REPUBLIC OF TÜRKİYE

ALTINBAŞ UNIVERSITY

Institute of Graduate Studies

Information Technologies

USING DATAMINING TECHNIQUES FOR DATA NORMALIZATION IN SMART CLASSROOMS

Abdullah Ragheb Hamid ALBAKER

Master's Thesis

Supervisor

Prof. Dr. Osman Nuri Uçan

Istanbul, 2022

USING DATAMINING TECHNIQUES FOR DATA NORMALIZATION IN SMART CLASSROOMS

Abdullah Ragheb Hamid ALBAKER

Information Technologies

Master’s Thesis

ALTINBAŞ UNIVERSITY

2022

The thesis titled USING DATAMINING TECHNIQUES FOR DATA NORMALIZATION IN SMART CLASSROOMS prepared by ABDULLAH AL BAKER and submitted on 21 /12/2022 has been **accepted unanimously** for the degree of Master of Science in Department of Computer Engineering.

Prof. Dr Osman Nuri UÇAN

the Supervisor

Thesis Defense Committee Members:

Prof .Dr. Osman Nuri UÇAN	Department of Electrical and Electronics Engineering, Altınbaş University	______________
Asst. Prof. Dr. Sefer KURNAZ	Department of Computer Engineering, Altınbaş University	______________
Prof. Dr. Hasan Hüseyin BALIK	Faculty of Electrical and Electronics, Yıldız Technical University	______________

I hereby declare that this thesis meets all format and submission requirements of a Master's thesis.

Submission date of the thesis to Institute of Graduate Studies: ___/___/___

I hereby declare that all information and data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Abdullah ALBAKER

Signature

ABSTRACT

USING DATAMINING TECHNIQUES FOR DATA NORMALIZATION IN SMART CLASSROOMS

Abdullah Ragheb Hamid ALBAKER

Altınbaş University

Supervisor: Prof. Dr Osman Nuri Uçan

Date: December/2022

Pages: 71

Under the circumstances, teachers have adapted to seeing their classrooms equipped with computers, so far it has remained rather passive. The computer is traditionally used as part of an educational sequence, for the purpose of learning (computer environments for human learning), to document itself, to play games for teaching, or to achieve rhythmic automation of a digital learning tool. We argue here that some advances in computer technologies applied to teaching and learning could change these roles. Therefore, the current work aims to clarify the use of machine learning algorithms in the teaching and learning process through experiments conducted on a set of synthetic data and in order to favour the reproduction of these experiments. Our results revealed patterns of information that would be impossible to obtain quickly without the support of machine learning techniques. Thus, it was possible to supplement the instructions and accessible materials for reproducing the experiments with the considered techniques of artificial intelligence, as well as others.

KEYWORDS: E-LEARNING, K-MEANS, AI, ML.

TABLE OF CONTENTS

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

ABBREVIATIONS

RFID	:	Radio Frequency Identification
MCDM	:	Multi-Criteria Decision-Making
ACORN	:	Automatic Classroom Observation Recognition Network
SAW	:	Simple Additive Weeighting
SPS	:	Selective Problem Solving
TOPSIS	:	Technique for Order Preference by Similarity to the Ideal Solution
CRS	:	Classroom Response Systems
ARS	:	An audience Response System

1. INTRODUCTION

1.1 BACKGROUND

While teachers are now used to seeing their classrooms equipped with computers, these have so far remained fairly passive their manipulation remains a voluntary, controlled and considered act. The computer is traditionally used as part of an educational sequence, for the purpose of familiarization with (computer environments for human learning), to document itself, to play games for teaching, or to achieve a rhythmic automation of a digital teaching aid. We argue here that certain advances in computer techniques applied to teaching and learning could change these roles.

For around twenty years, context-aware rooms have been deployed, often still as prototypes. These classrooms are often presented as "intelligent", that is to say equipped with sensors responsible for measuring a certain number of phenomena taking place in the classroom and in return proposing their analysis, either in real time or retrospectively. We wish to show that these SCs, if they can shed light on the teaching-learning processes in their authentic classroom context, must however be designed and used within an ethical framework, the importance of which we will explain.

1.1.1 What Are Smart Classrooms

These intelligent classrooms, also known as "ambient" classrooms, are at least equipped with video and audio sensors (respectively their substitute eyes and ears), and effectors capable of responding to perceived stimuli, and a great “behind the scenes” computing power, in order to collect, process and even interpret what is at stake in the teaching and learning activities, according to the perspective(s) selected.
Beyond that, the number of sensors available today makes it possible to envisage a considerable spectrum of analyses. Thus, equipping the teacher with an eye tracker allows both to "see his gaze", in other words to see what he sees, and to see "inside him", since his pupillary diameter can provide information on the cognitive load it bears. The installation of position tracers (based for example on RFID technology) makes it possible to increase the precision of the

determination of the relative positions. And if certain sensors (electroencephalography, sensors of electrodermal activity) remain to this day still too invasive, it seems possible that their miniaturization will soon lead them out of experimental laboratories and into classrooms.

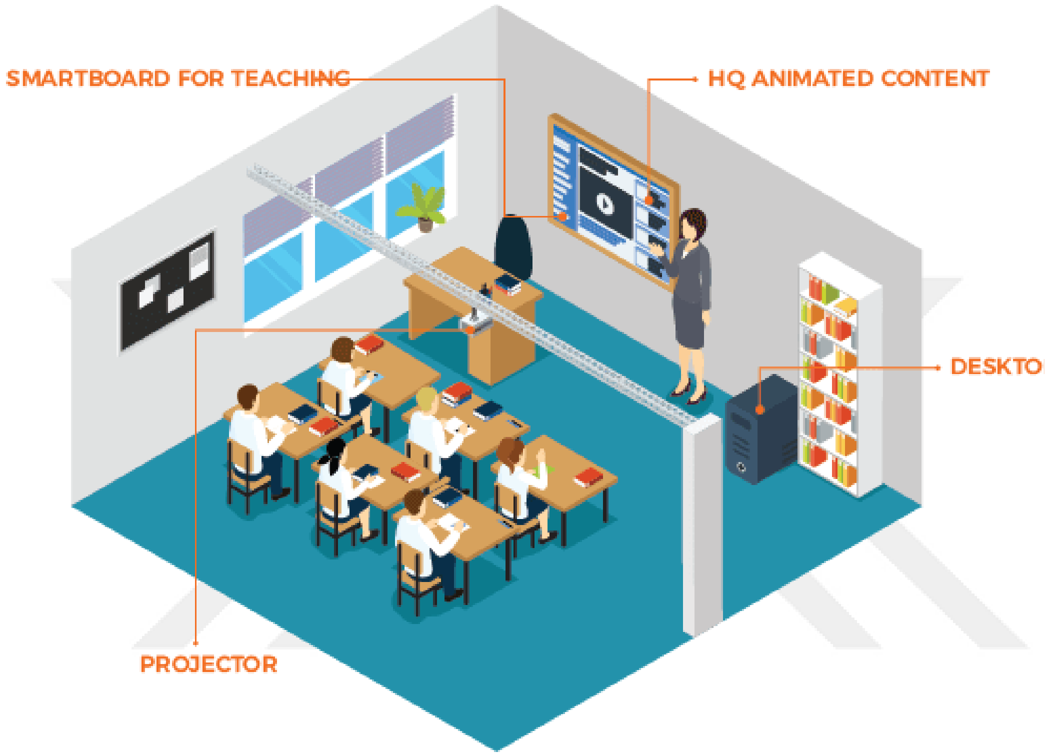


Figure 1.1: Smart classrooms with IT management [11]

In other words, smart classrooms are augmentations of labs, including educational sciences and cognitive psychology, in the authentic context of classrooms. As such, they make it possible to measure phenomena whose capture in the laboratory has hitherto violated the ecology of situations. They therefore open up the prospect of considerably enriching our understanding of the processes involved in teaching and student activities.

If they are currently more widespread in higher education, their decreasing cost as well as the richness of the analyzes offered make them usable at any level of the educational continuum, as demonstrated by Montebello (2019). However, most of the experiments conducted in the United States of America or in the People's Republic of China are always presented “for the good of the students”, and are surrounded by ethical precautions which seem insufficient to us. After presenting a typology of the different types of smart rooms, we will examine the risks of switching, intentional or not, from the adaptive monitoring of students to the monitoring of their bodies and minds.

1.1.2 A Typology of Smart Classrooms

The term "intelligent class" has covered, over time, classes with very different functionalities. In the 2000s, a smart classroom was a classroom where information could be transmitted from one class to another (Abowd, 1999), for example for distance learning purposes; or even a class where information from an interactive whiteboard is disseminated. Then appeared more recently, still under the same designation, classes equipped with computers and interconnected whiteboards, allowing the tracing of student activity by the teacher, in the form of dashboards. Even more recently, and thanks to advances in research in signal processing, rooms tracing characteristics related to teacher-student interactions are built and tested.

A. Tool rooms

The “tool rooms”make it possible to manage low-level tasks, such as controlling the lighting atmosphere or the temperature (like the iClass, cf. Ramadan, 2010). They are also able to facilitate the work of the teacher and the learners (the camera follows the teacher or zooms in on the area of the screen watched by the teacher, to retransmit it in videoconference). The Open Smart Classroom (Suo et al., 2009) offers this type of functionality, even augmenting it with a Chinese-Japanese machine translation system, thus allowing students separated by the language barrier to chat with each other. From a distance. Toolrooms can more prosaically authorize the automation of so-called administrative tasks, such as counting students,

B. Dashboard rooms

The room-dashboards" allow the teacher to supervise the activity of the pupils, by means of computerized dashboards, tracing their activity. This can be perceptible through computer traces (and therefore available in face-to-face or distance modalities) or through analog traces, resulting from the physical behavior of the students: what are they involved in? What emotions do they feel? Are they confused? Enthusiasts? Attentive? The sensors of these classes are able to inform their users and researchers about a large number of phenomena: the positions of the participants, their movement, their posture, their gestures, their attention, their commitment to the task (computerized or not) and even an inference of their emotions, perceptible through their facial expressions. In addition to these visual data, the audio sensors are able to determine the sound environment, the distribution of the sound, the prosody of the interactants, even their verbalization, through automatic speech recognition techniques (speech to text). Combined (therefore multimodal), these data offer an unprecedented descriptive power in the perception, by the machine and then by the human, of what is happening in the classrooms.

A high school in Hangzou, in the People's Republic of China, uses computer vision to analyze the emotions and actions of its students (Ngoc Anh et al., 2019). A dashboard informs the teacher in real time, in order to allow him to adapt his teaching as closely as possible to the activity of his pupils. But we could go even further, and assign each participant a complete dashboard of their activity in class, in particular to evaluate it. The PLACE device (Tissenbaum & Slotta, 2019) is another illustration of these dashboard rooms: students, in groups, solve problems by displaying the procedures and solutions on their shared television, which can also be seen by the teacher. , who plays the role of a “wandering facilitator”. The teacher also has a tablet which reports, group by group, the results of the exercises and their traces, the system playing the role of both a distributor of exercises and a supervisor of the progress of the pupils. It is noteworthy that increasingly important parts of the teacher's organizer-facilitator role could be delegated to the machine, as in the ClassMate device (Leonidis, 2010), which monitors and supervises the delivery of educational content. Finally, the Smart Classroom System (Gligoric et al., 2015) monitors the students' degree of attention in real time, by analyzing their eye movements and their oral production (analysis of the characteristics of speaking turns), the movement of the teacher also being recovered with an accelerometer.

C. Thermometer rooms

The “thermometer rooms”analyze the aforementioned elements of the context of the room to better characterize it, mainly for the purposes of feedback for to give experience to teachers, or for the validation of empirical constructs, when entrusted to researchers. This last type of room is only distinguished from the previous one by the role entrusted to the effectors, thus deciding the temporality of the feedback: immediate for the dashboard rooms, deferred in the thermometer rooms, influencing in our opinion differently the ecology of the scene and the integrity of its participants. This type of room therefore studies a few parameters of what could be called the "class climate", i.e. the socio-emotional level of teacher-student interactions, interactions that are very difficult to observe directly,

Thus the ACORN system (Automatic Classroom Observation Recognition Network) of Ramakrishnan et al. (2020) proposes to analyze audio-video recordings of class sequences in order to measure two characteristic dimensions of the class climate, respectively the positive climate (defined by warmth, respect and pleasure communicated by verbal and non-verbal interactions) and negative (defined by the overall level of negativity expressed in the classroom), using object recognition, emotion and speech analysis techniques. Interestingly, James et al. (2019) arrived at a similar result with a much more frugal approach, based solely on verbal and prosodic cues picked up by a microphone carried by the teacher.

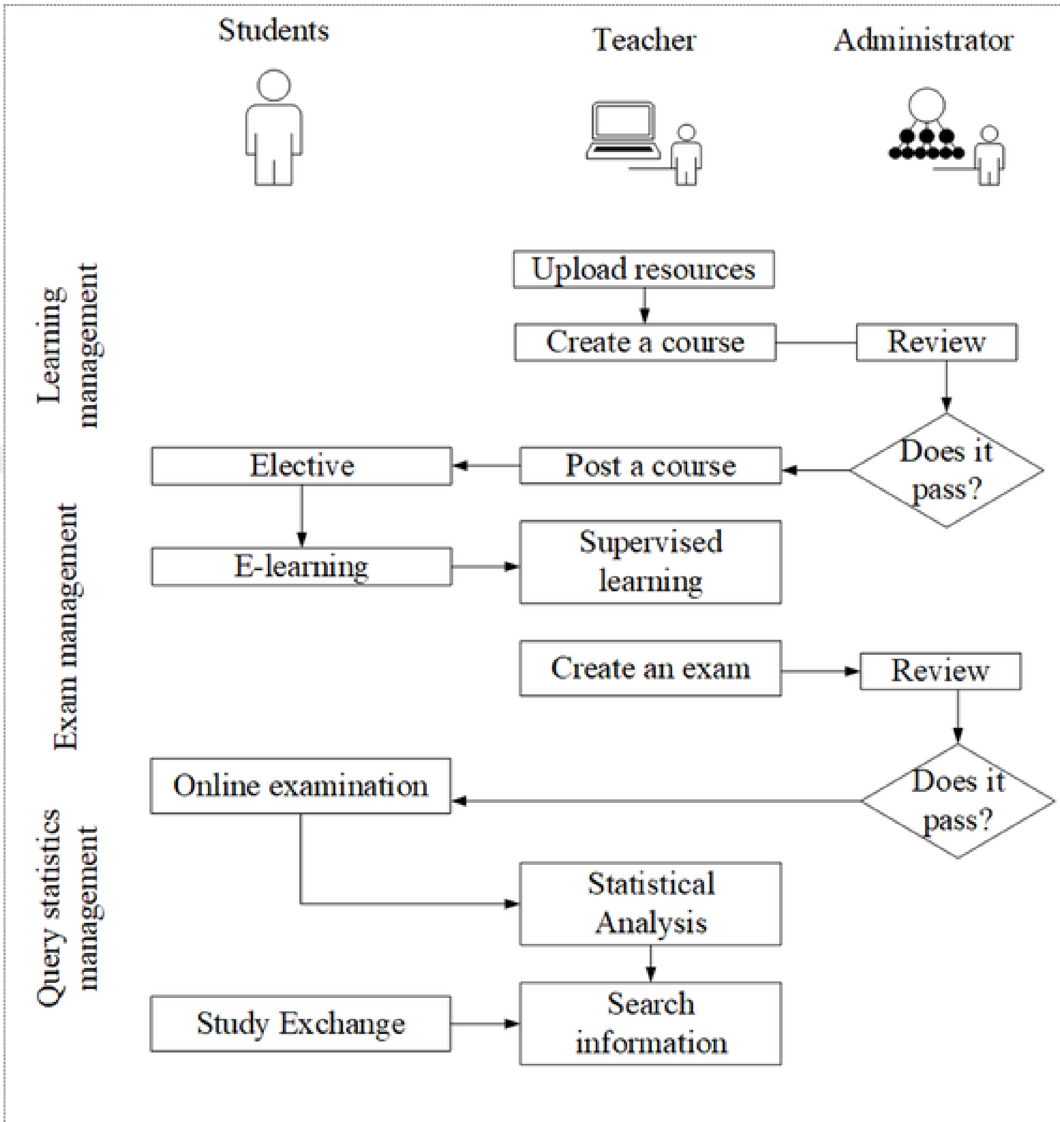


Figure 1.2: Topology of SCs [5]

1.1.3 What Should Context-Sensitive Classrooms Be Used For

The use of 2SCs is therefore wide, from the simple reaction to a stimulus (adjustment of light, recognition of student identity) to panoptic supervision of class activity, i.e. to adapt in real time the rhythm of the class, or to enrich the retrospective reflection around a session. For Montebello (2019), smart classrooms are likely to offer new learning opportunities, in addition to adapting to the differences between learners. They would thereby alter the professional roles and responsibilities of the teacher. Beyond the incantatory aspect of such a hypothesis, moreover poorly explained ("the role of educators and the educators themselves must change and adapt to adapt to a new era" (our translation, id ., p. 123), the demonstrative power of the Montebello review (2019), as well as those of studies relating to smart classrooms, can only be judged by the yardstick of two insurmountable criteria: their usefulness for the participants and their ethics , both conditions of their acceptability for the participants, and therefore of their ecology, essential to the smooth running of the teaching and learning processes.

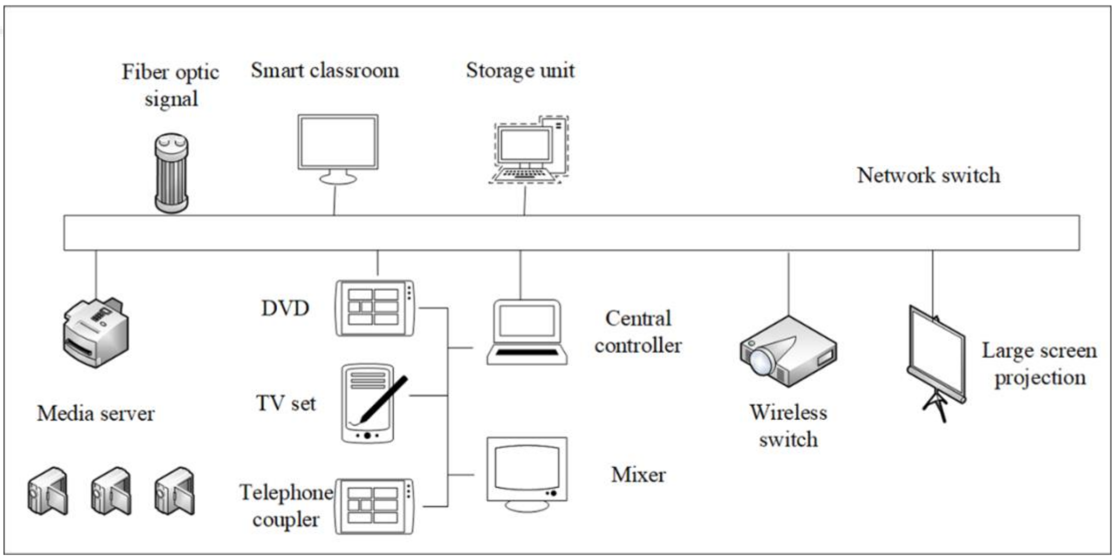


Figure 1.3: SCs connected devices [5]

We know that these criteria are burdened by the biases of facial recognition algorithms (eg, Zou & Schiebinger, 2018), the latter reproducing the biases of the humans who trained them, but more generally we can wonder about two roles of 2SCs: monitor or inform practices? as well as those of studies relating to smart classrooms can, in our opinion, only be judged on the basis of two insurmountable criteria: their usefulness for the participants and their ethics, both conditions of their acceptability for the participants, and therefore of their ecology, which is essential for the smooth running of the teaching and learning processes. We know that these criteria are burdened by the biases of facial recognition algorithms (eg, Zou & Schiebinger, 2018), the latter reproducing the biases of the humans who trained them, but more generally we can wonder about two roles of 2SCs: monitor or inform practices? as well as those of studies relating to smart classrooms can, in our opinion, only be judged on the basis of two insurmountable criteria: their usefulness for the participants and their ethics, both conditions of their acceptability for the participants, and therefore of their ecology, which is essential for the smooth running of the teaching and learning processes. We know that these criteria are burdened by the biases of facial recognition algorithms (eg, Zou & Schiebinger, 2018), the latter reproducing the biases of the humans who trained them, but more generally we can wonder about two roles of 2SCs: monitor or inform practices? Both conditions of their acceptability for the participants, and therefore of their ecology, essential to the smooth running of the teaching and learning processes. We know that these criteria are burdened by the biases of facial recognition algorithms (eg, Zou & Schiebinger, 2018), the latter reproducing the biases of the humans who trained them, but more generally we can wonder about two roles of 2SCs: monitor or inform practices? Both conditions of their acceptability for the participants, and therefore of their ecology, essential to the smooth running of the teaching and learning processes. We know that these criteria are burdened by the biases of facial recognition algorithms (eg, Zou & Schiebinger, 2018), the latter reproducing the biases of the humans who trained them, but more generally we can wonder about two roles of 2SCs: monitor or inform practices?

A. To watch

The opportunities offered by 2SCs to better understand the phenomena at work in learning (attention, emotions, interactions) could prove to be counter-productive when their initiators (primarily the students) feel oppressed by the capture and/or processing devices. Similarly, from a support instrument for the individualization of teaching and feedback enriched by conclusive data, these devices can slip imperceptibly into the control and monitoring of participants. Let's just suppose that a monitoring loop informs the teacher in real time of the emotions and the distribution of students' attention, it is possible that this loop is used cautiously by the teacher, but it is also possible that this loop inconveniences learners, who would feel hunted in their heart of hearts. This perceived surveillance could therefore feed them into legitimate behaviors of simulated conformation, avoidance, masking or cheating (Crooks, 2019). What would the pedagogical relationship have to gain? What would teachers feel if the data on the atmosphere, interactions and attentions of their class were reported to their hierarchical supervision, for example in order to instruct inspection feedback? Because socio-emotional interactions are crucial in the teaching-learning process (Creemers & Tillema, 1987), their authenticity must be safeguarded, and the sensitive device materialize this integrity. Moreover, it is not enough for the system to be "integrated" in its purposes,

B. enlightens

Thus, we propose, following Martinez-Maldonado et al. (2020), that these rooms, like Learning Analytics, adopt a translucent approach to design. This translucence can be based on three pillars. First, obtaining the informed consent of the owners of this data, first and foremost the learners. Informed means that they must be informed in understandable terms of what is captured, then of the treatments and analyzes applied. This precaution, by giving a view through the device, makes it possible to shed light on the issues and purposes. Then the renunciation of all immediate feedback, in order not to risk disturbing the ecology of the class, characterized in particular by Doyle (2011) as unpredictable, this unpredictability may represent a condition of the sincere commitment of teachers and students. Finally by obfuscating the data, that is to say by irrevocably anonymizing their source. This obfuscation is moreover within the reach of the same artificial intelligence techniques as those implemented to analyze the data, provided they were designed in this way (eg, Petrova et al., 2020).
On the other hand, this protective obfuscation of the psychological, physical and pedagogical integrity of the protagonists as well as the temporalization of the feedback to the teacher could sound the death knell for the individualization of real-time teaching assisted by the machine, out of traces. Perceived objectives in ILEs. Moreover, the pedagogical differentiation induced by the feedback mediated by artificial intelligence, by making the comparison between students more immediate and more visible, would probably prove to be counterproductive (Galand, Hospel, & Baudoin, 2014), whereas a "classic" differentiation, less intrusive and more respectful of the privacy of the protagonists (high expectations combined with a diversification of the types of activities assessed, defense, positivity, variations in groupings,
Thus constrained, smart rooms are likely to become instruments for understanding what is at stake (and probably what is at stake) in classrooms. By retrospectively enlightening teachers and their practice of evidence-based data, by informing researchers of these results, in other

words by feeding a loop of research-action feedback beneficial to all the protagonists, these rooms could become ecological laboratories that respect their participants and their environment. The richness of teacher-learner relationships.

1.2 PROBLEM STATEMENT

Numerous daily duties need the use of multi-criteria decision-making procedures (MCDM), also known as multiple attribute decision-making (MADM), to resolve choice dilemmas and decide the most effective course of action given a variety of criteria and prospective solutions. This is required for the completion of countless other daily duties.

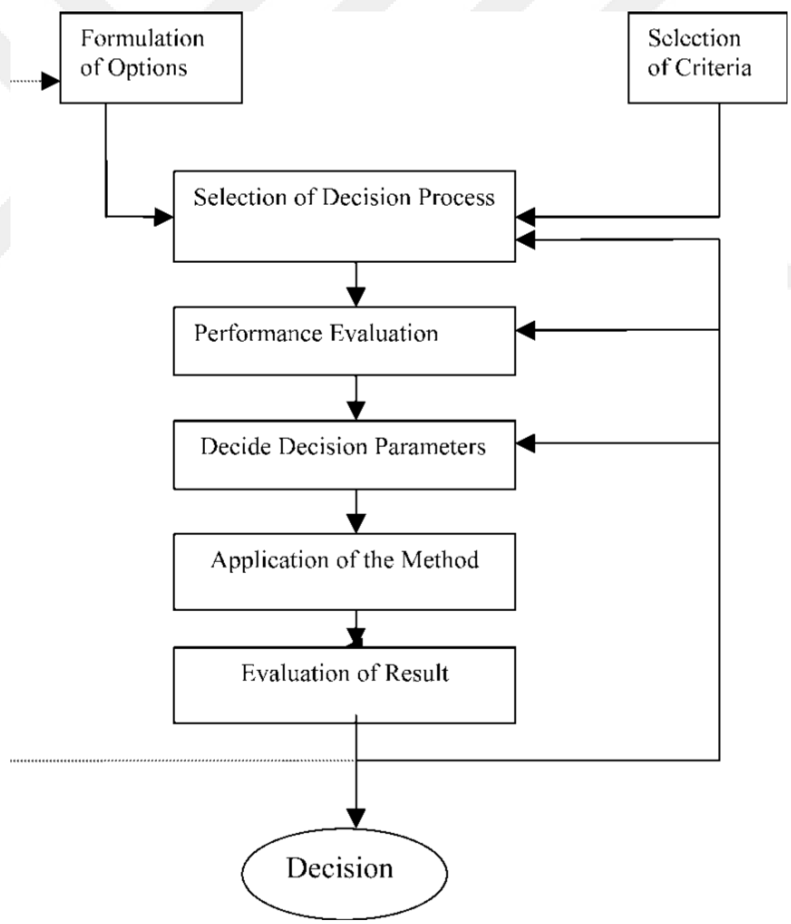


Figure 1.4: multi-criteria decision-making procedures [6]

A decision matrix for a multi-criteria decision-making (MCDM) problem is made up of a finite number of options Ai (I = 1, m), a set of criteria Cj (j = 1, n), their relative weights, and matrix cell elements (rij), which reflect the rating for alternative Ai in relation to criteria Cj. The process of aggregating and ranking may be very complicated because each criterion may be stated in a number of qualitative and quantitative units, such as size (qualitative), degrees, kilos, or meters. The reason behind this is that each criterion may be expressed in a different unit. Therefore, normalization is required to produce dimensionless and similar criterion values in preparation for aggregation into a final score. To reiterate quickly, the first step in modelling and applying the vast majority of MCDM strategies to address choice issues is to choose an appropriate normalization technique. In published investigations, several normalizing strategies of diverse sorts have been provided. 31 approaches for normalizing data to provide metric-free criteria for use in MCDM decision making issues have been identified. Despite the widespread use of MCDM approaches, the research community has not yet settled on a single normalization strategy that is superior in every circumstance. Because normalization changes criterion values to interval values while maintaining about the same magnitude, this technique may result in varied rankings for the several options available. As a direct result of this, the order or rating of things will change. If the normalization approach employed for a decision problem is not well-suited to the MCDM strategy in use, the best choice might be disregarded. The concept of offering proper normalization techniques for use with MCDM approaches would increase the reliability of decision-making solutions. Unfortunately, almost no study has been done on the analysis of normalization procedures for MCDM approaches. Consequently, the task of identifying and proposing the optimal normalization technique that is, the one that most precisely represents the input or raw data remains unsolved. In this research, the consequences of using four distinct normalization methods were examined. The SAW method, which decides which normalization method is the most effective by using a Ranking Consistency Index (RCI). In another investigation, the TOPSIS approach was used with four different normalization techniques. Using consistency criteria, the researchers evaluated and contrasted the effects of each normalizing approach to determine whether or not each normalization method was appropriate.

1.3 OBJECTIVES

Enhancement of the E-learning process by using data mining to conduct an analysis of data currently existing in the E-learning environment in order to forecast student performance and provide instructors and students with recommendations based on those predictions. This section includes detailed details about the study's final design and methodology.

i. Consider the methods, procedures, and activities involved in data mining.
ii. Create the framework that will serve as the foundation for all data mining tools.
iii. Research and evaluate popular online learning environments.
iv. You must determine which data mining technique is the most successful.
v. Develop a system that can predict how well enrolled students will succeed in their courses (SPS).
vi. The goal of testing the SPS prototype, an anonymous sample of students from a Dammam institution will be employed.
vii. Utilize the SPS model to give youngsters with guidance and educators with forecasts.

1.4 CONTRIBUTION

The present work aimed to demonstrate the use of machine learning algorithms in the teaching-learning process through experiments performed on a set of synthetic data and in order to favor the reproduction of these experiments. Our results revealed patterns of information that would be impossible to obtain quickly without the support of machine learning techniques. Thus, it was possible to complete accessible instructions and materials for reproducing the experiments with the considered artificial intelligence techniques, in addition to others.

1.5 THESIS STRUCTURE

This thesis is composed of five sections. The following provides an overview of each chapter: The Introduction: An Overview This section defines the research question and hypothesis, as well as the issue statements and the motivation for doing the study. Then, an outline of the research approach is presented, followed by an introduction to the thesis's structure. Context and

Review of Related Literature in Chapter 2 In this chapter, we will explore the literary works that serve as the foundation for this study, as well as the knowledge required to resolve issues related to this thesis. It gives a taxonomy for MCDM approaches and other normalizing techniques, introduces and explores the relevance of these techniques in MCDM, and concentrates on the most prevalent normalization strategies now in use. The literature review includes previous publications addressing the evaluation of normalizing approaches in MCDM as well as suggestions for normalizing tactics for a variety of challenges. In addition, a number of topics pertaining to dynamic systems and collaborative networks are offered for your consideration. The proposed evaluation framework is provided, examined, and illustrated using examples and case studies in Chapter 3. This chapter's objectives are to make the issue more accessible and to provide a deeper understanding of the methodologies used. This chapter also includes well-known evaluation methods and metrics, along with the rationale for including them into the suggested assessment framework.

2. LITERATURE REVIEW

According to the conclusions of a literature review conducted by Zawacki-Richter et al. [1], multicultural student characteristics need more attention than they now get. Ferguson [2] discussed excellent ways to learning analytics. The term "learning analytics" refers to research on student behavior undertaken with the purpose of assessing academic progress and predicting future outcomes. Students participate in a variety of activities, including, but not limited to, online networking connections, quizzes, assignments, discussion board updates, and other activities. Similar to this, Romero and Ventura [3] focused their data mining investigation on academic records. In the course of their research, Papamitsiou and Economides [4] emphasized the many applications for learning analytics. Taking roll involves a significant amount of time. A substantial amount of study has been undertaken on the various methods for automatically collecting attendance. Patel and Priya [5] investigated student attendance concerns and investigated possible solutions based on RFID and facial recognition technologies. Just as traditional classrooms need a range of networked technology, so too does education provided through distant learning. This is required to guarantee that effective instruction is provided at both the main school and any satellite campuses. Recent research by Kele s et al. [6] focuses on the utilization of LMSs in combination with distant education (LMSs). An event has occurred upsurge in the use of augmented reality (AR) in classrooms recently, particularly in secondary schools. Chen et al. [7] have produced a paper on the use of augmented reality (AR) in "smart classrooms." [8] Given that education has been practiced for millennia, it is only natural that its many forms and techniques have been the subject of considerable discussion. The development of technology in the realm of education has a direct effect on the current educational methods. Although technology may be advantageous in a variety of ways, it may also make education more challenging. Therefore, it is of the highest significance to assess how each new technology and its associated pedagogical approach affects education as a whole. If a piece of technology is to be used in an educational setting, it must be able to explain how it will improve the learning environment. Moreover, a number of research investigating the impact of the employment of various technologies on the results of educational initiatives have been published. Glover et al. [9] performed study to determine how the usage of multimedia displays affects various teaching methods. According to the results of a study conducted by Higgins and colleagues [10], the use of digital whiteboards in specific educational contexts is on the rise. Martin et al. [11] investigated the needs of participants in a basic science program that used digital whiteboards. About half of the time, instructors utilize digital whiteboards, according to the survey's findings. According to the results of yet another poll [12], the vast nearly all educators feel assured about their abilities to use digital whiteboards in the classroom. Fies and Marshall [13] explored the challenges that may come from the usage of CRSs in the classroom. Students emailed their contributions to the instructor using a common format known as CRS. The responses are collected and graded by the teacher. Kay and LeSage [14] did a study that was quite comparable to this one, evaluating the impact of audience response systems (ARSs). After digital questions have been provided, an audience response system (ARS) allows participants to respond using their mobile devices. In addition to the widespread use of mobile computing devices such as tablets, iPads, and smartphones, Georgiev et al. [15] suggested "Mlearning" (mobile learning). [Bibliography required] Wang et al. [16] investigated the influence of mobile education on students' final grades. The authors argue that the usage of mobile devices for educational purposes will increase students' engagement and level of interest in class. Wu et al. [17] found that the frequency of M-learning in secondary schools was larger than in elementary schools. According to Abachi and Muhammad [18], people in the teaching profession and those in the learning profession embrace m-learning. In a similar line,

Ha and Kim [19] performed research and found that the use of Twitter microblogging on mobile phones for academic engagement boosts student achievement. According to the results of a research done by Parker and Burnie [20], a growing proportion of students are using digital and online resources to acquire business skills. According to the authors, business schools use multimedia technology often because it boosts students' learning experiences and teachers' efficacy. Zhou [21] performed study to evaluate the hypothesis that smart classrooms in which students have the ability to take an active part in their own education result in higher academic performance. The influence of a smart classroom environment on pupils' ability to learn has been the topic of a number of further studies. Abdul et al [22]. Further classify EDM techniques into six groups: data extraction, prediction, connection mining, structure discovery, and model discovery, in addition to hybrid methods. They showed the ML or DM algorithms that were used in each category, organized some prior research according to the EDM techniques they used, and said that the most generally used approach for prediction is the EDM method. This technique employs the machine learning (ML) classification, regression, and density estimation methods. Said et al [23]. Offered an overview of the EDM methodologies used in over ten research in EDM domains between 2016 and 2019 and addressed the educational difficulties that these studies sought to address. Said et al. also revealed the particular methods utilized to predict students' success in these studies. These algorithms consist of, among others, decision tree (DT), support vector machine (SVM), nave Bayes (NB), logistic regression (LR), and artificial neural network (ANN). The authors Hanan et al [24]. used the literature survey method to summarize close to 400 studies published between 2000 and 2017 in the fields of computer supported learning analysis (CSLA), computer supported predictive analysis (CSPA), computer supported behavioral analysis (CSBA), and computer supported visualization analysis (CSVA). Then, they provided a list of ten EDM solutions for difficulties in the realm of education. Classification, grouping, data visualization, statistics, association rules, and regression analysis were among these techniques. In addition, they claimed that ML's classification system provides accurate estimates of academic progress. The authors Ashish et al. utilized the approach of literature review to synthesize over 160 publications that used clustering algorithm in the area of EDM during the last three decades (1983–2016). In addition, they identified the educational challenges that academics attempted to tackle in these studies, the particular algorithms used, the datasets utilized, and the datasets' sources [25]. They divided the educational challenges that the clustering algorithm may tackle into five distinct categories: evaluating the motivation and conduct of students, identifying individual learning styles, using digital resources to study, and acquiring information together. Angelista. did an analysis of more than 300 ongoing research on the application of soft computing approaches in EDM. Among these methods are ANN, DT, Bayesian, random forest (RF), and support vector machines (SVM). They listed research into the application of soft computing techniques in various contexts, such as evaluating and predicting student performance in learning, evaluating the quality of user interaction with systems, developing personalized recommendations for students' education, and enhancing managerial decision-making [26]. More than sixty percent of EDM research has been conducted on the topic of forecasting student performance using a classification algorithm, indicating that this is one of EDM's greatest obstacles [27]. The statistical findings of this investigation indicate that this is the case. Romero et al. offered a thorough analysis of EDM research, introducing the most important international conferences, vital publications, and highly cited works in the field. In addition, he provided a succinct explanation of the overall structure of EDM, as well as the data produced by various educational contexts and the most important instruments used. Numerous researches have studied whether EDM is beneficial for correctly predicting students' performance levels. A literature assessment of DM approaches for predicting student achievement from 2002-201510 revealed that students' grade point average (GPA), homework performance, quiz scores, class attendance, and other kinds of self-evaluation were the most influential indicators [28]. This was followed by the gender, age, socioeconomic standing, and other demographics of the students. Machine learning algorithms, including as DT, ANN, naive Bayes, SVM, and soon, are often used in the process of developing a prediction model of students' performance. Using the SLR approach, Abdallah et al. evaluated 62 publications from 2020 to 2020. These articles were retrieved from IEEE Xplore, web of science, ACM, and more databases. Their objectives were to establish (1) how learning outcomes are predicted, (2) how predictive analytics models are created, and (3) which elements have the most impact on learning outcomes. Saa et al. [31] conducted a comprehensive examination of 36 out of 420 research publications published between 2019 and 2018. The authors of this paper conducted their analysis using the SLR method. According to the study's results, the most influential criteria fall into four main categories: students' past academic accomplishment, their e-Learning activity, their demographic and social information, and their instructors' pedagogical techniques. Prior academic success is the most significant aspect to consider. Garcia et al. [32] presented a prediction approach based on the GARCH model. As a result, the major objective of this research is to predict one day in advance the times of day when the price of energy would fluctuate significantly. During the pilot phase, the method was evaluated in unrestricted marketplaces in continental Spain and the state of California. It is also crucial to note that Malo ET altechnique, [33].'S presentation was significant since it was related to the varying power costs over time. In the Nordic power markets, multivariate GARCH models are used for dynamic hedging, and the author of this study investigates a range of specification tests for these models. In order to evaluate the success of the hedging strategy, the unconditional and conditional ex-post variances were also compared. Weron et al. [34] offered a total of twelve parametric and semi-parametric time series models that may be used to forecast the price of energy. In this study, forecasting intervals were supplied and analyzed, and both conditional and unconditional coverage was considered. They concluded that the confidence intervals generated by semi-parametric models were superior to those generated by parametric models. Near the turn of the century, it was commonplace to produce forecasts using linear approaches. In the modern society, however, such actions function as standards by which others are assessed.

Taylor et al. [35] examined six distinct univariate time series algorithms with the purpose of forecasting the quantity of energy that customers in Rio de Janeiro and England and Wales would use. These included an autoregressive integrated moving average (ARIMA) model and exponential smoothing (both for double seasonality), an artificial neural network, a regression model that used a previous principal component analysis, and two naïve techniques as references. The exponential smoothing strategy was the most effective, and the regression model performed well in its capacity to predict demand in England and Wales. Neupane et al. [36] created their model using an ANN, and they carefully picked the inputs. To make a determination based on these inputs, a wrapper approach for feature selection was used. The concept was assessed using data from the power markets in Australia, New York, and Spain; it was determined that the results acquired using this approach were superior to those obtained using the PSF technique. Attempts have also been made to solve the challenge of selecting input characteristics for load forecasting by using an artificial neural network (ANN) [37]. Using the latest and most cutting-edge prediction algorithms, the authors examined the performance of four alternative feature selection processes. The data utilized in this research were collected in Australia during a two-year period. The actual findings were much superior than those predicted by the exponential smoothing models. Li et al. [38] addressed the nonstationarity of the load time series by proposing a wavelet transform and an ELM with weights initially obtained by an artificial bee colony approach to anticipate the load time series in New England and North America based on the wavelet series. This was done to better comprehend the nonstationary character of the load time series. The authors demonstrated that predicting mistakes might be decreased by utilizing an optimization technique to define the weights in ELM. Ismail et al. [39] presented a rule-based mathematical model that can predict electrical system peak load demand. This model is used for forecasting. In this inquiry, Malaysian information was used. For an appropriate evaluation of the outcomes, both regression and SARIMA models were used. Utilizing the wavelet transform and ARIMA models, Conejo et al. [40] suggested a novel method for forecasting future power price that takes use of the wavelet transform. Consequently, they used the wavelet transform to deconstruct the time series into a collection of component series that behaved more appropriately. After implementing the wavelet transform in reverse, ARIMA models were used to forecast future values for the new series. This method represents a substantial improvement over their previously revealed efforts, which can be seen here. Aggarwal et al. [41] also did research and made projections on the price of power. To do this, each day was segmented and a multiple linear regression (MLR) was run on either the original series or the component series generated by the wavelet transform. This enabled them to identify the most correct series. In addition, the variables included as inputs to the regression model varied per category. Pindoriya et al. [42] developed an adaptive wavelet-based neural network with the intent of forecasting short-term power price time series for the Spanish and Californian markets (AWNN). The neurons in the neural network's hidden layer relied on wavelets, the shape of which changed dependent on the information being used for training. As a consequence of their success in modelling non-stationary and high-frequency signals using their approach, the authors concluded that their method converged more rapidly and outperformed competing strategies in predicting future electricity prices. PJM was the intended audience for this communication.

3. MATERIALS AND METHODS

3.1 INFORMATION AND COMMUNICATION TECHNOLOGIES FOR EDUCATION

Information and Communication Technologies offers its services to many sectors, including the one that interests us: education. This is how a new expression appeared in 1992: Information and Communication Technologies for Education (TICE). In this case, it is a question of processing the information content of a digital document for teaching and learning purposes by applying the processes common to ICT. Thierry Karsenti says about ICTE that they "break the model of integrated teaching around three pillars: — unity of place (classroom),

i. Unit of time (the same timetable for everyone),
ii. Unity of action (same content according to the same terms)".

It seems necessary to recall that we are really talking here about tools for information and communication technologies. In other words, we are interested in educational actions integrating the use of digital tools. Indeed, ICTE is not intended to have lessons evaluated by changing the essential content, but rather the didactic approaches that allow the acquisition of knowledge. We therefore discuss the general pedagogical added value(s). The pedagogical added value is defined as the capacity of these to modify the terms of the didactic contract which governs the triptych pupil, teacher, knowledge during an activity. The contributions of technology are measurable, thanks to the enhancement of shared resources, the collaboration of the educational community, the gain in autonomy, the development of hypothetico-deductive thinking and therefore of critical thinking.

Planning Lessons with ICT	A lot	Somewhat	A little	None
-Write a text using Microsoft Word processing programme.	41,79 %	17,91 %	14,92 %	25,37 %
-Email a file to another teacher and send messages.	17,91 %	17,91 %	13,43 %	38,80 %
-Create PowerPoint presentation with simple animation of pictures or video and integrate into the lesson activities.	11,94 %	14,92 %	14,92 %	46,26 %
-Participate in educational discussions in forums, blogs social networks on the net like Facebook or Twitter.	8,95 %	5,97 %	13,43 %	59,70 %
-Download and install software on a computer	11,94 %	8,95 %	23.38 %	41.79 %

Planning Lessons with ICT

Figure 3.1: planning lessons with Information and Communication Technologies [22]

ICT have empirically demonstrated their interest in producing better quality resources and increasing student motivation. Numerous polls and surveys carried out with teachers and education professionals have mainly corroborated this interest in integrating ICT in teaching. Playful and interactive, they are stimulating for students. Most of the research granted to ICTE the same pedagogical added value: they are practical when the pedagogical scenario is mastered and allows the productivity of the pupils.

3.1.1 Online Learning

As can be observed, the emergence of ICTE induces changes in teaching practices, teachers therefore having to appropriate collaborative work environments which functionalize the continuity of training outside the space-time of the class. . After a difficult start in the 1990s, e-learning is now a device present in several training models and regardless of the public, school or professional. E-learning refers to anything that broadcasts, interacts, or communicates using a local or wide area network or the Internet, such as distance education, access to sources through downloading, or consultation on the Internet. It may include taught systems, self-training systems, synchronous or asynchronous systems, or a combination of the aforementioned aspects.3.1.2 Data Mining in Education.

Web technologies have thus enabled the development of so-called hybrid courses, i.e. based on both face-to-face courses and a digital work environment. Given the ongoing digital transformation of our training, educational platforms such as Moodle are becoming privileged teaching-learning spaces within the University and they are full of information on the uses and profiles of our students. These data are commonly called digital traces, a kind of logging of the actions carried out by our users through the services and applications available to them. In Educational Sciences, as in most disciplines, data sources are a valuable raw material for conducting empirical studies. Higher Education has associated big data techniques with its reflective practice for the last decade, speaking of Educational Data Mining. This consists of using the massive data collected in our information system. Our Observatory of Digital Uses has set up a set of dashboards to better manage our services, make optimization decisions in line with uses. By cross-referencing all of the institution's data sources, we are able to extract major trends that promote understanding of our environments and our educational activities to identify their impact on the behaviour and learning of our students. At the same time, we must respond to a request from our user population.

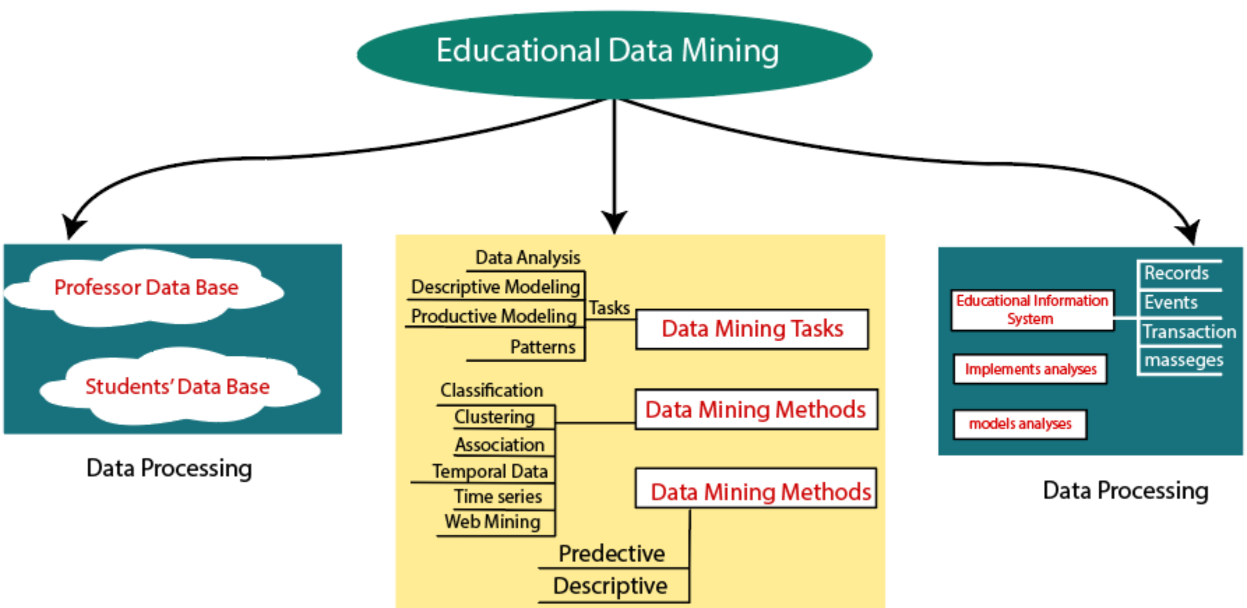


Figure 3.2: Applications of datamining in educational aspects [15]

3.1.3 Virtual Classrooms

Indeed, in this context, a new virtual classroom space is described, designed by our teachers for the use of Moodle. With regard to what is done in direct face-to-face, the class is led by the teacher who is able to adapt his practices according to his observations of these moments of class, the human-to-human interactions that are played out, student apprehension. However, through our educational platforms, the relationship is different and calls on the human-machine interface. One of the biases of e-learning is to believe that putting everything online is enough to "learn", even though there is confusion between published information and accessible information. The teacher then needs "tools to help manage" his virtual classroom since part of the interactions are done through the instrumentation of Moodle. Supiot tells us that the supervision activity is replaced by "the digitization and traceability of data replacing the eye of the foreman". As we noted above, the actions of learners on their digital work environment are logged in our databases. A posteriori, the teacher can therefore know exactly what has been achieved and how. He must therefore be able to support his learners in this teaching-learning model integrating ICT. The objective of this course is therefore to contribute to this effect through a device presenting in a synthetic and effective way the actions of the learners in a course space. Pedagogical teams must therefore be able to consult the traces left by their students in a readable way to better understand the challenges of digitization and the role of tutor assigned to them.

3.2 ARTIFICIAL INTELLIGENCE

The first summer conference on artificial intelligence (AI) took place at Dartmouth College in New Hampshire, USA, in 1956. John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon were the conference's organizers. This conference's proposal called for "ten men, to conduct a research on the subject of artificial intelligence for two months." Since its inception, AI has caused controversy, from the definition of its goals and methodology to the very name of the field, which some have deemed pretentious. Exaggerated promises and related disappointments occasionally resulted from ignorance of the concepts underlying intelligence on the one side, and the practical limits of computer processing power on the other.

The merging of the fields of Artificial Intelligence (AI) and Distributed Systems gave rise to Distributed Artificial Intelligence (DAI) (SD). This integration distinguished the IAD as a branch of computing with unique characteristics rather than as a subfield of AI. The IAD is distinct from the SD field in that it aims to develop ways for collaboration amongst system participants rather than concentrating on problems with distributed processing and increasing computing efficiency. In contrast to AI, IAD offers fresh and more in-depth viewpoints on knowledge representation, organizing, coordinating, and solving problems.

The theory of IAD enables for research in two domains that are frequently studied: distributed problem solving (RDP) and multi-agent systems (MAS).

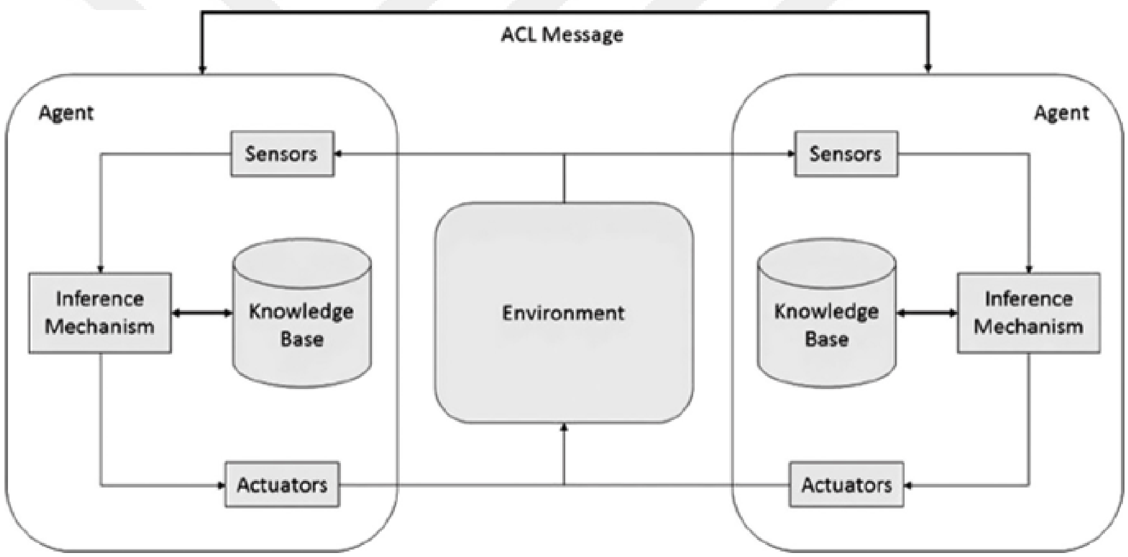


Figure 3.3: distributed artificial intelligence structure [12]

3.2.1 Distributed Problem Solving

Connects numerous agents in a robust and cohesive manner to tackle a certain challenge. The agents required to create the environment and produce the solution are identified from the problem. Its overall goal is to create the reasoning and knowledge representation methods required for problem-solving nodes connected by a network that is loosely linked work together successfully to solve a challenging distributed challenge (Bittencourt, 2001). SMAs, on the other hand, focus on the agent, its internal characteristics, and how it behaves in the environment.

Therefore, it can be claimed that the generality of the environment is the fundamental distinction between RDP and SMA. RDP's environment was created with a specific issue in mind,

Important aspects of the RDP approach include the fact that, even when agents cooperate, it is not necessary for them to explicitly state their skills and objectives. The designer implicitly represents this; in most cases, the designer makes all of the decisions about the description and breakdown of jobs. In most instances, if the job is divided appropriately by the designer, no conflict will arise even with some dynamic task decomposition; nevertheless, even in this extreme scenario, the methods utilized are largely reliant on the application domain. Even though agents can communicate, there is no need for a complicated discourse to accomplish the aim as this is a vital situation and any conflict would depend strongly on the application domain. No new agents may be dynamically becoming a part of society. As a result, this kind of system cannot be regarded as an open system (Sichman et al., 1992).

3.2.2 Multi-Agent System

In the IAD community, there is currently no agreement on what an agent is. Due to the lack of consensus, definitions are contextualized to fit the application's goals or the viewpoint of each researcher. For instance, Etzioni and Weld describe an agent as "anything that can sense its environment through sensors and act in this environment using actuators," which includes "a computer program that behaves in a manner akin to a human agent, such as a travel agent or an insurance agent.”.

In MAS, a group of agents' behaviour and interactions are first of concern. Although there is a wealth of literature on agent models, reactive and cognitive agents represent the two primary schools of thought.

Reactive agents don't have memory structures, complicated symbolic reasoning, or an explicit internal knowledge representation. With these limitations, a reactive agent can only perceive the external environment and responds in a way that has been predetermined by the programmer depending on environmental inputs. In a society of reactive agents, the system's intelligent behaviour results from the interaction of each agent's fundamental behaviours. This kind of society is based on illustrations of biological and ethological groups, such an ant colony. This form of society typically has a large number of agents, thousands or perhaps hundreds. The

Brooks subsumption architecture is one of many efforts in the literature on MAS based on reactive agents. Cognitive agents can plan their future activities because they have memory, an explicit representation of the environment and other agents, and a sophisticated system of cooperation. And communication. A small group of persons, typically no more than two or three dozen, come together to establish a society of cognitive agents.

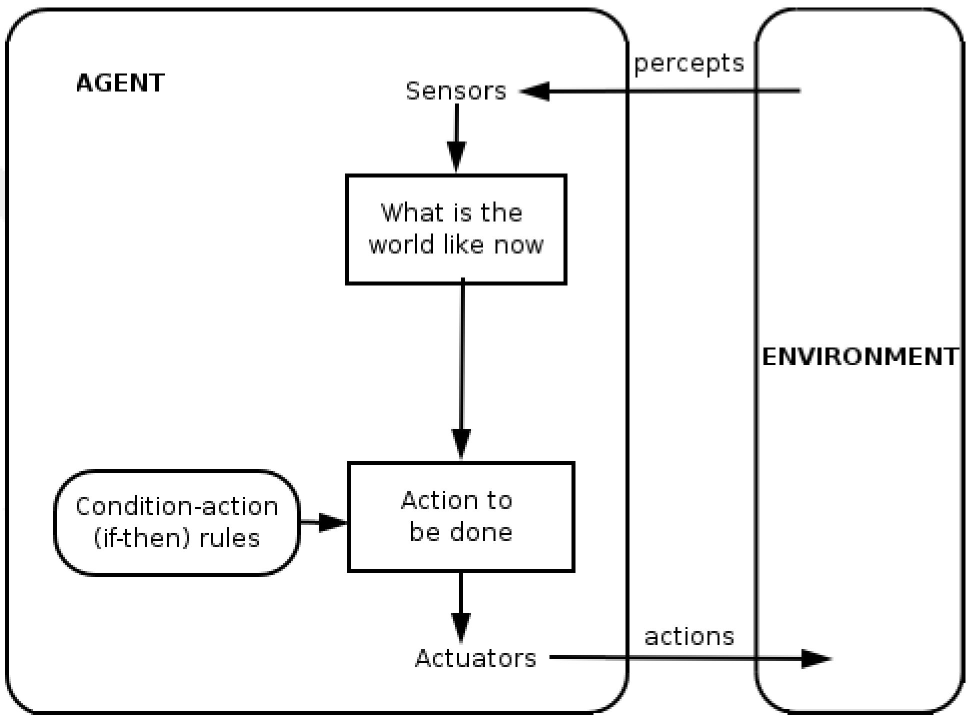


Figure 3.4: structure of MAG [41]

Typically, the metaphor of human social groups is used to structure this form of civilization, where teams of experts can work together to solve problems. Important aspects of the SMA approach include the fact that the agents, not the designer, decompose tasks. Agents are autonomous, meaning they can have their own local goals, and there can only be a little amount of dynamic reorganization, meaning that agents can choose what they can do to alter their behavior in order to better complete their responsibilities. As a result, conflicts can frequently occur because there are both local and global goals.

To clarify each agent's involvement in the efforts to solve the problem, complicated communication must also be organized.

When appropriate, an agent may join or quit the partnership. The other agents will add the new agent's capabilities and objectives to their knowledge base if they are introduced into society. This is done to maintain a clear portrayal of all agents' objectives and potential.

Agents must update their internal environment models whenever the environment changes. Mobile robots that can navigate through unfamiliar and foreign environments are an example of agents that must deal with a dynamic environment (Sichman et al., 1992).

In MAS, agents coexist in a shared environment and cooperate with one another to accomplish the goal. Graph: (agents) (problem to be solved) (problem solved) (solution)

A hybrid method, or the creation of an agent that is neither entirely cognitive nor totally reactive, has been discussed by several researchers. This method would involve an agent with two or more subsystems: a cognitive one that could develop plans and make decisions and have a symbolic representation of the world; and a reactive one that could respond to environmental occurrences. Without using sophisticated thinking.

3.2.3 Generic Model

There was a hybrid architecture displayed. The three levels of this generic model—reactive, instinctive, and cognitive—are functionally comparable to the three parts of the architecture used in Sloman's agent. The model is doubly devoted to the evolutionary approach: both the reactive level and the model's design, which allows the components at each level to evolve one after the other, are based on evolutionary mechanisms. Here is a quick rundown of each level.

The reactive level attempts to simulate a basic animal, like an insect. It is made up of an evolutionary environment with the following components: patterns taken from an information source.

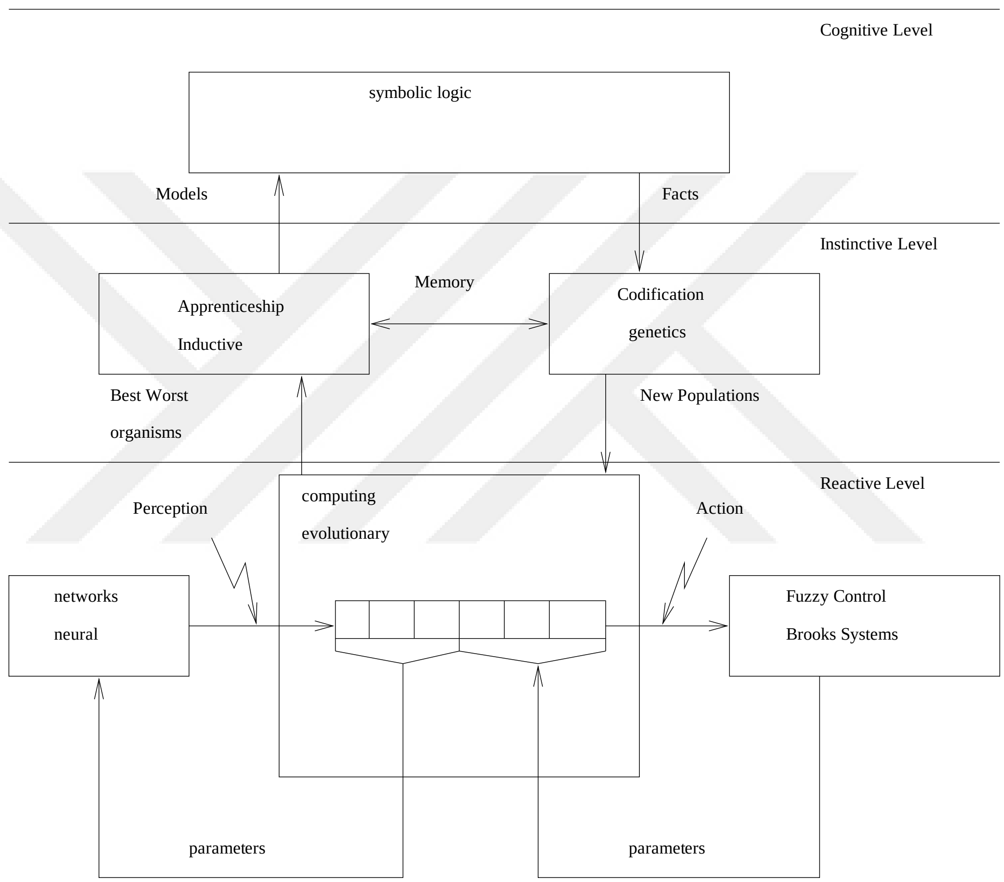


Figure 3.5: Generic Model for Cognitive Agents

A population of reactive agents that connect perception and action, causal control that results in activities in the same external environment, and an external environment. High parallel activity at this level causes a quick cycle of perception and action. The most effective community agents, as determined by the fitness function, are able to act at the conclusion of each cycle.

Long-term memory is incorporated into the model at the instinctive level. The populations in the environment that contribute to greater performance in a particular situation can be found using the reactive level. It is feasible to create a broad description of a particular population, a sort of reserve, by extracting the features of the best and worst agents from these populations using a fitness function genetics. These descriptions make up the majority of this memory, and the act of "memorization" entails the introduction of a new population at the reactive level, from which the agents are genetically coded in accordance with these generic descriptions, also formed at the reactive level. This memory's long-term effects are comparable to "breeding" and "domesticating" the reactive agent population. The instinctive level is characterized by a lengthier cycle of action that requires a few iterations of the same circumstance to be successful. The reactive and instinctual levels are meant to mimic more complex species, including mammals. The manipulation of descriptions created at the instinctive level is an issue of the cognitive level. This level is built around two complementing tasks: learning about important situation descriptions and coming up with fresh action plans. The primary benefit of the cognitive level is that it enables the development of a second language-based channel of communication between the agent and the environment. With the use of this language, abstract circumstances may be described, allowing us to conceive them and piece together the pertinent social relationships into a unified internal theory that can be accurately recreated at the reactive level.

3.3 MACHINE LEARNING

The goal of machine learning (ML), a branch of artificial intelligence, is to build systems that can learn on their own and develop computational models of learning processes. A learning system is a computer software that decides based on experiences gained from solving problems successfully in the past. The many Machine Learning systems can be categorized according to their description language, mode, paradigm, and learning style thanks to specific and shared properties. Learning involves a variety of activities, such as gaining new knowledge, improving motor and cognitive skills through practice and instruction, organizing knowledge, creating useful representations of knowledge, and discovering new facts and theories through learning from experiments and observations. The secret of human intelligence's superiority is learning. Man can pick up new motor and cognitive abilities. Even without vision, a person can learn the same motor abilities. Vision is the primary way to learn motor skills. The primary goal of ML is to research and computationally model learning processes in all of their various manifestations. Increased machine learning capabilities will lead to a considerable rise in artificial intelligence. Finding a general definition of learning can be challenging. The following list defines learning: Learning is the capacity to get better at something." For instance, this query may be used to identify a specific physiognomy. This term is behavioral in that it views the actor who develops himself as a black box and evaluates his performance from an outside perspective. The knowledge required to identify physiognomies, for instance, might be stated in terms of classification rules that enumerate the traits of a particular physiognomy.

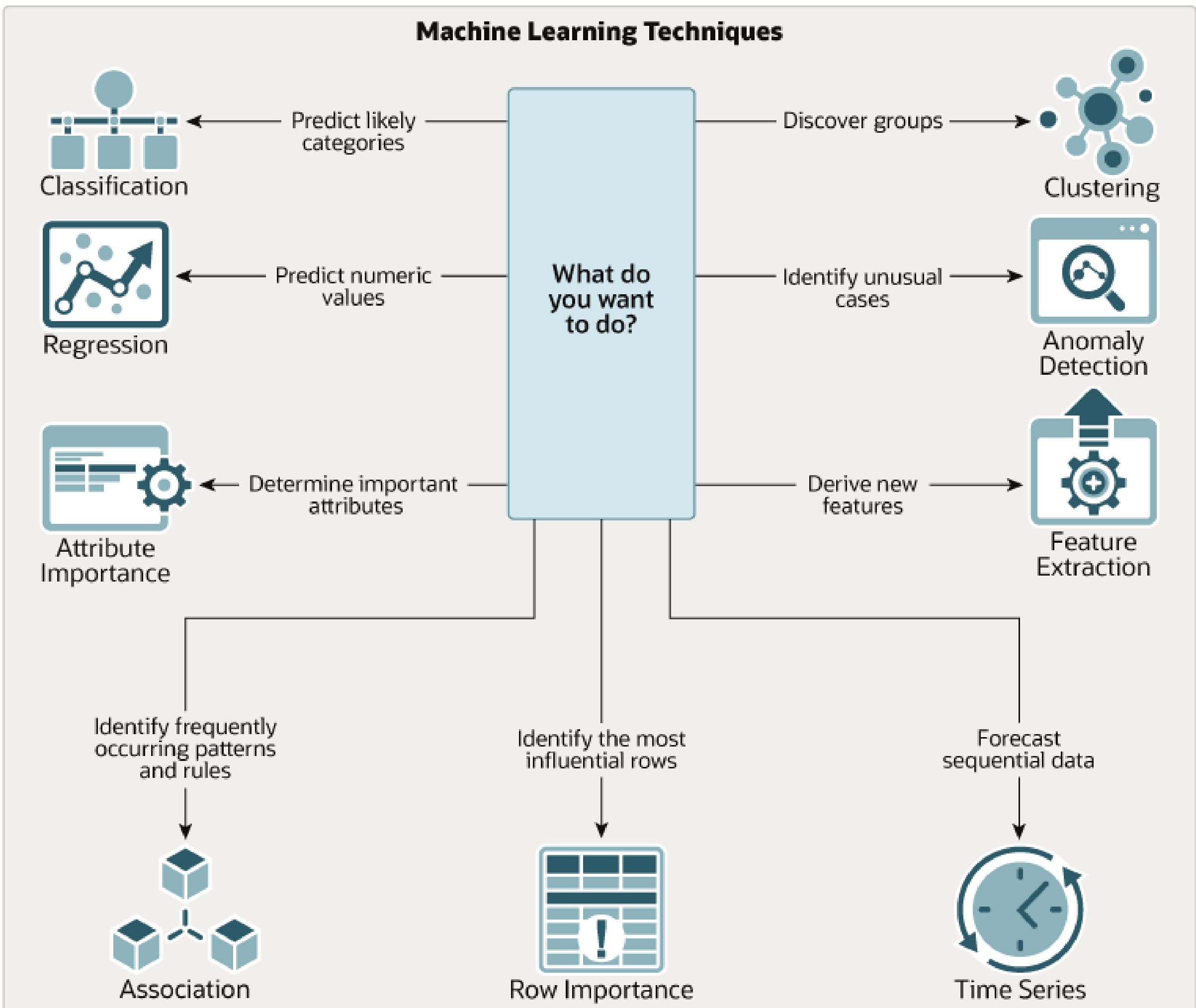


Figure 3.6: machine learning technologies [15]

The ability to recognize faces can be improved by improving classification rules. In this view, learning is the ability to acquire new knowledge. There are several ML paradigms, such as:

A. Symbolic: Symbolic learning systems seek to learn by constructing symbolic representations of a concept through the analysis of examples and counterexamples of that concept. Symbolic representations typically take the form of a logical expression, decision tree, rule set, or semantic network.

B. Statistical: Statistical researchers have created several classification methods, many of them similar to methods later developed by the ML community. The general idea is to use statistical models to find a good approximation of the concept to be learned. Among the statistical methods, Bayesian learning stands out, which uses probabilistic models based on prior knowledge of the problem.

C. Connectionist: Neural networks are mathematical constructions inspired by biological models of the nervous system. The biological metaphor with the neural connections of the nervous system has interested many researchers, and the analogies with biology have led many researchers to believe that neural networks have great potential in solving problems that require intense human sensory processing.

An interesting point about human beings is related to their ability to make accurate generalizations from facts. Humans are able to find patterns just by observing a real-world process. In computing, this can be obtained from a set of examples, provided by the user or by a real-world process, and through inductive inference, which, despite being one of the most used resources by the brain in the production of new knowledge, must be used carefully.

3.3.1 Inductive Learning

Induction is the form of logical inference that allows general conclusions to be drawn from particular examples. It is characterized as the reasoning that goes from the specific to the general, from the particular to the universal, from the part to the whole. In Inductive Learning, the learner acquires a concept by making inductive inferences about the facts presented. Hypotheses generated by inductive inference may or may not preserve the truth. The process that uses observations to discover rules and procedures is called induction. The induction process is indispensable to human beings, as it is one of the main means of creating new knowledge and predicting future events. It was through inductions that Kepler discovered the laws of planetary motion, that Mendel discovered the laws of genetics, and that Archimedes discovered the principle of the lever. One can dare to say that induction is the most used resource by human beings to obtain new knowledge. Despite this, this resource must be used with due care because, if the number of observations is insufficient or if the relevant data are poorly chosen, the rules obtained may be of little or no value. There are two forms of induction learning. In Learning by Examples, the learner induces the description of a concept by formulating a general rule from the examples and counter-examples provided by the teacher or the environment. The teacher already has knowledge of the concept and thus he can help the learner by selecting relevant examples to learn a given concept. The learner's task is to determine the general description of a concept by analyzing individual examples given to it. This strategy is also known as supervised learning.

Table 3.1: Inductive vs deductive learning [7]

Inductive learning	Transductive learning
Train the model and label unlabelled points which we have never encountered.	Train the model and label unlabelled points which we have already encountered.
Builds a predictive model. If new unlabelled points are encountered, we can use the initially built model.	Does not build a predictive model. If new unlabelled points are encountered, we will have to re-run the algorithm.
Can predict any point in the space of points beyond the unlabelled points.	Can predict only the points in the encountered testing dataset based on the observed training dataset.
Less computational cost.	Can become more computationally costly.

In Observational Learning, the learner analyzes given or observed entities and tries to determine whether some subsets of these entities can be usefully grouped into certain classes (ie concepts). As there is no teacher who already has the knowledge of the concept to provide meaningful examples of the concept to be learned, this strategy is also called unsupervised learning.

3.3.2 Reinforcement Learning

The use of interaction with the environment to learn certain tasks is the first idea that comes to mind when dealing with natural learning. Throughout life, undoubtedly the greatest source of knowledge of human beings is their interaction with the environment in which they are inserted. Learning through interaction is the fundamental idea that underpins some theories of learning and intelligence. Reinforcement Learning (RL) is synonymous with interaction learning, since the agent learns directly from the interaction with the environment where it is inserted (Shabani et al., 2003). In this process, the learner is not told what action to take, instead he must figure out which action will return the best reward. According to (Kaelbling et al., 1996) the standard model of reinforcement learning consists of: a set of environmental states (S), a set of possible actions for the agent (A) and a set of scalar reinforcement signals, commonly {0.1}. In this model, an agent is connected to the environment via perception (p) and action (a), as shown in Figure 5.1. At each interaction, the agent receives information (i) that indicates the current state(s) of the environment. After the agent knows the current state, it performs an action that will modify the state of the environment, and the value of this state transition is informed to the agent by a reinforcement value r, called reward.

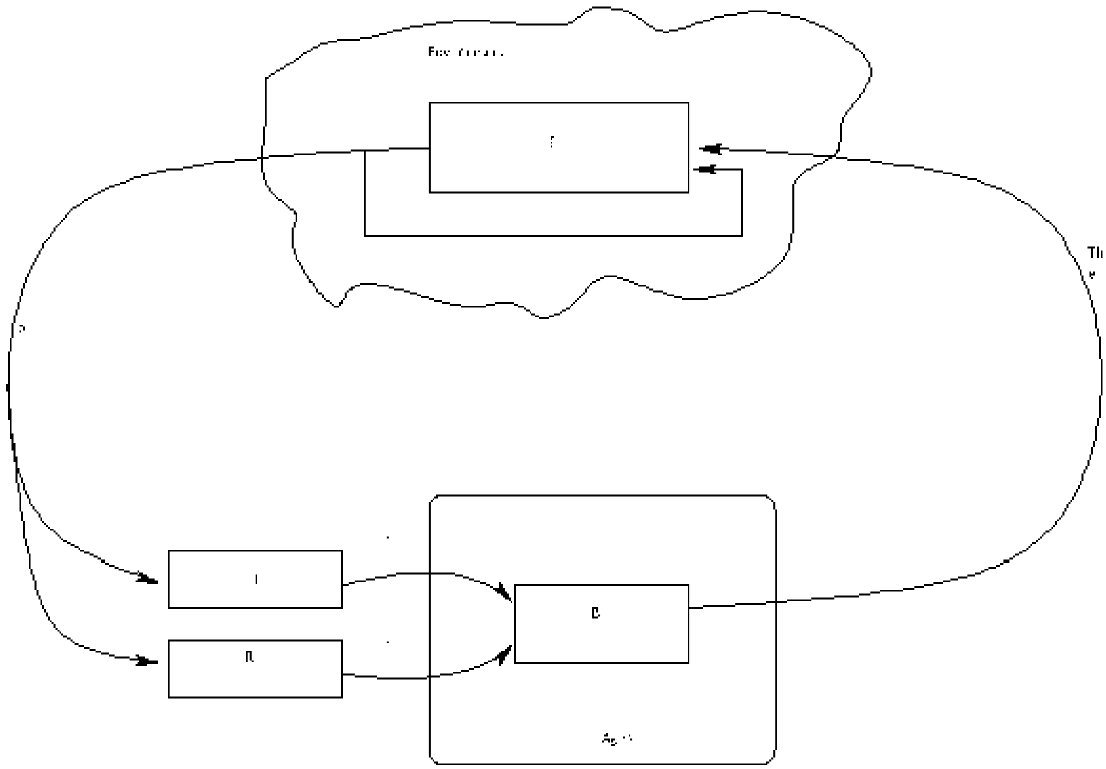


Figure 3.7: Standard model of Reinforcement Learning

Reinforcement learning is defined not by characterizing learning algorithms, but by the problem to be learned. Any algorithm that satisfies the resolution of a problem can be considered a reinforcement learning algorithm (Sutton and Barto, 1998). The basic idea is simply to capture the most important aspects of the real problem that a learning agent faces while interacting with the environment to achieve a goal. RL differs from supervised learning, as the latter consists of learning through examples and these depend on some external supervisor to be introduced into the system. This is an important type of learning, but by itself it is not suitable for interactive learning. In interaction problems, it is sometimes impractical to obtain examples that are both correct and that represent all kinds of situations. In many domains, reinforcement learning is the only practicable way to train a program that involves a high level of complexity.

3.3.3 Markov Decision Processes

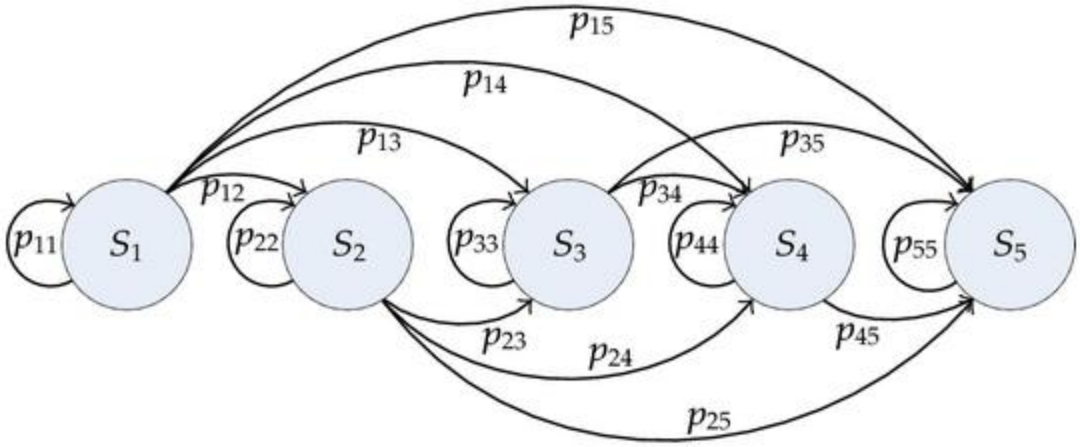
A large part of the work carried out in reinforcement learning assumes that the interaction of the agent with the environment can be modeled as a Markov decision process (MDP), in this way an agent step happens as follows:

At time t, the agent is in any state, s S, and chooses an action, a A s, according to a learning policyπ. After performing this action, the agent reaches another state, s , at tt 1 with a probability of Pssa , probability of reaching states , being in the state taking an action a. To perform this transition, the agent receives a reward, rt 1, given by Rass , probability of receiving a reward r when reaching states , starting from states, taking an action a .

i. A finite discrete Markov process in reinforcement learning consists of:
ii. A finite set of S states.
iii. A finite set of actions A.

A state transition function, Pssa Pr st 1 s st s at a, that is, the probability of reaching s at time = t 1 given an action a that was taken in states at time t.

A reward function that, from the sas triples, generates a numerical value for the agent, through Rass. Is true. That is, the probability distribution of the states inserted in t 1 is conditionally independent of the previous events of st at - knowing the current state and the action taken, are sufficient factors to define what will happen in the next step. In reinforcement learning this condition is also used for the reward function,



$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} \\ 0 & p_{22} & p_{23} & p_{24} & p_{25} \\ 0 & 0 & p_{33} & p_{34} & p_{35} \\ 0 & 0 & 0 & p_{44} & p_{45} \end{bmatrix} \text{ where } p_{ij} \geq 0 \text{ and } \sum_j p_{ij} = 1$$

Figure 3. 8: example of Markovian chain [43]

3.3.4 Elements of Reinforcement Learning

In addition to the agent and the environment, four important sub-elements for reinforcement learning can be identified:

i. a learning policy
ii. a reward function,
iii. an evaluation function and
iv. a model of the environment.

A. Learning Policy

a learning policyπit is the mechanism used to select the actions to be taken in a given state In other words, a policy$\pi$is the organization of states, s S, and of actions, a A s , for the probability$\pi sa$to take an action a while in the s state. This policy is what psychology calls a "stimulus-response" set of rules or associations. To optimize the agent's interaction with the environment, the future consequences of a learning policy must be known.

B. Reward Function

In a reinforcement learning process, the goal is defined through a reward function. An agent's goal is to maximize the total rewards received. Roughly speaking, the function maps the perceived states of the environment to a single value, usually numerical, called a reward. This function depends on the environment and also on some parameters, such as the actions taken by the agent and the states that the environment has already reached. The reward function defines the events that are or are not correct for the agent. An immediate reward defines the problem faced by the agent, so it must necessarily be a predetermined role.

C. Evaluation Function

While the reward function indicates whether or not the agent performed a good action right after it was performed, the state evaluation function, V s, describes its long-term behavior, that is, the value of a state corresponds to the total rewards that an agent expects to accumulate from the current state until reaching a final state. Most reinforcement learning algorithms are based on estimation of evaluation functions. An evaluation function is a function that estimates how good it is for the agent to be in a certain state, if a pi policy is applied. Assessment functions are defined respecting a particular learning policy. The value of a state's subject to a policyπ, $V^{\pi}s$, is the expected value when starting in a state if a policy has been followed since then.π. V^{π}ssecond can be defined as:

$$V^{\pi}s\sum\pi sa\sum Pssa\ Rass_{=}\gamma V^{\pi}s \qquad (3.1)$$

Whereγ0 1 is the discount factor that is used to weight future rewards. ifγ1 corresponds to immediate reward, otherwise, ifγ0, then it corresponds to the expected sum of future rewards. Equation is known as the Bellman equation for V^{π}, it expresses the relationship between the value of the current state and its successors. The Bellman equation simply averages over all possibilities, weighing each state according to its probability of being reached. The initial state value s must equal the expected value of its successor plus the expected payoff along the way.

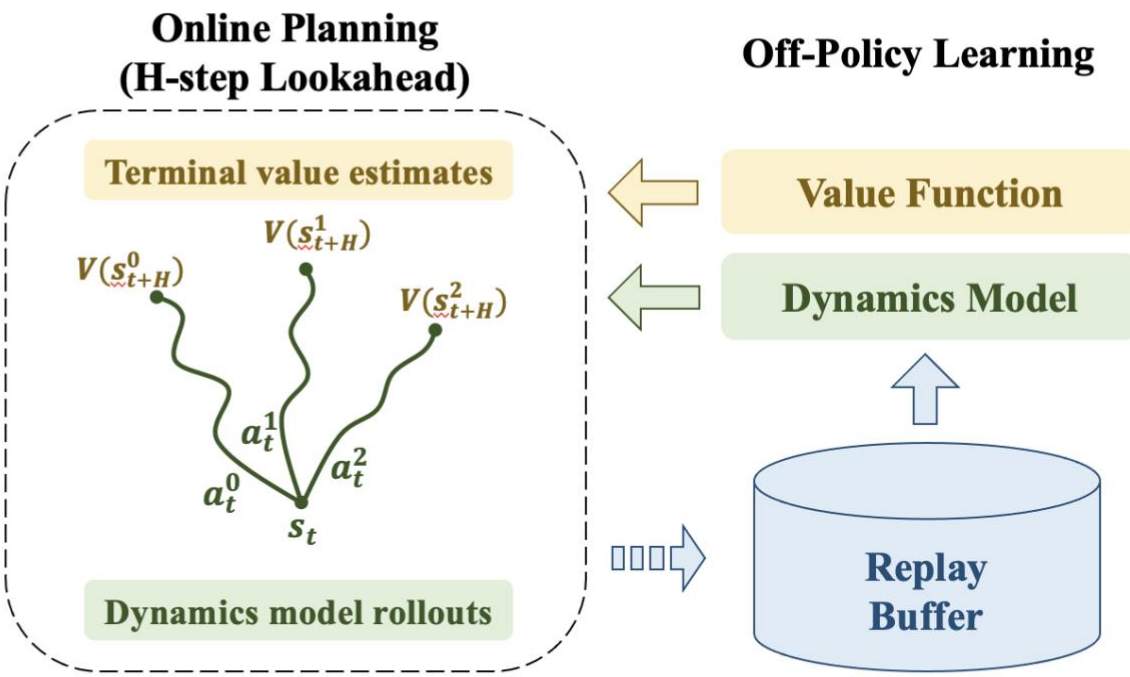


Figure 3.9: Bellman model for RL [45]

D. Environment Model

The fourth element of a reinforcement learning system is a model of the environment. These models are used to plan, what is intended to be a way of deciding the course of an action, considering possible future situations, before they are actually experienced. The merging of models and planning in reinforcement learning systems is a relatively new development. In a general reinforcement learning problem, the agent's actions determine not only the immediate reward, but also (through probability) the next state of the environment. In some cases the agent has to learn from delayed rewards, that is, it has to take a long sequence of actions, receiving insignificant rewards, problems with this type of reward are well modeled as Markov Decision Processes (MDP).

3.3.5 Finding A Learning Policy Given an Environment

This section will explore techniques to determine an optimal learning policy, knowing the model of the environment where the agent is inserted. The model consists of knowing the probabilities of state transitions Pssa and the reward function Rssa. To find an optimal policy, we first have to find an optimal evaluation function.

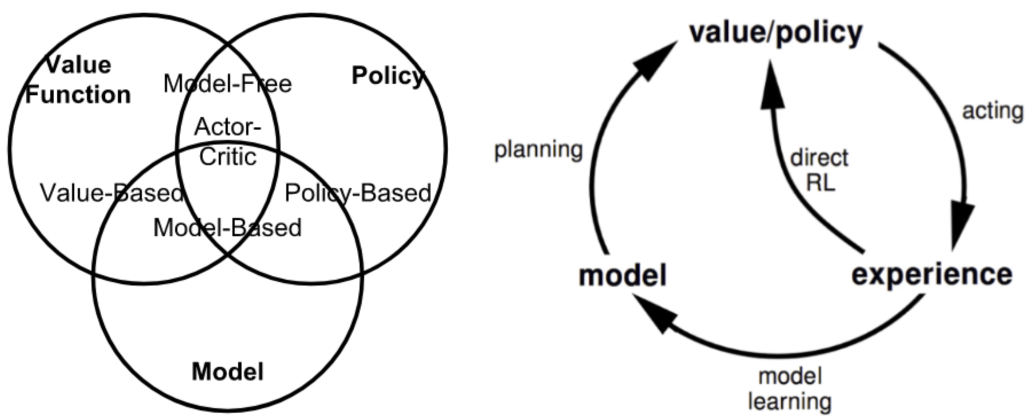


Figure 3.10: Value/Policy optimization [13]

What defines the value of the optimal evaluation function of a state's is the instantaneous reward added to the value of the next state, believing to have performed the best possible action. Given the optimal evaluation function of a state:

A. Iteration of Reviews

One way to find an optimal learning policy is to find an optimal evaluation function. This can be determined using a simple algorithm called Evaluation Iteration (Algorithm 1), which demonstrates how to converge to correct values of

V:

Algorithm 1 Iteration of Reviews

---

initialize V s arbitrarily

repeat

for$s$ $s$and the*THE***of**

$V\ s \quad max\sum Pss$

s

s**end for until**policy

good enough $\quad Rass \quad \gamma V\ s$

B. Policy Iteration

Another algorithm used to find a learning policy (Algorithm 2), deals directly with the learning policy, and then indirectly finds an optimal evaluation function. Once the value of an evaluation of a certain state that is subject to a learning policy is known, it can be considered that this value can be improved by changing the first action taken. If this is possible, change the learning policy to take the new action. This step ensures the performance of the new policy. When the action cannot be improved, this means that you have an optimal learning policy.

Algorithm 2 Policy Iteration

choose an arbitrary

policyπ**repeat**

π:π

compute from value function of

policyπ

solve the linear equations

$V\ s{:}\sum Pss\pi s \quad R\pi sss \quad \gamma V\ s$

$s\ s$

for$s$ $s$improve the policy of

$\pi s{:} \quad argmax\sum Pss\pi s \quad R\pi sss \quad \gamma V\pi s$

s s**end**

for

untilπ π

3.4 LEARNING A POLICY: WITHOUT KNOWING THE ENVIRONMENT

When the environment model is not known in advance, reinforcement learning is concerned with obtaining an optimal learning policy. For this, the agent has to interact directly with the environment to obtain information that will be processed by a certain algorithm to produce an optimal policy. There is a class of algorithms called Time Difference Methods, these algorithms use insights from algorithm 1 to adjust the estimated values of the evaluation functions of a given state based on the immediate reward and the estimated value of the next state. Algorithms that fall into this class will be discussed below. From this point on, when talking about the value of a state, this value refers to the value of the evaluation function of this state.

3.4.1 Adaptive Heuristic Critic and TD (Λ)

The Critical Adaptive Heuristic algorithm is an adaptation of algorithm 2, which uses the TD 0 algorithm instead of linear equations to calculate the value of the evaluation function of a given state. For the critic to learn the value of a policy, it was defined that sars is an experience tuple that summarizes a single state transition in the environment. The policy value is learned using the Sutton TD 0 algorithm. Where s is the state visited, its estimated value is updated to approximate rγV s, where r is the value of the instant reward received and V s is the estimated value of the next state. The main idea is that rγV $s$is a sample of the value of V s, and is closer to the correct value because it incorporates the real payoff r. If the learning rateαis well-adjusted and the policy is kept fixed, TD 0 is a guarantee to converge to an optimal evaluation function.

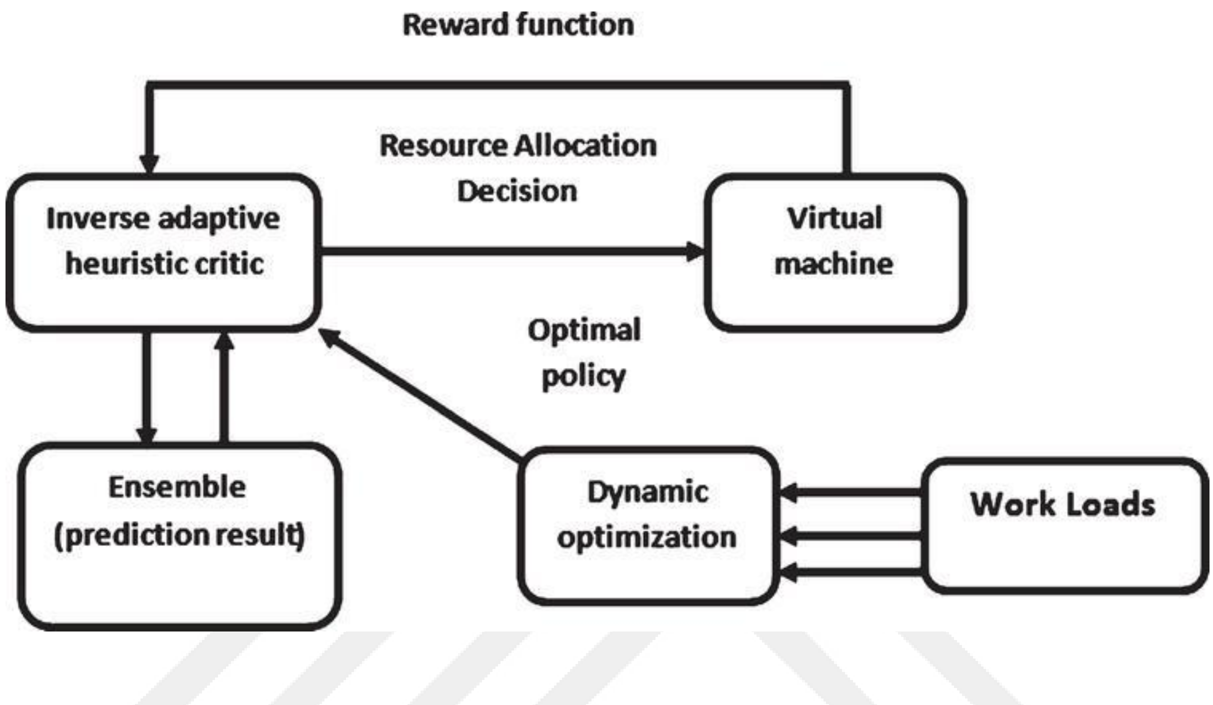


Figure 3.11: Reward function in AHC [22]

3.4.2 Q-Learning

The Q-learning algorithm (Watkins, 1989) is easy to implement. To understand this algorithm some equations have to be presented. First, Equation 5.7 defines the value of a state-action or Q-value evaluation function.

According to Equation 5.4 which can be reduced to: V s maxaQ sa . From these two equations comes another one

The Q-values can be estimated using the same method as for TD 0 (Equation 5.6), and can also be used to define the policy. For an action can be chosen only by taking the action that has the highest value among the Q-values for the current state. This type of update was used in the implementation of the Q-Learning algorithm.

3.4.3 Exploration X Exploitation

One of the dilemmas faced when it comes to reinforcement learning algorithms is the Exploration X Exploitation paradigm. In its literal translation, the two words mean the same thing, but in reinforcement learning, an Exploration approach means that the agent will take the stance of an explorer agent of the environment, that is, it will always be trying to explore paths that have not yet been explored. Experienced. The Exploitation approach, on the other hand, is a behaviour aimed at maximizing the values of its evaluation function, that is, the agent will always take the action that returns a better reward. In the Exploration approach, the agent acts randomly, thus being able to explore the entire environment where it is inserted, while the Exploitation may end up finding a path that reaches the agent's final goal, but that is not optimal, but because it does not act in an exploratory way. You will never know all the possible paths. An ideal approach for an agent would be one that is in between. The agent could act in a more exploratory way when only aware of the environment in which it is inserted, and from the moment that it had the model of the environment formulated, it would act in a way to maximize its evaluation function.

3.5 APPLICATIONS OF REINFORCEMENT LEARNING

One of the reasons that reinforcement learning is so popular is that it is treated as a theoretical tool for studying the principles of an agent learning to act. But it has also been used by countless researchers as a computational tool for building autonomous systems that improve on their own experiences. The applications of this learning reach several areas such as robotics, industrial manufacturing and computer games.

Reinforcement learning has been extensively explored in RoboCup. In (Riedmiller et al.,) brings a version of a team of a group of researchers who learned basic movements through reinforcement learning. Among the movements learned can be highlighted:

i. Kick: the player can kick the ball so that it reaches any speed, between 0 and 2.5m/s, in a certain direction;

ii. ball interception: the player learns to intercept the ball in motion, taking into account the stochastic environment of the RoboCup;
iii. Dribble: the player learns to run without losing control over the ball;
iv. To the ball: the player learns to stop the ball at high speed, this can be translated into "football" as "kill the ball".

4. PROPOSED METHODOLOGY

To apply machine learning in teaching and learning contexts, algorithms will be used that perform different tasks and on a set of synthetic data, that is, data generated from computational models. Thus, the tasks and corresponding algorithms selected for this study are: clustering with K-Means; classification with decision tree and linear regression with the least squares function. K-Means is one of the simplest and most popular clustering algorithms based on the representation paradigm. Therefore, this algorithm was chosen for use in this work, considering the purpose of initiating educators from different areas in the use of machine learning. The intuition of the K-Means algorithm is based on first, randomly choosing a starting position for the*k*centroids (each representing a distinct group) and from there, iterate over the association of each point to the group that has the closest centroid and move the centroid to the midpoint of its group members.

These iterations are done until the position of all centroids converge (no longer moving or reaching a pre-established amount of movement). More formally, given a data set*D*with*n*ospots*pi*(which represent entities such as students) in a space*d*-dimensional, and given the number*k*of desired groups, the objective of this algorithm is to partition the data set into*k*groups (cluster) denoted as*C = {C1, C2, ..., Ck}*. Furthermore, in each group*Ci*there is the centroidμi, a dot representative average of all points in the group, that is, $\mu_i = 1/n_i \sum_{p_j \in C_i} p_j$
Where*ni*= |*Ci*|,corresponds to the number of points in the group*Ci*.

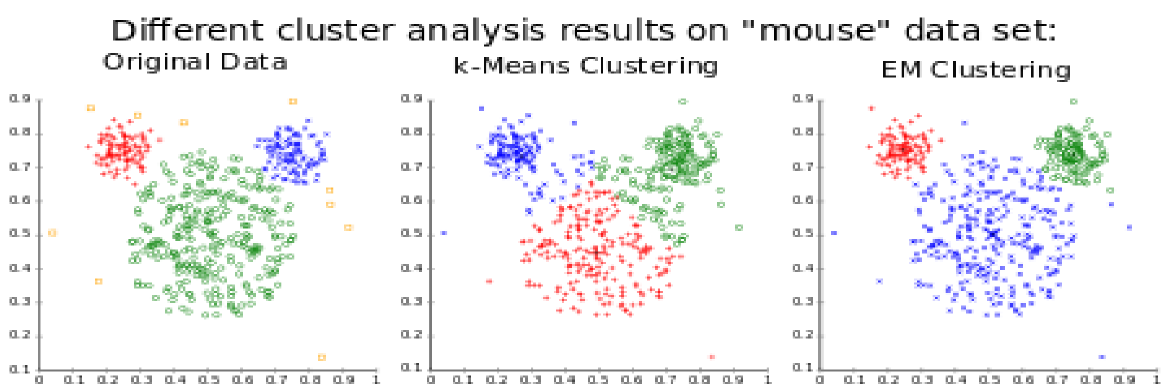


Figure 4.1: Data clustering using K-means algorithm [7]

4.1 DATA SET

The dataset used in this work consists of records that represent properties of 396 students, generated by computational models (synthetic data). These properties are described in Table 4.1, and were generated by two different mathematical models.

The first model (labeled in Table 1 as RL) is a sparse random linear combination of properties, with noise. These properties simulate the performance of students as a function of the quality attributed by them to the learning objects used in a period of a course.

The second model (labeled in Table 1 as CL) corresponds to 10 attributes with values ranging from 1 to 10 representing students' preference for 10 types of learning objects.

Table 4.1: Synthetic dataset properties

Label	Description
student	student identifier
pref_audio	Audio-based object preference
pref_image	Preference for images
pref_infographic	Preference for charts, diagrams
pref_games	Preference for game-based activities
pref_webpgs	Preference for web pages
pref_table	preference for tables
pref_text	Preference for reading texts
pref_tutorial	Preference for teaching tutorials
pref_video	Preference to watch videos
pref_forum	Preference for participation in forums

WILL	academic performance Index
pref_avg	Preference estimate

Note that it was decided to use synthetic data for these properties due to the difficulty of collecting or finding all of them available. Furthermore, for the purpose of this work, the interpretation of the values contained in these data is not relevant, as it does not influence the achievement and quality of the results. The experimentation made use of synthetic data, modeled considering the following hypothetical data collection scenarios:

Figure 4.2: Simulated classroom data

4.1.1 Preferences for Types of Learning Objects

i. Offer 10 different types of learning objects (eg image, infographic, games, web pages, table, text, tutorial, video and forum).

ii. Ask each student to assign a grade from 0 to 10 for each type of learning object, according to their preference.

4.1.2 Satisfaction with the Use of Learning Objects and Learning Performance

i. Plan for a lesson an activity that makes use of different types of learning objects.

ii. Offer each student a different learning object, randomly selected from those available.
iii. After the student's interaction with the object, evaluate his performance in understanding the content of the class. Repeat up to step 3 for 10 lessons.
iv. Analyze the objects that were offered over the 10 classes and calculate the percentage of satisfaction for each of them, for each student. To do this, consider the grade given by each student in the Scenario 1 experiment of each object he used in Scenario 2.

Also consider the corresponding grade that the student obtained in each activity (performance). Note that these scenarios can be adapted to meet other objectives of the teaching and learning process, or to make use of a different data collection methodology. In addition, it is important to remember that, in a real scenario, experiments need to be authorized by a research ethics committee.

4.3 MACHINE LEARNING ALGORITHMS

The algorithms used in this work were selected among the most popular for each of the three tasks considered. They are machine learning algorithms that can be easily implemented in different programming languages or in data mining tools such as Weka, Octave, Neural Designer, among others. Next, we describe each of these algorithms in their respective tasks.

4.3.1 Clustering

In choosing the algorithm for clustering preferences for learning objects, the possibility of forming groups of approximate sizes was considered, as normally occurs in a classroom. Thus, it becomes possible, in addition to identifying similar preferences for recommending types of objects, it also makes it possible to have a reference to suggest group members in collaborative activities oriented to different activities, as in methodologies such as rotation by stations

To choose the number of groups of students, the elbow method was used. The elbow method allows searching for the best number of groups for a clustering, which is the one that has the lowest possible value for the sum of squares within the groups. Within-clusters sum-of-squares – wcss). Figure 4.3 illustrates the different values of*toilets*, for group amounts ranging from 1 to 10.

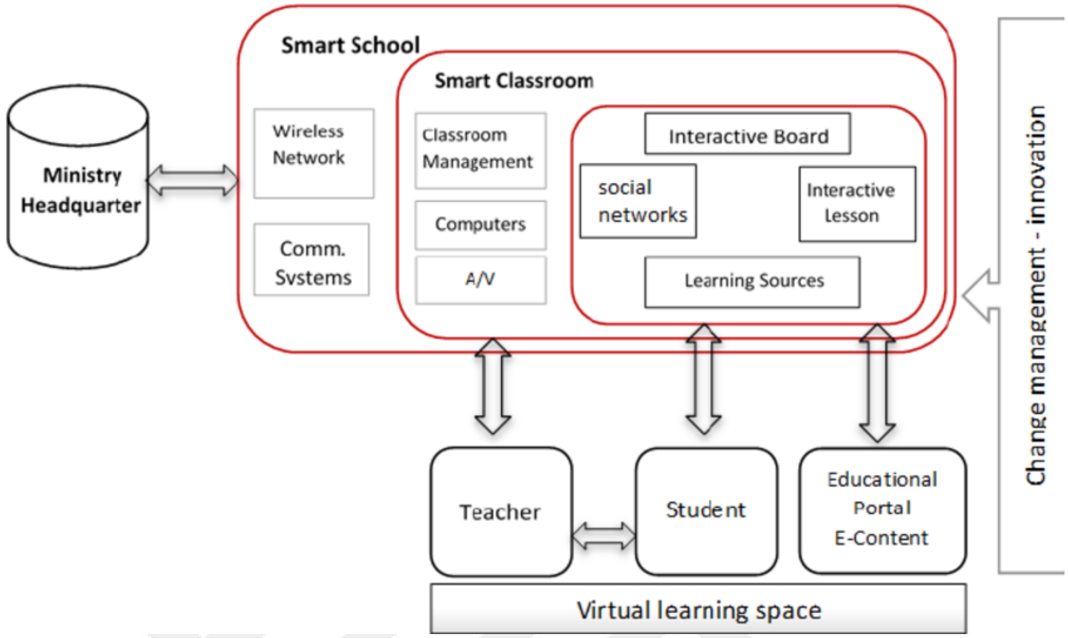


Figure 4.3: clustering of the proposed classroom

By analyzing the clusters on the synthetic data set presented in Figure 4.4, it is observed that those who have the lowest possible value for to Ile shave 5, 6, 8, 9 or 10 groups. Among these, the smallest number of clusters was chosen, that is, with 5 groups. This means that five distinct profiles can be identified based on students' responses about their preferences for types of learning objects. Although cluster analysis considers 10 different types of learning objects, each representing a dimension, only a few dimensions can be easily visualized simultaneously due to the limitations of human perception. On the other hand, the examples in Figure 5, which illustrate different perspectives of the groups in 3 dimensions, show that these five groups are very well defined, which corroborates the low value of the toilets. In practice, this means that it is important to customize the types of learning objects for students who have profiles like these from our synthetic data, as they have diverse preferences, but which can be optimally met through the use of cluster analysis.

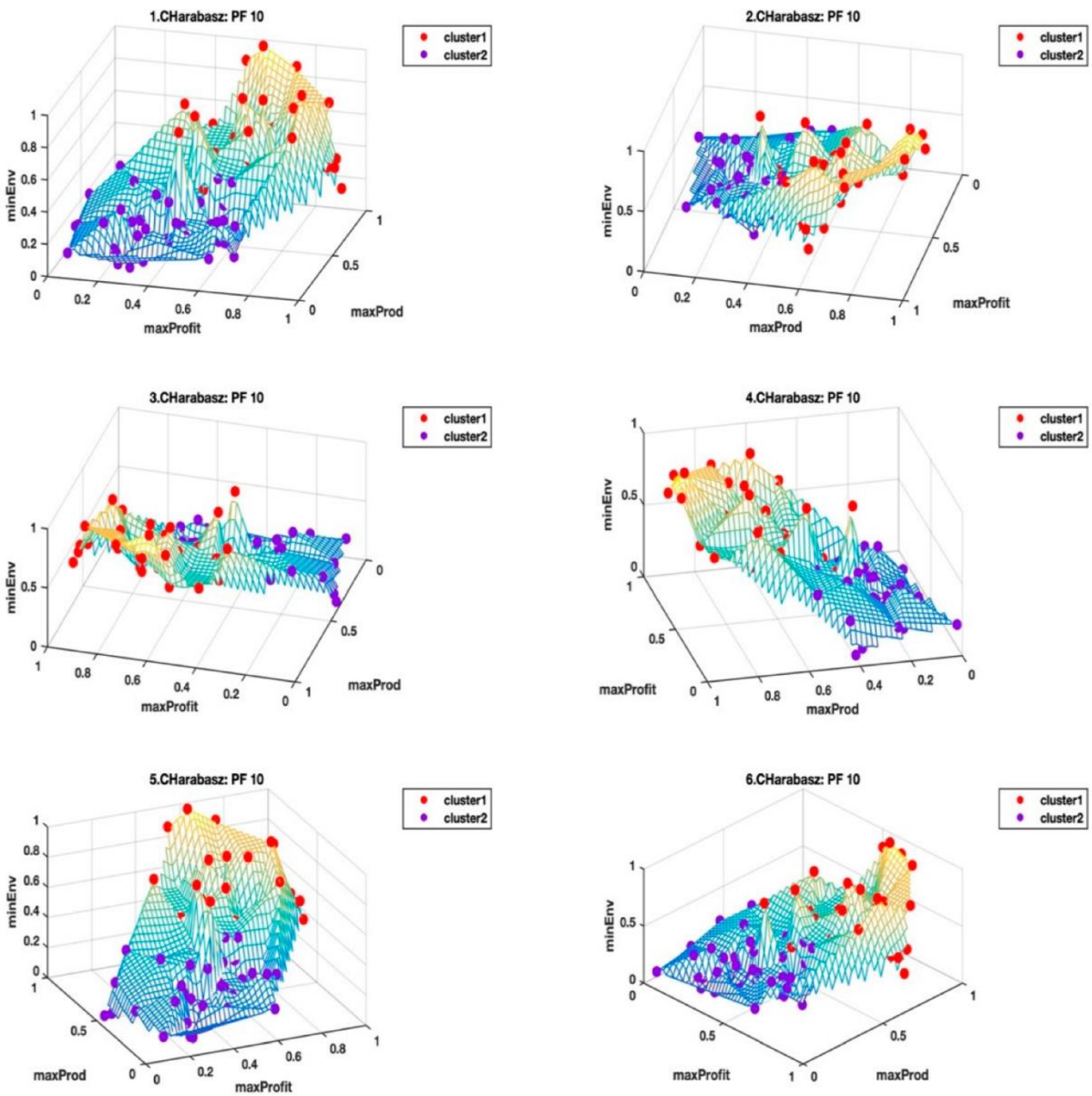


Figure 4.4: Three-dimensional visualization of the clustering of preferences, under three of the possible combinations of types of learning objects

In this way, it is also possible to do this analysis manually by the teacher, using a few hours to work on the data provided by the students to obtain additional information that contributes to their class. In addition, note that the characteristics of synthetic data such as the ones we used facilitate the interpretation of their visualization, which is favourable to the demonstration purpose of this work. Finally, a machine learning algorithm can be used on a dataset with many more student records, more dimensions and with values not as well behaved as the ones we used. Even so, it is possible to obtain statistics and extract good quality groupings and their visualizations in a few seconds, which would be unfeasible to be done in a short time without the use of adequate techniques.

4.3.2 Classification

For the data classification task, the decision tree technique was chosen (*decision tree*), which is considered one of the most intuitive and highly interpretable. This technique can be defined as a decision node that contains a test, a question, about each data attribute. For each result of this test there is a link to a certain set of data records. Each set has a test-to-end-result structure, as in Figure 4.5, where the algorithm sorts students into three groups based on the test by checking each one's grade. The method used to generate the tree was recursive partitioning

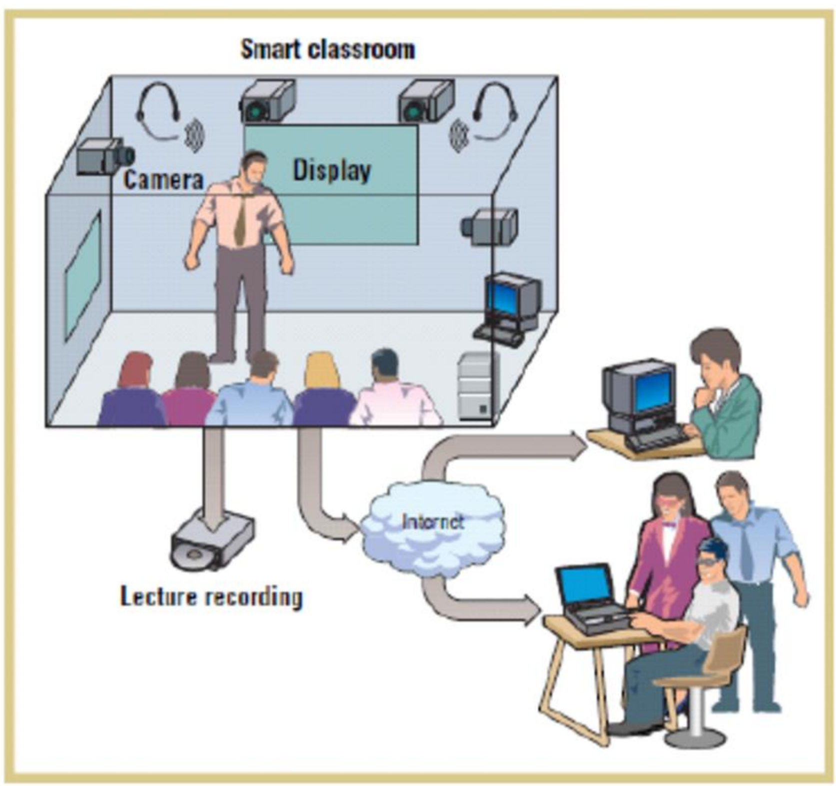


Figure 4.5: Classification and monitoring data in the proposed smart classroom

4.3.3 Linear Regression

Linear regression is a model that allows estimating the value of a variable in a hyperplane P-dimensional referring to explanatory variables (Menezes et al., 2018). More formally we have:

$$Q_w = \sum_{i=1}^{n} w_i (Y_i^* - \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip})^2$$

Where Y is the response component, $\beta 1, \beta 2, \ldots, \beta p$ are the estimators that minimize the weighted least square's function (MQP) and wi is the weighted deviation. This method is interesting to be applied when there is a high degree of relationship between the explanatory variables. For this, a correlation study must be carried out.

4.4 RESULTS

Next, the results of executing each of the machine learning techniques on the synthetic data set of students are presented. In addition, this dataset was made publicly available. The summary of preferences contained in this dataset is shown in Figure 4.6.

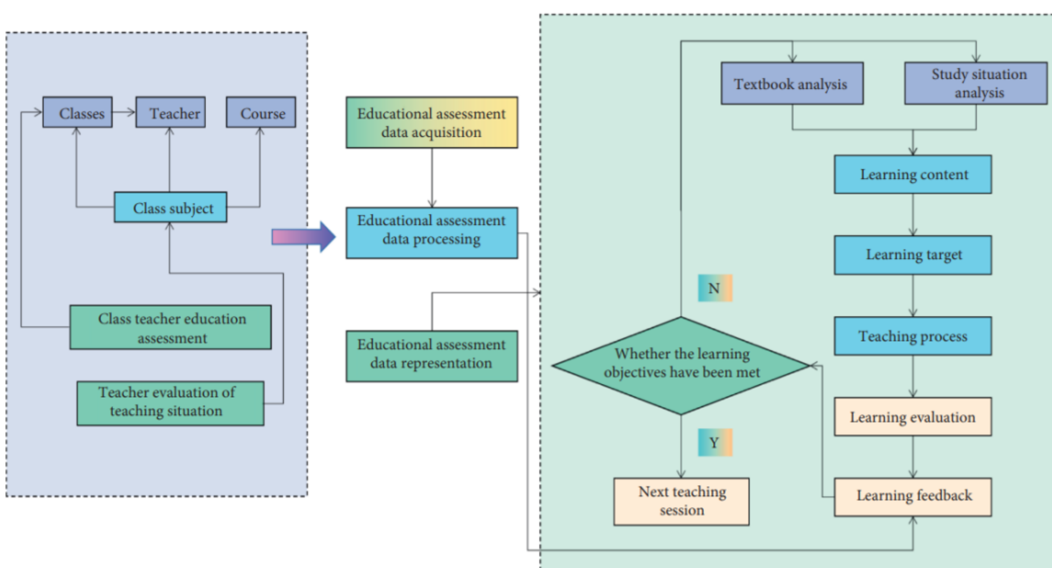


Figure 4.6: General statistics of preferences by type of learning objects

Students to groups of close sizes based on their preferences without the support of a suitable algorithmic technique to identify the groups and allow an individual analysis, as described below. To answer about which main preference profiles exist in the data, each group of students was analyzed individually. Thus, in addition to the general statistics of this dataset, the preferences shared by students belonging to the same group are described below. In addition, from the cluster analysis it was possible to verify which objects prevail in the preferences and which ones would be fundamental to satisfy the students. It was observed that at least one of the video, audio or forum objects are preferred by all groups, since in all groups one or more of these objects received a score above 8 in relation to preference. For students belonging to groups with these descriptions, content in more appropriate media can be directed as well as collaborative activities that involve the use of these media. A natural situation in the analysis of preference over learning objects is to offer students that object that has the highest estimate of preference of the grades in the opinion consultation of these students. This can be easily obtained in the general statistics by analysing the highest mean (circle in the boxplot in Figure 6) and median values. Thus, according to Figure 4.6, objects of type's video and games are among the favourites. However, this general and premature analysis missed something important that can be verified in the cluster analysis: a considerable part of the students rejects precisely objects of this type, despite having higher general averages. This can be verified by observing each group individually, in particular, analysing in Figure 4.8 the lowest preference indicated in groups 1 and 4 for objects of the type video and in group 3 by objects of the type games. Cases like these require personalization of learning strategies to maximize learning satisfaction. Note that this analysis also allows to identify objects scored by all groups with reasonable scores, as is the case of the object of type forum.

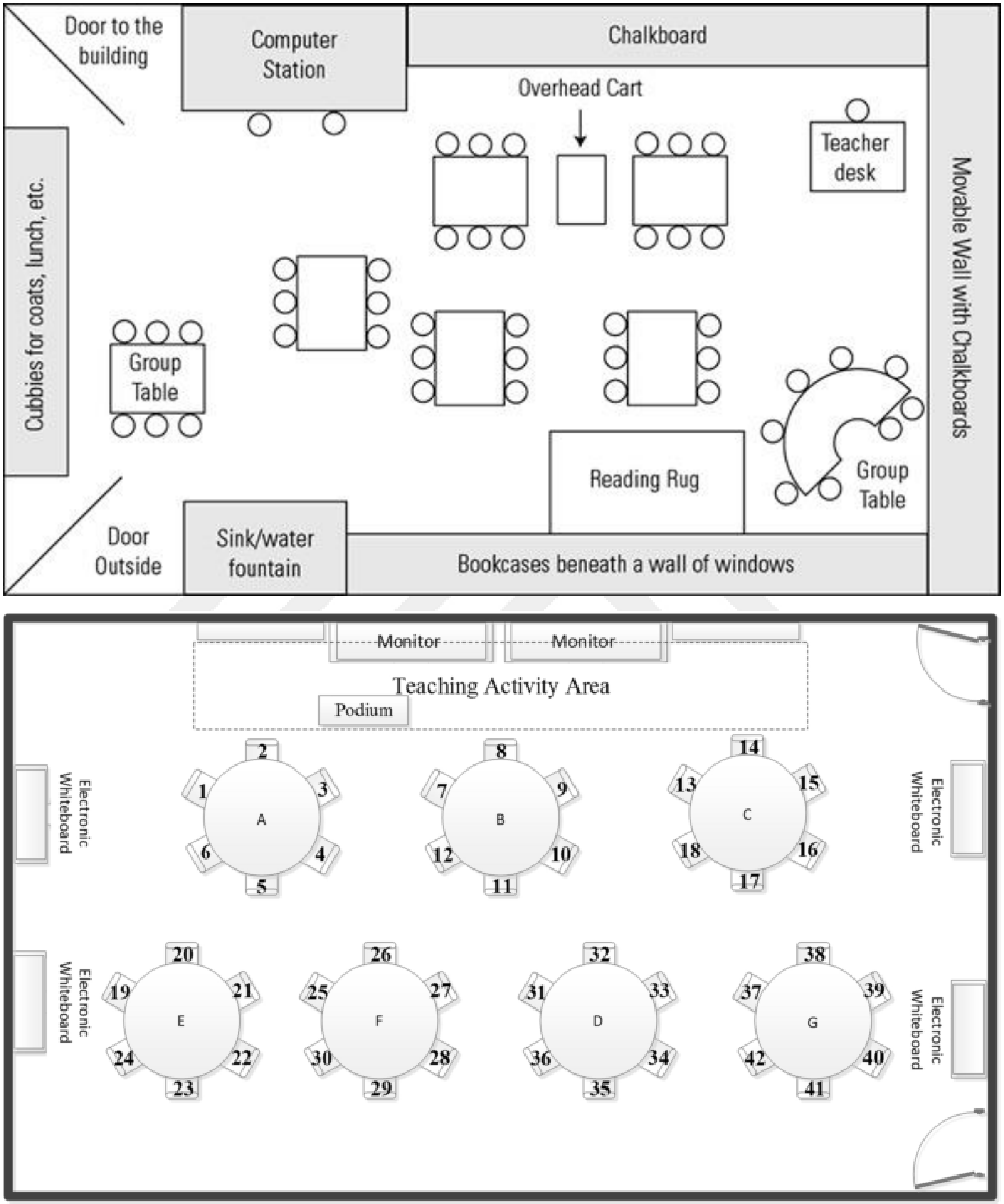


Figure 4.7: Statistics of preferences by type of learning objects of each group

In Figure 4.7 it is possible to observe the decision tree obtained from the preference data of students combined with the labels of the group to which they were associated by the clustering technique. This model allows new students to be classified and then associated with a group without having to perform the clustering technique on the data of all other students again. In addition to the reduced computational cost, the use of this model allows us to understand how the student's properties were considered to direct him to his group.

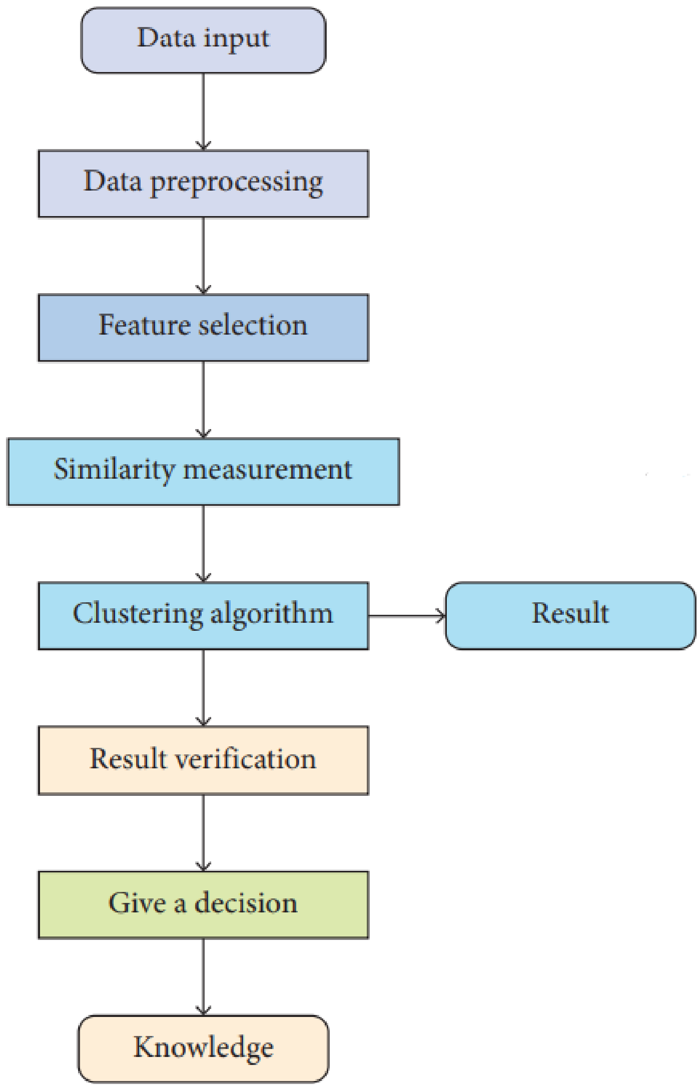


Figure 4.8: Decision Tree generated by the recursive partitioning method

Figure 4.8 illustrates the application of linear regression on synthetic data on student performance in a discipline as a function of the satisfaction of their preferences for the learning objects used in the discipline.

Thus, in the classroom, such an application can help the teacher to know if the student's learning is as expected (estimated by the model) considering his preference for types of learning objects.

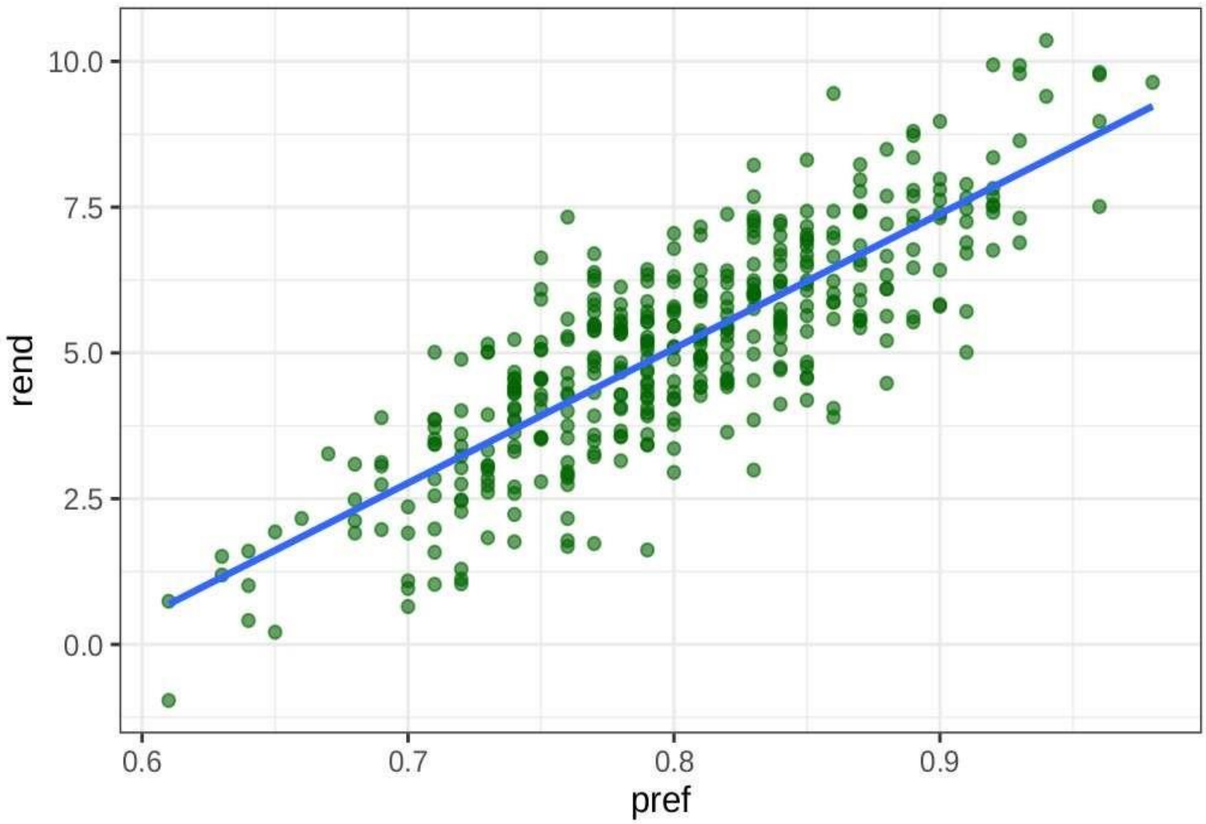


Figure 4.9: Regression model: income and preferences

Note that, when working with real data, in addition to allowing verifications like the ones presented above, it is still possible to interpret the values contained in this data and obtain additional results, such as peculiar preference patterns. Furthermore, it becomes possible for the information discovered with the support of algorithms to be applicable to a specific teaching-learning context.

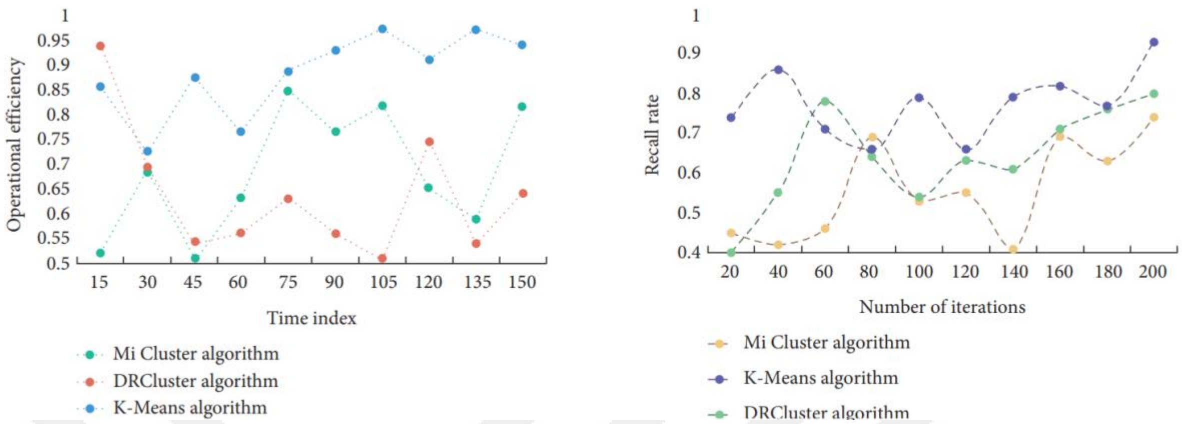


Figure 4.10: Operational efficiency of the proposed K-means algorithm

4.5 EVALUATION METRICS

The number of correctly classified instances a (yes) divided by the number of predicted instances will give the Precision a, is computed as 4/8, while Recall is 4/9, the number of correctly classified instances a (yes) divided by the total number of true instances a. The precision and recall for the other class instances b is the just reverse. Various other measures of calculation for performance evaluation are such as listed below. Entropy, is normally a split criterion used in ID3, C4.5 (J48) Decision Tree algorithms. The sample is completely homogeneous the entropy is 0 and if the sample is an equally divided it has entropy of 1.

CONFUSION MATRIX

		Detected	
		Positive	**Negative**
Actual	**Positive**	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

$$\text{Accuracy}=\frac{TP+TN}{TP+FN+FP+TN} \qquad \text{-----(1)}$$

$$\text{Precision}=\frac{TP}{TP+FP} \qquad \text{-----(2)}$$

$$\text{Recall}=\frac{TP}{TP++FN} \qquad \text{-----(3)}$$

$$\text{F-measure}=2*\frac{\text{Precision}-\text{Recall}}{\text{Precision}+\text{Recall}} \qquad \text{-----(4)}$$

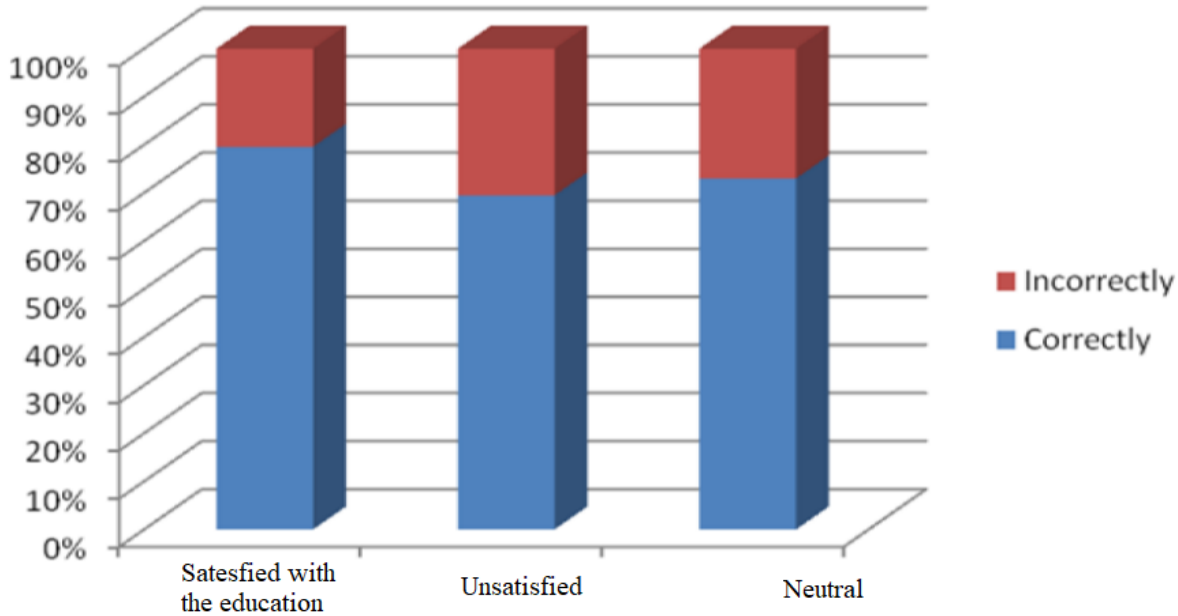


Figure 4.11: classification of student satisfaction with the education

5. CONCLUSION

The present work aimed to demonstrate the use of machine learning algorithms in the teaching-learning process through experiments performed on a set of synthetic data and in order to favour the reproduction of these experiments. Our results revealed patterns of information that would be impossible to obtain quickly without the support of machine learning techniques. Thus, it was possible to complete accessible instructions and materials for reproducing the experiments with the considered artificial intelligence techniques, in addition to others. A limitation of the models generated in our work is that they do not aim to present the behaviour of students in a real environment. Despite this, these data satisfactorily met the object of the work and this limitation can be easily overcome by reproducing our experiments on data collected from a real educational environment, preferably the same where they will be used. Another limitation of the work arises from the dynamic nature of the educational environment, which may require, for example, additional skills to mine your data that go beyond understanding the techniques presented. In future works, we intend to perform new machine learning tasks, as well as demonstrate other solutions based on artificial intelligence to solve other types of problems in the teaching and learning process.

REFERENCES

[1] Zawacki-Richter, Olaf, Eva Maria Bäcker, and Sebastian Vogt. "Review of distance education research (2000 to 2008): Analysis of research areas, methods, and authorship patterns." *International Review of Research in Open and Distributed Learning* 10.6 (2009): 21-50.

[2] Vieira, Marcos Alves, and Sergio T. Carvalho. "Configuraç ao de Espaços Inteligentes para Sistemas Ubıquos de Monitoramento de Pacientes Domiciliares." *Anais da III Escola Regional de Informática de Goiás (ERI-GO 2015)* (2015): 19-30.

[3] Romero, Cristobal, and Sebastian Ventura. "Data mining in education." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 3.1 (2013): 12-27.

[4] Papamitsiou, Z.; Economides, A.A. Learning analytics for smart learning environments: A meta-analysis of empirical research results from 2019 to 2015. Learn. Des. Technol. Int. Compend. Theory Res. Pract. Policy 2016, 1–23.

[5] Patel, Unnati A., and S. Priya. "Development of a student attendance management system using RFID and face recognition: a review." *International Journal of Advance Research in Computer Science and Management Studies* 2.8 (2014): 109-119.

[6] Keleş, Mümine Kaya, and Selma Ayşe Özel. "A review of distance learning and learning management systems." *Virtual learning* (2016).

[7] Chen, Peng, et al. "A review of using Augmented Reality in Education from 2011 to 2016." *Innovations in smart learning* (2017): 13-18.

[8] Glover, Derek, et al. "The interactive whiteboard: a literature survey." *Technology, Pedagogy and Education* 14.2 (2005): 155-170.

[9] Higgins, Steve, Gary Beauchamp, and Dave Miller. "Reviewing the literature on interactive whiteboards." *Learning, Media and technology* 32.3 (2007): 213-225.

[10] Martin, Susan F., Edward L. Shaw, and Lynda Daughenbaugh. "Using smart boards and manipulatives in the elementary science classroom." *TechTrends* 58.3 (2014): 90-96.

[11] Arumugam, Raman, et al. "Technology acceptance on Smart Board among teachers in Terengganu using UTAUT model." *Asian Social Science* 10.11 (2014): 84-91.

[12] Fies, Carmen, and Jill Marshall. "Classroom response systems: A review of the literature." *Journal of Science Education and Technology* 15.1 (2006): 101-109.

[13] Kay, Robin H., and Ann LeSage. "Examining the benefits and challenges of using audience response systems: A review of the literature." *Computers & Education* 53.3 (2009): 819-827.

[14] Georgiev, Tsvetozar, Evgenia Georgieva, and Angel Smrikarov. "M-learning-a New Stage of E-Learning." *International conference on computer systems and technologies-CompSysTech*. Vol. 4. No. 28. 2004.

[15] Wang, Minjuan, et al. "The impact of mobile learning on students' learning behaviours and performance: Report from a large blended classroom." *British Journal of Educational Technology* 40.4 (2009): 673-695.

[16] Wu, Wen-Hsiung, et al. "Review of trends from mobile learning studies: A meta-analysis." *Computers & education* 59.2 (2012): 817-827.

[17] Abachi, Hamid R., and Ghulam Muhammad. "The impact of m-learning technology on students and educators." *Computers in human behavior* 30 (2014): 491-496.

[18] Ha, Ilkyu, and Chonggun Kim. "The research trends and the effectiveness of smart learning." *International Journal of Distributed Sensor Networks* 10.5 (2014): 537346.

[19] PArkEr, BETTy, and DAviD BurniE. "Classroom technology in business schools: A survey of installations and attitudes toward teaching and learning." *AACE Review (formerly AACE Journal)* 17.1 (2009): 45-60.

[20] Zhou, Baiyu. "Smart classroom and multimedia network teaching platform application in college physical education teaching." *International Journal of Smart Home* 10.10 (2016): 145-156.

[21] Algarni, Abdulmohsen. "Data mining in education." *International Journal of Advanced Computer Science and Applications* 7.6 (2016).

[22] Aleem, Abdul, and Manoj Madhava Gore. "Educational data mining methods: A survey." *2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT)*. IEEE, 2020.

[23] Charitopoulos, Angelos, Maria Rangoussi, and Dimitrios Koulouriotis. "On the use of soft computing methods in educational data mining and learning analytics research: A review of

years 2010–2018." *International Journal of Artificial Intelligence in Education* 30.3 (2020): 371-430.

[24] Salloum, Said A., et al. "Mining in educational data: review and future directions." *The International Conference on Artificial Intelligence and Computer Vision*. Springer, Cham, 2020.

[25] Dutt, Ashish, Maizatul Akmar Ismail, and Tutut Herawan. "A systematic review on educational data mining." *Ieee Access* 5 (2017): 15991-16005.

[26] Shahiri, Amirah Mohamed, and Wahidah Husain. "A review on predicting student's performance using data mining techniques." *Procedia Computer Science* 72 (2015): 414-422.

[27] Rastrollo-Guerrero, Juan L., Juan A. Gómez-Pulido, and Arturo Durán-Domínguez. "Analyzing and predicting students' performance by means of machine learning: A review." *Applied sciences* 10.3 (2020): 1042.

[28] Wohlin C, Prikladniki R. Systematic literature reviews in software engineering.Inf Softw Technol. 2013;55(6):919-920.

[29] Namoun, Abdallah, and Abdullah Alshanqiti. "Predicting student performance using data mining and learning analytics techniques: A systematic literature review." *Applied Sciences* 11.1 (2020): 237.

[30] Garcia, Reinaldo C., et al. "A GARCH forecasting model to predict day-ahead electricity prices." *IEEE transactions on power systems* 20.2 (2005): 867-874.

[31] Malo, Pekka, and Antti Kanto. "Evaluating multivariate GARCH models in the Nordic electricity markets." *Communications in Statistics—Simulation and Computation®* 35.1 (2006): 117-148.

[32] Weron, Rafał, and Adam Misiorek. "Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models." *International journal of forecasting* 24.4 (2008): 744-763.

[33] Taylor, James W. "Density forecasting for the efficient balancing of the generation and consumption of electricity." *International Journal of Forecasting* 22.4 (2006): 707-724.

[34] Omar, Madiah Binti, et al. "Smart Grid Stability Prediction Model Using Neural Networks to Handle Missing Inputs." *Sensors* 22.12 (2022): 4342.

[35] Koprinska, Irena, Mashud Rana, and Vassilios G. Agelidis. "Correlation and instance based feature selection for electricity load forecasting." *Knowledge-Based Systems* 82 (2015): 29-40.

[36] Li, Song, Peng Wang, and Lalit Goel. "Short-term load forecasting by wavelet transform and evolutionary extreme learning machine." *Electric Power Systems Research* 122 (2015): 96-103.

[37] Ismail, Z., A. Yahya, and K. A. Mahpol. "Forecasting peak load electricity demand using statistics and rule based approach." *American Journal of Applied Sciences* 6.8 (2009): 1618.

[38] Conejo, Antonio J., et al. "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models." *IEEE transactions on power systems* 20.2 (2005): 1035-1042.

[39] Aggarwal, S. K., L. M. Saini, and Ashwani Kumar. "Price forecasting using wavelet transform and LSE based mixed model in Australian electricity market." *International Journal of Energy Sector Management* (2008).

[40] Pindoriya, N. M., S. N. Singh, and S. K. Singh. "An adaptive wavelet neural network-based energy price forecasting in electricity markets." *IEEE Transactions On power systems* 23.3 (2008): 1423-1432.

[41] Galar, Mikel, et al. "A survey of fingerprint classification Part II: Experimental analysis and ensemble proposal." *Knowledge-Based Systems* 81 (2015): 98-116.

[42] Fan, Shu, et al. "Forecasting electricity demand by hybrid machine learning model." *International Conference on Neural Information Processing*. Springer, Berlin, Heidelberg, 2006.

[43] Tan, Zhongfu, et al. "Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models." *Applied energy* 87.11 (2010): 3606-3610.

[44] De Felice, Matteo, and Xin Yao. "Short-term load forecasting with neural network ensembles: A comparative study [application notes]." *IEEE Computational Intelligence Magazine* 6.3 (2011): 47-56.

[45] Zhang, Rui, et al. "Short-term load forecasting of Australian National Electricity Market by an ensemble model of extreme learning machine." *IET Generation, Transmission & Distribution* 7.4 (2013): 391-397.

[46] Koprinska, Irena, et al. "Combining pattern sequence similarity with neural networks for forecasting electricity demand time series." *The 2013 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2013.

[47] Shen, Wen, et al. "An ensemble model for day-ahead electricity demand time series forecasting." *Proceedings of the fourth international conference on Future energy systems*. 2013.

[48] Jetcheva, Jorjeta G., Mostafa Majidpour, and Wei-Peng Chen. "Neural network model ensembles for building-level electricity load forecasts." *Energy and Buildings* 84 (2014): 214-223.

[49] Kaur, Amanpreet, Hugo TC Pedro, and Carlos FM Coimbra. "Ensemble re-forecasting methods for enhanced power load prediction." *Energy Conversion and Management* 80 (2014): 582-590.

[50] Jensen, Lasse, and Flemming Konradsen. "A review of the use of virtual reality head-mounted displays in education and training." *Education and Information Technologies* 23.4 (2018): 1515-1529.

[51] Minogue, James, and M. Gail Jones. "Haptics in education: Exploring an untapped sensory modality." *Review of Educational Research* 76.3 (2006): 317-348.

[52] Wiggins, Grant P. *Assessing student performance: Exploring the purpose and limits of testing*. Jossey-Bass, 1993.

[53] Rice, William. *Moodle 2.0 e-learning course development a complete guide to successful learning using Moodle*. Packt Pub.,, 2011.

[54] Pardo, Abelardo, and Carlos Delgado Kloos. "Using third party services to adapt learning material: A case study with Google forms." *European Conference on Technology Enhanced Learning*. Springer, Berlin, Heidelberg, 2009.

[55] Deng, Hui, Feng Wang, and Bo Liang. "A low-cost OMR solution for educational applications." *2008 IEEE international symposium on parallel and distributed processing with applications*. IEEE, 2008.

[56] Weaver, Ruth, and Brian Chalkley. "Introducing objective tests and OMR-based student assessment: A case study." *Journal of Geography in Higher Education* 21.1 (1997): 114-121.

[57] Vimal, P. P., and C. K. Kumbharana. "Analysis of different examination patterns having question answer formulation, evaluation techniques and comparison of MCQ type with one word answer for automated online examination." *International Journal of Scientific and Research Publications* 6.3 (2016).

[58] Ala-Mutka, Kirsti M. "A survey of automated assessment approaches for programming assignments." *Computer science education* 15.2 (2005): 83-102.

[59] Martin, Joel, and Kurt VanLehn. "Student assessment using Bayesian nets." *International Journal of Human-Computer Studies* 42.6 (1995): 575-591.

[60] Shermis, Mark D., et al. "Automated essay scoring: Writing assessment and instruction." *International encyclopedia of education* 4.1 (2010): 20-26.

[61] Park, Chan Jung, and Jung Suk Hyun. "A peer-assessment system connecting on-line and a face-to-face smart classroom." *Life Science Journal* 11.7 (2014): 700-705.

[62] Batane, Tshepo. "Turning to Turnitin to fight plagiarism among university students." *Journal of Educational Technology & Society* 13.2 (2010): 1-12.

[63] Graven, O.H.; MacKinnon, L.M. A Consideration of the Use of Plagiarism Tools for Automated Student Assessment.*Educ*. 2018, 51, 212–219.

[64] Ahmed, R. A. "Overview of different plagiarism detection tools." *International Journal of Futuristic Trends in Engineering and Technology* 2.10 (2015): 1-3.