” *IOP Conference Series: Earth and Environmental Science* 562(1).

Neri, Mario Arrigoni. 2001. “Ontology-Based Learning Objects Sequencing.”

Neri, Mario Arrigoni, and Marco Colombetti. 2009. “Ontology-Based Learning Objects Search and Courses Generation.” *Applied Artificial Intelligence* 23(3): 233–60.

Neuroni, Alessia C. et al. 2019. “Public Value Creation in a Smart City Context: An Analysis Framework.” *Public Administration and Information Technology* 35: 49–76.

Ning, Zhaolong et al. 2018. “Green and Sustainable Cloud of Things : Enabling Collaborative Edge Computing.” *IEEE Communications Magazine* PP: 1–7.

Noy, Natalya F., and Deborah L. McGuinness. 2001. “A Guide to Creating Your First Ontology.” *Biomedical Informatics Reseach*: 7–25. http://bmir.stanford.edu/file_asset/index.php/108/BMIR-2001-0880.pdf.

Oliveira, Álvaro, and Margarida Campolargo. 2015. “From Smart Cities to Human Smart Cities.” *Proceedings of the Annual Hawaii International Conference on System Sciences* 2015-March: 2336–44.

Olson, David L., and Desheng. Wu. 2016. *Predictive Data Mining Models.*

Pal, M. 2005. “Random Forest Classifier for Remote Sensing Classification.” *International Journal of Remote Sensing* 26(1): 217–22.

Palma, H. (2020), “Dataset wayuu”, Mendeley Data, V1,doi: 10.17632/fd4rz92zbj.1.

Papazoglou, Michael P, Paolo Traverso, Schahram Dustdar, and Frank Leymann. 2007. “Service-Oriented Computing: State of the Art and Research Challenges.”

Paskaleva, Krassimira, and Ian Cooper. 2020. *Innovations in Co-Created Smart City Services*.

Pe, Alejandro. 2016. 17 The Turkish Online Journal of Distance Education *Book Review Educational Data Mining: Applications and Trends*.

Petkovic, Dragutin et al. 2016. “Using the Random Forest Classifier to Assess and Predict Student

Learning of Software Engineering Teamwork." *Proceedings - Frontiers in Education Conference, FIE* 2016-Novem.

Petrolo, Riccardo. 2016. "Semantic-Based Discovery and Integration of Heterogeneous Things in a Smart City Environment To Cite This Version : HAL Id : Tel-01403844 Semantic-Based Discovery and Integration of Heterogeneous Things in a Smart City Environment."

Polese, Francesco. 2012. "Toward a Service (Eco) Systems Perspective on Value Creation." 3(September): 12–25.

Psomadaki, Ofilia I, Charalampos A Dimoulas, George M Kalliris, and Gregory Paschalidis. 2018. "Digital Storytelling and Audience Engagement in Cultural Heritage Thessaloniki." *Journal of Cultural Heritage*. https://doi.org/10.1016/j.culher.2018.07.016.

Ramírez-Donoso, Luis, Mar Pérez-Sanagustín, and Andrés Neyem. 2018. "MyMOOCSpace: Mobile Cloud-Based System Tool to Improve Collaboration and Preparation of Group Assessments in Traditional Engineering Courses in Higher Education." *Computer Applications in Engineering Education* 26(5): 1507–18.

Ray, Julie Basu. 2017. "Collaborative Interdisciplinary Teaching and Technologies and Smart Devices." *2017 5th IEEE International Conference on MOOCs, Innovation and Technology in Education (MITE)*: 77–82.

Ray, Santosh, and Mohammed Saeed. 2018. "Applications of Educational Data Mining and Learning Analytics Tools in Handling Big Data in Higher Education." *Applications of Big Data Analytics: Trends, Issues, and Challenges*: 135–60.

Rees, Reinout Van. 2003. "Clarity in the Usage of the Terms Ontology, Taxonomy an Classification." *Cib Report* 284: 432–39.

Rego, Conceição, and M Isabel Sánchez-hernández. 2019. *New Paths of Entrepreneurship Development: The Role of Education, Smart Cities and Social Factors*.

Reis, Rachel Carlos Duque et al. 2018. "Affective States in Computer-Supported Collaborative Learning: Studying the Past to Drive the Future." *Computers and Education* 120(April 2016): 29–50.

Rezgui, Kalthoum, Hedia Mhiri, and Khaled Ghédira. 2014. "An Ontology-Based Profile for Learner Representation in Learning Networks." *International Journal of Emerging Technologies in Learning* 9(3): 16–25.

Rochet, Claude. 2014. "Les Villes Intelligentes, Enjeux et Stratégies Pour de Nouveaux Marchés Le Programme MUST : Management of Urban Smart Territories." *CSDM 2014 Paris* (November). https://www.researchgate.net/profile/Claude_Rochet/publication/267928154_Les_villes_intelligentes_enjeux_et_strategies_pour_de_nouveaux_marches_Le_programme_MUST_Management_of_Urban_Smart_Territories/links/545da1130cf295b5615e70e2.pdf.

Rodríguez Bolívar, Manuel Pedro. 2019. "The Relevance of Public Value into Smart Cities." *Public Administration and Information Technology* 35: 3–13.

Romero, C., and S. Ventura. 2007. "Educational Data Mining: A Survey from 1995 to 2005." *Expert*

Systems with Applications 33(1): 135–46.

Romero, Cristbal, and Sebastin Ventura. 2010. “Educational Data Mining: A Review of the State of the Art.” *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews* 40(6): 601–18.

Romero, Cristobal, and Sebastian Ventura. 2013. “Data Mining in Education.” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 3(1): 12–27.

Romero, Cristobal, and Sebastian Ventura. 2020. “Educational Data Mining and Learning Analytics: An Updated Survey.” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10(3): 1–21.

Rutkauskiene, Danguole, Daina Gudoniene, and Rytis Maskeliunas. 2016. “Smart Education and E-Learning 2016.” 59: 291–301. http://link.springer.com/10.1007/978-3-319-39690-3.

S. Bechhofer F. van Harmelen, J Hendler I Horrocks D L McGuinness, and L A Stein. 2004. “OWL Web Ontology Language Reference, Http://Www.W3.Org/TR/2004/REC-Owl-Ref-20040210/.” (February). http://www.w3.org/TR/2004/REC-owl-ref-20040210/.

Sá, J. A. S. et al. 2016. “Lightning Forecast Using Data Mining Techniques On Hourly Evolution Of The Convective Available Potential Energy.” (March): 1–5.

Salma, Najar. 2014. “Adaptation Des Services Sensibles Au Contexte Selon Une Approche Intentionnelle Najar Salma To Cite This Version : HAL Id : Tel-00989775 Salma Najar Adaptation Des Services Sensibles Au Contexte Selon Une Approche Intentionnelle.”

Scekic, Ognjen, and Stefan Nastic. 2018. “Blockchain-Supported Smart City Platform for Social Value Co-Creation and Exchange.” *IEEE Internet Computing* PP(c): 1.

Scheuer, Oliver, and Bruce M Mclaren. 2011. “Ducational Ata Ining.”

Schleicher, Johannes M., Michael Vogler, Schahram Dustdar, and Christian Inzinger. 2016. “Application Architecture for the Internet of Cities: Blueprints for Future Smart City Applications.” *IEEE Internet Computing* 20(6): 68–75.

Schuurman, Dimitri, Bastiaan Baccarne, Lieven De Marez, and Peter Mechant. 2012. “Smart Ideas for Smart Cities: Investigating Crowdsourcing for Generating and Selecting Ideas for ICT Innovation in a City Context.” *Journal of Theoretical and Applied Electronic Commerce Research* 7(3): 49–62.

Sciarretta, Giada, Roberto Carbone, and Silvio Ranise. 2016. “A Delegated Authorization Solution for Smart-City Mobile Applications.” *2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a Better Tomorrow, RTSI 2016*.

Shao, Xiwen. 2012. “Study on Security Issue of Internet of Things Based on RFID.” *Proceedings - 4th International Conference on Computational and Information Sciences, ICCIS 2012*: 566–69.

Shen, Chang Xiang et al. 2007. “Survey of Information Security.” *Science in China, Series F: Information Sciences* 50(3): 273–98.

Shrivastava, Shubhi, Iti Mathur, and Nisheeth Joshi. 2018. “An Ontology Development for University.” *2018 International Conference on Advanced Computation and Telecommunication, ICACAT*

2018: 1–8.

Singh, Himanshu. 2021. Practical Machine Learning with AWS *Practical Machine Learning with AWS: Process, Build, Deploy, and Productionize Your Models Using AWS*.

"Smart Education for Smart Cities: Visual, Collaborative & Interactive." https://hub.beesmart.city/en/solutions/smart-people/smart-education/viewsonic-smart-education-for-smart-cities (January 5, 2022).

Tett, Robert P., and Patrick J. Murphy. 2002. "Personality and Situations in Co-Worker Preference: Similarity and Complementarity in Worker Compatibility." *Journal of Business and Psychology* 17(2): 223–43.

Tokody, Daniel, and Imre Janos Mezei. 2017. "Creating Smart, Sustainable and Safe Cities." *SISY 2017 - IEEE 15th International Symposium on Intelligent Systems and Informatics, Proceedings*: 141–45.

Trilles, Sergio, Andrea Calia, Óscar Belmonte, Joaquín Torres-Sospedra, et al. 2017. "Deployment of an Open Sensorized Platform in a Smart City Context." *Future Generation Computer Systems* 76: 221–33. http://dx.doi.org/10.1016/j.future.2016.11.005.

Trilles, Sergio, Andrea Calia, Óscar Belmonte, and Joaquín Torres-sospedra. 2017. "Deployment of an Open Sensorized Platform in a Smart City Context." *Future Generation Computer Systems* 76: 221–33. http://dx.doi.org/10.1016/j.future.2016.11.005.

Ullah, Mohammad Aman, and Syed Akhter Hossain. 2019. 814 Advances in Intelligent Systems and Computing *Ontology-Based Information Retrieval System for University: Methods and Reasoning*. Springer Singapore. http://dx.doi.org/10.1007/978-981-13-1501-5_10.

UN. 2008. "E c o n o m i c S o c i a l World Urbanization Prospects The 2007 Revision Executive Summary." (February). http://www.un.org/esa/population/publications/wup2007/2007WUP_ExecSum_web.pdf.

Urbieta, A et al. 2017. "Adaptive and Context-Aware Service Composition for IoT-Based Smart Cities." *Future Generation Computer Systems* 76: 262–74. http://dx.doi.org/10.1016/j.future.2016.12.038.

Uskov, Vladimir L. et al. 2017. "Building Smart Learning Analytics System for Smart University." *Smart Innovation, Systems and Technologies* 75(5): 191–204.

Uskov, Vladimir L., Jeffrey P. Bakken, Srinivas Karri, et al. 2018. 70 Smart Innovation, Systems and Technologies *Smart University: Conceptual Modeling and Systems' Design*.

Uskov, Vladimir L., Jeffrey P. Bakken, Robert J. Howlett, and Lakhmi C. Jain. 2018. "Innovations in Smart Universities." *Smart Innovation, Systems and Technologies* 70: 1–7.

Varia, Jinesh, and Sajee Mathew. 2014. "Overview of Amazon Web Services (Survey Report)." (January): 1–30. http://media.amazonwebservices.com/AWS_Overview.pdf.

Velte, Antonhy T. Velte, Toby J. Velte, and Robert Elsenpeter. 2010. *Cloud Computing: A Practical Approach*.

Venkatachalapathy, K, V Vijayalakshmi, and V Ohmprakash. 2020. "Educational Data Mining Tools : A Survey from 2001 to 2016." (September).

Verstegen, D. M.L. et al. 2018. “How Do Virtual Teams Collaborate in Online Learning Tasks in a MOOC?” *International Review of Research in Open and Distance Learning* 19(4): 39–55.

Vilajosana, Ignasi et al. 2013. “Bootstrapping Smart Cities through a Self-Sustainable Model Based on Big Data Flows.” (June): 128–34.

Vinod Kumar, T. M., and Bharat Dahiya. 2017. *Smart Economy in Smart Cities*.

Vranic, M, D Pintar, and Z Skocir. 2007. “The Use of Data Mining in Education Environment.” : 243–50.

Wan, Puspa Melati, and Abdul Halim Wan. 2020. *Clinical Sociology Moving from Theory to Practice*. http://link.springer.com/10.1007/978-3-030-49083-6.

Wang, Shouhong. 2008. “Ontology of Learning Objects Repository for Pedagogical Knowledge Sharing.” *Interdisciplinary Journal of e-Skills and Lifelong Learning* 4: 001–012.

Washburn, Doug, and Usman Sindhu. 2009. “Helping CIOs Understand ‘Smart City’ Initiatives.” *Growth*: 17. http://c3328005.r5.cf0.rackcdn.com/73efa931-0fac-4e28-ae77-8e58ebf74aa6.pdf.

Witten, Ian H, Eibe Frank, and Mark A Hall. *No Title*.

Yi, Fengji, Wenlong Fu, and Huan Liang. 2018. “Model-Based Reinforcement Learning: A Survey.” *Proceedings of the International Conference on Electronic Business (ICEB)* 2018-Decem: 421–29.

Yin, Chengjiu, and Yoshiyuki Tabata. 2009. “A Collaborative Learning Service for SNS in Ubiquitous Computing Environment.” (Figure 1).

Zhang, Xian Da. 2020. A Matrix Algebra Approach to Artificial Intelligence *A Matrix Algebra Approach to Artificial Intelligence*.

Zhuang, Fuzhen et al. 2021. “A Comprehensive Survey on Transfer Learning.” *Proceedings of the IEEE* 109(1): 43–76.

Zhuang, Rongxia et al. 2017. “Smart Learning Environments for a Smart City: From the Perspective of Lifelong and Lifewide Learning.” *Smart Learning Environments* 4(1).

Zine, Othmane, Aziz Derouich, and Abdennebi Talbi. 2019. “IMS Compliant Ontological Learner Model for Adaptive E-Learning Environments.” *International Journal of Emerging Technologies in Learning* 14(16): 97–119.

Zorić, Alisa Bilal. 2020. “Benefits of Educational Data Mining.” 6(1): 12–16.

Zygiaris, Sotiris. 2013. “Smart City Reference Model: Assisting Planners to Conceptualize the Building of Smart City Innovation Ecosystems.” *Journal of the Knowledge Economy* 4(2): 217–31.

Appendix 1 : The HBCT validation

Processing example from i=1 and k=2

Tamp table value	Iteration	i= 1		K= 2				
	Skills Matrix							
		C1	C2	C3	C4	C5		
	Learner1	0	1	0	0	0		
	Learner 2	1	1	0	0	0	BP	Compl_table
{1;2}	Tamp	1	1	0	0	0	0	Not yet
	Missed_comp	---	---	C3	C4	C5		

2 missed_comp found for Learner3 (k=3)

Tamp table value	Iteration	i= 1		K= 3				
	Skills Matrix							
		C1	C2	C3	C4	C5		
{1;2}	Tamp	1	1	0	0	0		
	Learner 3	0	0	1	0	1	BP	Compl_table
{1;2;3}	Tamp	1	1	1	0	1	0	Not yet
	Missed_comp	---	---	C3	C4	C5		

Tamp table value	Iteration	i= 1		K= 4		
	Skills Matrix					
		C1	C2	C3	C4	C5
{1;2;3}	Tamp	1	1	1	0	1
	Learner 4	0	0	0	0	0
{1;2;3}	Tamp	1	1	1	0	1
	Missed_comp	---	---	---	C4	---

Full combination

Tamp table value	Iteration	i= 1		K= 5				
	Skills Matrix							
		C1	C2	C3	C4	C5		
{1;2;3}	Tamp	1	1	1	0	1		
	Learner 5	1	0	0	1	0	BP	Compl_table
{1;2;3;5}	Tamp	1	1	1	1	1	1	{1;2;3;5}
	Missed_comp	---	---	---	C4	---		

Full combination

Tamp table value	Iteration	i= 1		K= 6				
	Skills Matrix							
		C1	C2	C3	C4	C5		
{1;2;3}	Tamp	1	1	1	0	1		
	Learner6	0	1	1	1	0	BP	Compl_table
{1;2;3;6}	Tamp	1	1	1	1	1	1	{1 ;2 ;3 ;6}
	Missed_comp	---	---	---	C4	---		

Full combination

Tamp table value	Iteration	i= 1		K= 7				
	Skills Matrix							
		C1	C2	C3	C4	C5		
{1;2;3}	Tamp	1	1	1	0	1		
	Learner7	0	1	0	1	1	BP	Compl_table
{1;2;3;7}	Tamp	1	1	1	1	1	1	{1 ;2 ;3 ;7}
	Missed_comp	---	---	---	C4	---		

(continued)

Example of processing from i = 10

Tamp table value	Iteration	i =	10					
	Skills Matrix							
Tamp table value		C1	C2	C3	C4	C5		
	Learner10	1	0	0	1	0		
							BP	Compl_table
{10}	Tamp	1	0	0	1	0	0	Not yet
	Missed_comp	---	---	---	---	---		

Tamp table value	Iteration	i =	10	K =	11			
	Skills Matrix							
Tamp table value		C1	C2	C3	C4	C5		
{10}	Tamp	1	0	0	1	0		
	Learner11	1	0	1	1	0	BP	Compl_table
{10;11}	Tamp	1	0	1	1	0	0	Not yet
	Missed_comp	---	C2	C3	---	C5		

Tamp table value	Iteration	i =	10	K =	12			
	Skills Matrix							
Tamp table value		C1	C2	C3	C4	C5		
{10;11}	Tamp	1	0	1	1	0		
	Learner12	0	1	1	0	1	BP	Compl_table
{10;11;12}	Tamp	1	1	1	1	1	1	{10;11;12}
	Missed_comp	---	C2	---	---	C5		

Tamp table value	Iteration	i =	10	K =	13			
	Skills Matrix							
Tamp table value		C1	C2	C3	C4	C5		
{10;11}	Tamp	1	0	1	1	0		
	Learner13	0	1	1	0	1	BP	Compl_table
{10;11;13}	Tamp	1	1	1	1	1	1	{10;11;13}
	Missed_comp	---	C2	---	---	C5		

Tamp table value	Iteration	i =	10	K =	14			
	Skills Matrix							
Tamp table value		C1	C2	C3	C4	C5		
{10;11}	Tamp	1	0	1	1	0		
	Learner14	0	1	1	0	0	BP	Compl_table
{10;11;14}	Tamp	1	1	1	1	0	0	Not yet
	Missed_comp	---	C2	---	---	C5		

Tamp table value	Iteration	i =	10	K =	15			
	Skills Matrix							
Tamp table value		C1	C2	C3	C4	C5		
{10;11;14}	Tamp	1	1	1	1	0		
	Learner15	0	1	1	0	0	BP	Compl_table
{10;11;14}	Tamp	1	1	1	1	0	0	Not yet
	Missed_comp	---	---	---	---	C5		

Tamp table value	Iteration	i =	10	K =	16			
	Skills Matrix							
Tamp table value		C1	C2	C3	C4	C5		
{10;11;14}	Tamp	1	1	1	1	0		
	Learner16	0	1	1	1	1	BP	Compl_table
{10;11;14;16}	Tamp	1	1	1	1	1	1	{10;11;14;16}
	Missed_comp	---	---	---	---	C5		

(continued)

Example of processing from i = 18

	Iteration	i =	18					
	Skills Matrix							
Tamp table value		C1	C2	C3	C4	C5		
	Et18	1	1	1	0	0		
							BP	Compl_table
{18}	Tamp	1	1	1	0	0	0	Not yet
	Missed_comp	---	---	---	---	---		

	Iteration	i =	18	K =	19			
	Skills Matrix							
Tamp table value		C1	C2	C3	C4	C5		
{18}	Tamp	1	1	1	0	0		
	Et19	0	1	1	0	1	BP	Compl_table
{18;19}	Tamp	1	1	1	0	1	0	Not yet
	Missed_comp	---	---	---	C4	C5		

	Iteration	i =	18	K =	20			
	Skills Matrix							
Tamp table value		C1	C2	C3	C4	C5		
{18;19}	Tamp	1	1	1	0	1		
	Et20	0	0	1	1	0	BP	Compl_table
{18;19;20}	Tamp	1	1	1	1	1	1	{18;19;20}
	Missed_comp	---	---	---	C4	---		

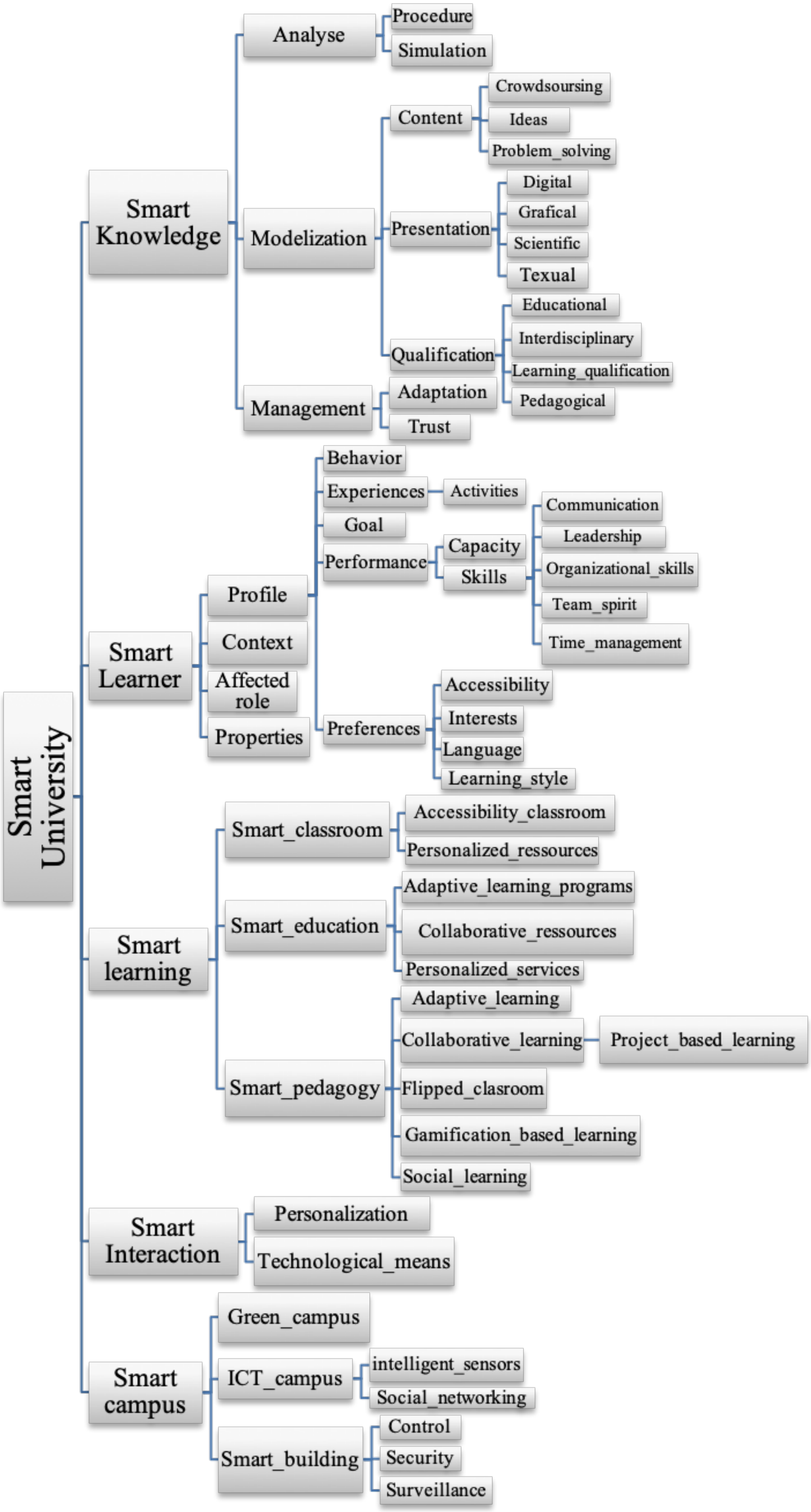
	Iteration	i =	18	K =	19			
	Skills Matrix							
Tamp table value		C1	C2	C3	C4	C5		
{18}	Tamp	1	1	1	0	0		
	Et20	0	0	1	1	0	BP	Compl_table
{18;20}	Tamp	1	1	1	1	0	0	Not yet
	Missed_comp	---	---	---	C4	C5		

	Iteration	i =	18	K =	18			
	Skills Matrix							
Tamp table value		C1	C2	C3	C4	C5		
ø	Tamp	0	0	0	0	0		
	Et19	0	1	1	0	1	BP	Compl_table
{19}	Tamp	0	1	1	0	1	0	Not yet
	Missed_comp	C1	C2	C3	C4	C5		

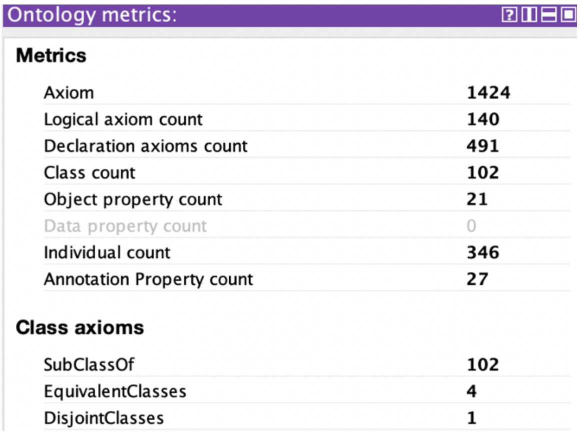
	Iteration	i =	18	K =	20			
	Skills Matrix							
Tamp table value		C1	C2	C3	C4	C5		
{19}	Tamp	0	1	1	0	1		
	Et20	0	0	1	1	0	BP	Compl_table
{19;20}	Tamp	0	1	1	1	1	0	Not yet
	Missed_comp	C1	---	---	C4	---		

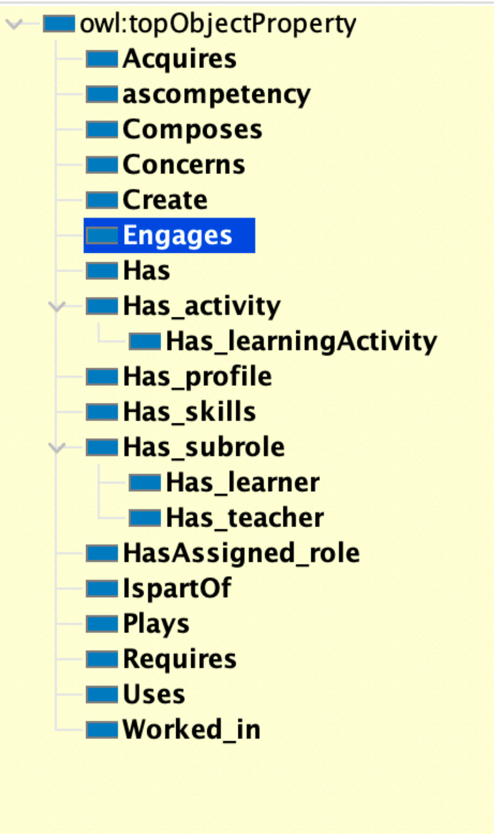
	Iteration	i =	18	K =	19			
	Skills Matrix							
Tamp table value		C1	C2	C3	C4	C5		
ø	Tamp	0	0	0	0	0		
	Et20	0	0	1	1	0	BP	Compl_table
{20}	Tamp	0	0	1	1	0	0	Not yet
	Missed_comp	C1	C2	C3	C4	C5		

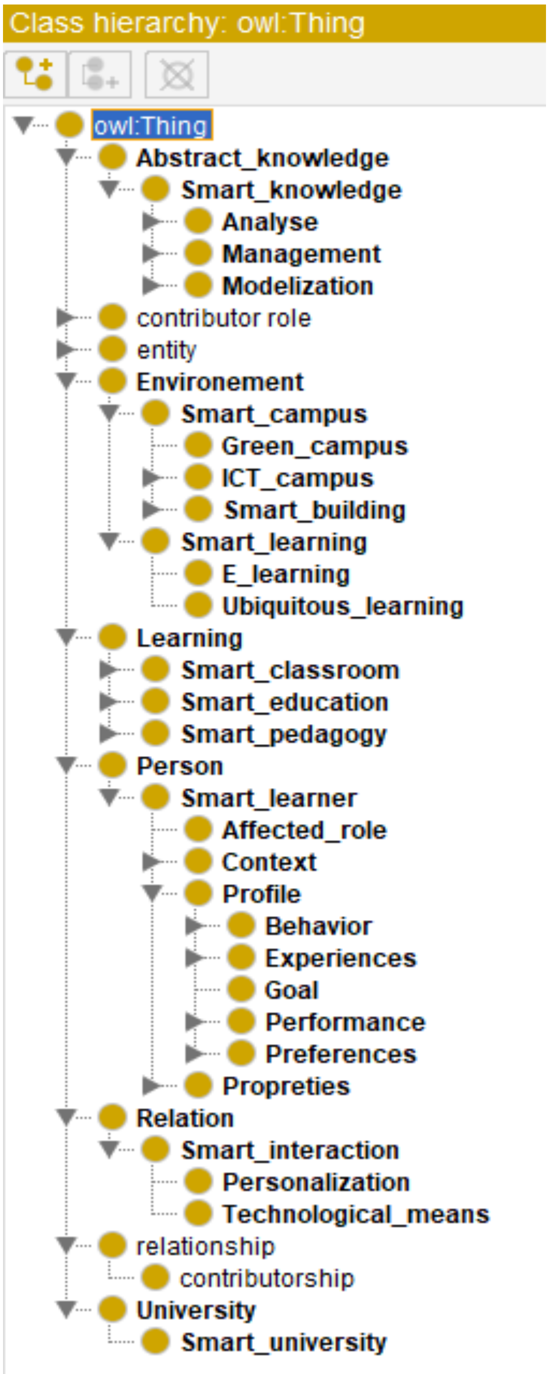
Appendix 2 : Smart University (SU) taxonomy



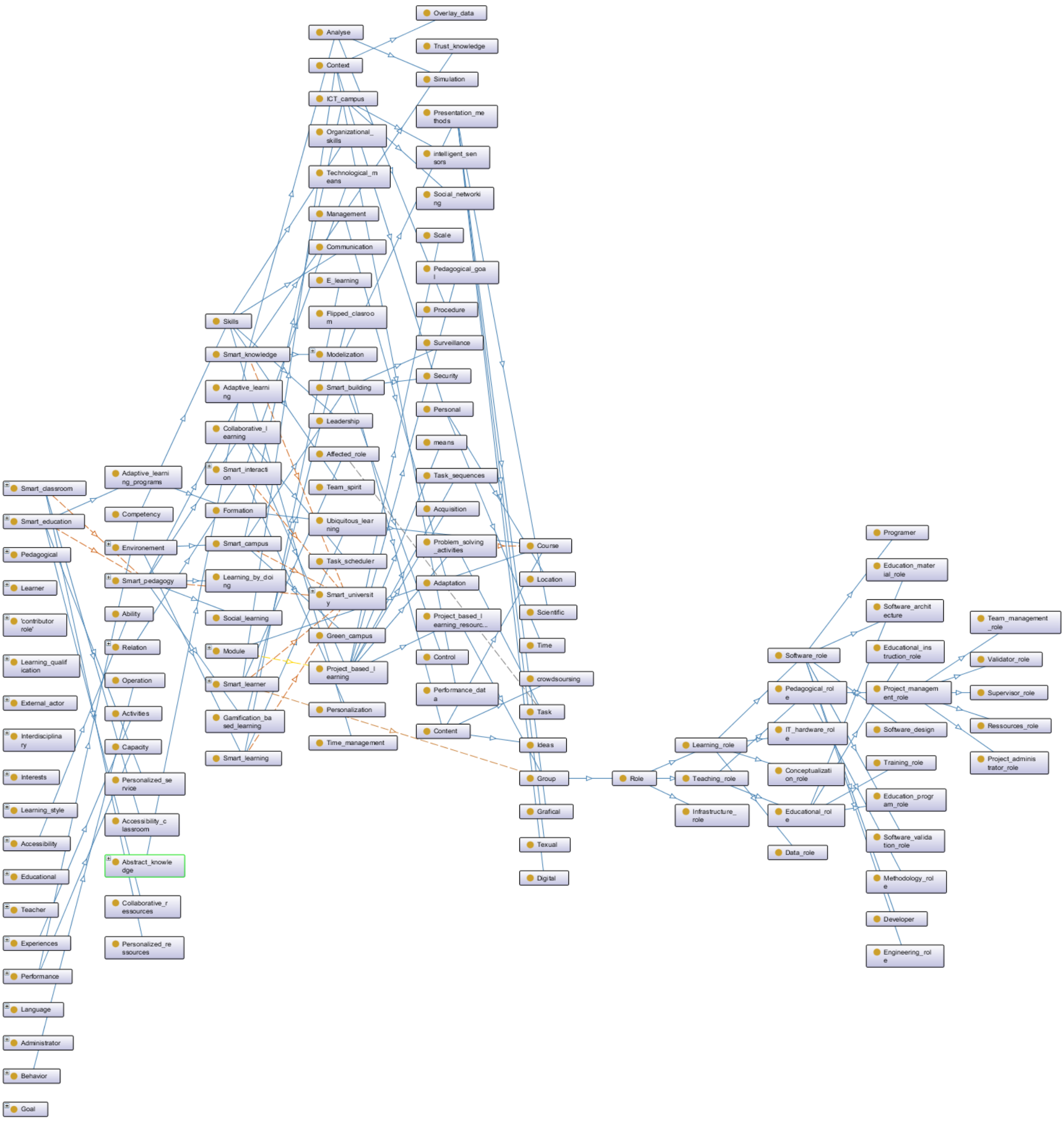
Appendix 3 : Concepts and relationships of the SU ontology







Appendix 4 : The SU ontology



Appendix 5 : Python implementation of the Random Forest algorithm

```
import pandas as pd

data=pd.read_excel("/Users/akhrifouidad/Downloads/Collaborative_learning.xlsx")

from sklearn.ensemble import RandomForestClassifier

Y=data["Classe"]

X=data.drop(columns=["Classe"])

from sklearn.model_selection import train_test_split

X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.2)

X_train.shape

X_test.shape

model=RandomForestClassifier()

model.fit(X_train,Y_train)

from sklearn import preprocessing

le=preprocessing.LabelEncoder()

data["Gender"]=le.fit_transform(data["Gender"])

Y_predict=model.predict(X_test)

Y_predict

model.score(X_test,Y_test)

pip install shap

shap_values = shap.TreeExplainer(model).shap_values(X_train)

shap.summary_plot(shap_values, X_train, plot_type="bar")
```