Multilevel Hybrid System Based on Machine Learning and AHP for Student Failure Prediction

Nawal Sael†, Touria Hamim†and Faouzia Benabbou†
Laboratory of Modeling and Information Technology Faculty of Sciences Ben M'SIK, University Hassan IICasablanca, Morocco

Summary
Student profile detection and performance prediction are both high-potential areas in the educational data mining domain. To meet these goals, multi-criteria analysis and machine learning techniques are employed to help with decision making when it comes to student failure and extracting useful, hidden and relevant information about students. In this paper, we propose a multilevel hybrid system for student performance prediction. This work combines two approaches: multi-criteria analysis via the method AHP (Analytic Hierarchical Process) and classification and multi-level prediction using machine learning techniques. Different classification techniques were compared, such as SVM, NB and DT, with the last one performing the best. An analysis using association rules was also conducted in order to detect the different hidden relationships between the scores obtained and the modules. To obtain a high performance in students' failure prediction, we successfully aggregated machine learning methods with feature selection and parameter optimization process. The results show that the student performance prediction is efficiently done and sufficient performance is obtained. Hence, our system is able to identify at-risk students, assess the adequacy of the courses or modules, and help tailor interventions to improve on student success.
Key words:
Educational data mining; prediction; profile analysis; student failure

1. Introduction

With the development of new technologies, the fields of education and pedagogy have undergone major beneficial changes. This development has contributed to the innovation of new learning methods that are based on an important component that is the profile of the learner. In fact, knowing the students better in relation to their style and the pace of learning, the courses in which they have more difficulties, are key parameters which have allowed the improvement of the quality of the teaching provided and a decrease in academic failure. Despite this great technological advance, universities around the world still suffer from several problems, namely the failure rate which can be quite significant, as well as student dropout and absenteeism. To overcome these problems, several studies have been conducted to better understand students profiles, in order to improve their working conditions, increase the success rates, and offer an adapted teaching process [1].

This kind of research is often related to the context of study: university courses, pedagogical system, as well as the social circumstances of the students. Currently, many university researchers are trying to exploit the availability of student history data in order to analyze the students' profile, reflecting the specificity of their educational institution, given that the student failure analysis depends heavily on the teaching environment and the student's context[5].
Indeed, different approaches were used to predict student profiles; such as: Machine learning techniques, deep Learning and multi-criteria analysis techniques.
Machine learning techniques are used in the educational domain for variant purpose such as: predicting student performance, analyzing student interaction with courses, and offering an adapted learning process
Multi-criteria analysis techniques allow us to benefit from expert knowledge so as to analyze learning situations and better describe student profiles. It can be exploited for reasons such as studying factors that contribute to student absenteeism, detecting the characteristics that embody a successful student and detecting groups' properties in order to help with pedagogical choices
In this paper, we propose a Multilevel Hybrid System for student failure prediction based on the combination of variant machine learning techniques and also one of the most used a multi-criteria decision making technique which is AHP. This system proposes multi-level prediction approaches that use the appropriate technique depending on the context of our University.
In this research, we first expose the architecture of our hybrid system then further develop the machine learning prediction steps. The rest of this paper is organized as follows: Section II explores the related works done in educational data mining and AHP. Section III introduces the theoretical background used in this study. Section IV describes the system's architecture and the machine learning steps process. Section V exhibits the results of the experiment following the application of data mining classification techniques and association rule mining. A discussion is elaborated in Section VI, and the conclusion and future work are given in Section VII.

Manuscript received September 5, 2019
Manuscript revised September 20, 2019

2. Related Works

Educational Data Mining (EDM) is the name dedicated to the application of data mining techniques in the educational domain. It is an emerging discipline that is concerned with the application of data mining techniques for the exploration of data produced in a learning context to better understand students and to improve their learning [6], [7].

In study [8], the authors propose a methodology for the implementation of a data mining project. The main aim of this study was to analyze the performances of different methods of data mining and class university students according to their academic performance results. The achieved results reveal that the decision tree classifier (J48) performs best (with the highest overall accuracy), followed by the rule learner (JRip) and the k-NN classifier, taking into consideration the fact that all tested classifiers are performing with an overall accuracy below 70 %.This means that the error rate is high and the predictions are not very reliable.

The authors of [9] provide an overview of data mining techniques to predict student's performance. The goal is to effectively improve student success. The neural network method gave the highest prediction accuracy due to the influence of the main attributes. These attributes are the hybridization of two characteristics, which are internal and external evaluations.

In [10], the authors use data mining in an educational environment. The analysis shows the potential of the association rules mining algorithm in improving the efficiency of academic planners, and reveals some hidden patterns of failed student modules. The analysis serves as a basis for academic planners to make academic decisions and help reduce the failure rate.

In[11], a classification task is used in the students' database to reduce the failure ratio and take appropriate steps for exams in the next semester. This study helps students and teachers improve the student division using the decision tree method.

The study in [12] aims to identify students who need special attention in order to reduce the ratio of failure and take the appropriate action at the right time. This study shows that the academic performance of students does not always depend on their own efforts but that there are other factors that have a significant influence on their performance, such as living location and medium of teaching with a probability of 0.7862 and .07225 successively.

The authors in [13] mainly worked on existing data from the university database. The study describes the characteristics that differentiate first-year career choices from a university sample: demographics, personality, social support and socio-economic characteristics. This analysis uses the decision tree algorithm in order to predict student's results and, subsequently, the choice of orientation.

The research goal in [14] is to guide students who have a real need for support, and provide an optimal distribution of teaching resources in order to curb academic failure. For that, the research aimed to classify students, as early in the academic year as possible. The study presents the results of the application of discriminant analysis, neural networks, random forests and decision trees. The prediction rates obtained in validation are not remarkable. However, discriminant analysis - and, to a lesser extent, neural networks and random forests - seemed to be able to lead to interesting results.

In addition to the data mining techniques, other approaches also allow the analysis of student profiles and facilitate the decision support. Indeed, several research projects propose to exploit the multi-criteria analysis techniques in educational domain.

The authors in[15]used fuzzy AHP to classify the factors that reduce student absenteeism in engineering schools. They analyze variant criteria such as: family problems, health, lack of motivation, psychic factors, evaluation system…. [16]also used the Fuzzy AHP method for the different purpose of evaluating student projects in order to choose the best one. The criteria adopted are: content, design, technique and presentation in addition to linguistic variables that are transformed in Fuzzy value using rule base tables. The technique is also adopted in [17] to evaluate the skills of Chinese teachers and in [18] to evaluate whether the innovative education can be used to build the future of high-quality talent.

Several other researchers used the AHP method in the educational domain. In fact, study [19] classifies the factors that lead Indian students to feel stress. [20] Compared three branches of computer science in order to find out the best one for students. [21] Used the method to choose the best students; students were evaluated on variants criteria describing their academic performance such as: personal skills, extra-curricular activities. Other research studies explore the potentials of AHP to analyze student success. Indeed, the authors in [22] used the method to rank the factors that have the most effect on the success of online learning.

In our research, we benefit from the AHP method and machine learning techniques in order to develop hybrid and multi-level prediction systems that allows an interesting description of students' profiles and also detect the most vulnerable ones. We first used the AHP to analyze the factors that affect student failure in their first semester. Secondly, we benefited from the potentials of supervised learning techniques to predict the final grades of students depending on their marks in the first semester, and used association rule mining technique to find any possible relationship between a student's final marks and that of certain modules, with the aim to detect those that have a greater effect on the failure of students.

3. Background

In this section, we briefly define the theoretical background of the concepts and techniques exploited in our study.

3.1 AHP

AHP is a multi-criteria analysis technique which attempts to benefit from mathematics and expert knowledge to classify criteria and found knowledge. It is used in almost all domains and in various situations to help in decision-making. AHP is one of the most used multi-criteria methods of decision support integrating several criteria and arriving at a justified choice of technology. The decision is then said to be rational, systematic and justified. This method is developed by Thomas L. Saaty in 1980. It consists of breaking down any decision-making problem into a hierarchy of sub-problems that can be analyzed independently. The advantages of this method include:

- Its ability to simplify complex situations.
- The choice of criteria and the performance rating are often simply made and understandable.
- The method streamlines the process leading to choices.
- The method is a useful negotiation tool for debates between users.

3.2 Feature Selection

Feature selection (FS) is a method with the objective of selecting or extracting an optimal subset of relevant characteristics for a criterion previously set. FS reduces the size of the sample space and makes the dataset more representative of the problem. Indeed, its main objectives are to facilitate visualization and understanding of the data, reduce the storage space needed, reduce learning and use time and identify the relevant factors.

Three main approaches are used on feature selection:

- Filters approaches: solely based on the dataset characteristics, these methods select the variables independent of the method that will use them.
- Wrappers approaches: unlike filter approaches that completely ignore the influence of selected variables on the performance of the learning algorithm, "wrappers" approaches use the learning algorithm as an evaluation function.
- Embedded Approaches: These methods perform variable selection during the learning process. The subset of variables thus selected is chosen as to optimize the learning criterion used.

3.3 Machine Learning Techniques

1) Decision Tree: A decision tree is a decision support tool that represents a set of choices in the form of a tree. The different possible decisions are presented in the leaves of the tree, and are reached according to decisions made at each stage[23], [24]. The C4.5 is decision tree algorithm that is based on ID3 and has some more potential. It allows:
   a. Managing attributes with missing values,
   b. Post-pruning the tree to avoid over-fitting;
   c. Manipulating continuous values (by "discretizing" them when setting into a tree)[25], [26].

2) Naive Bayes: A Naive Bayes classifier is a probabilistic classifier based on the application of Bayes theorem with the naive hypothesis; that is, the explanatory variables (Xi) are assumed conditionally independent of the target variable. Despite this strong assumption, this classifier appears very efficient in many real-world applications and is often used on data flows for supervised classification [27], [28].
3) Support vector Machine: SVM or Support Vector Machine is a linear model for classification and regression problems. It can solve linear and nonlinear problems and work well for many practical problems. The idea of SVM is simple, which is to create a line or a hyperplane that separate the data into classes. in [29][30], the authors demonstrated that the Support Vector Machine method has acquired the highest prediction accuracy in identifying students at risk of failing.
4) Association Rules: This model extracts frequent motifs, associations, correlations, or co-occurrence links of data expressed as conditional implication rules. It considers the data set as instances (rows) composed of a set of values called items. Association rules algorithms extract significant links between occurrences of values in the same instances [5], [31], [32].

3.4 Performance Metrics

A detection model is not perfect, as it can make prediction errors. Validation step measured the relevance of the generated model, before using it. Based on the confusion matrix, several measures are derived to quantify the performance of a classifier according to different points[33]–[35]:

- Accuracy, is the proportion of misclassified, it estimates the probability of misclassifying an individual randomly selected in the population when applying the prediction model.

- Precision represents the number of correctly classified data.
- Recall represents the amount of data correctly classified, such as positive, for the amount of positive data in the data corpus.
- F-measure which is a popular measure that combines precision and recall.

To better evaluate our model and have more information about prediction error, in this paper, we used and compared the machine learning techniques using these four coefficients.

3.5 Parameter optimization

It provides the optimal parameter values to determine the best model by comparing the suitability of the performance of the different iterations. The parameters change according to the algorithm used. For each iteration, the technique uses different parameter values until the optimal combination of parameters; to build the model; is found.

4. Our Proposed System

4.1 System Description

Our proposed prediction system is structured in three major levels (Fig. 1); each level is composed of a set of steps which use the most appropriate technique to the context. The first level is done when the student first begins university.
We analyze student global characteristics and calculate student risk factor depending on AHP classification results. This risk factor detects the students that are at high risk of failing their first semester, in [36], we have presented the implementation of AHP steps. After the first semester results, a prediction process is done using machine learning techniques to detect the students that are at high risk of failing their year depending on their semester 1 marks. A deep analysis is also offered to study the vulnerability level of students.
In these two steps we first realize feature selection to select the most important attributes, and used decision tree C4.5 algorithms that proved its performance compared to other machine learning techniques and realize predictors' optimization to improve the performance of our model. The last level in our system consists of the application of association rule mining techniques to detect the module that affects students' failure the most. Figure 1 displays our system process.

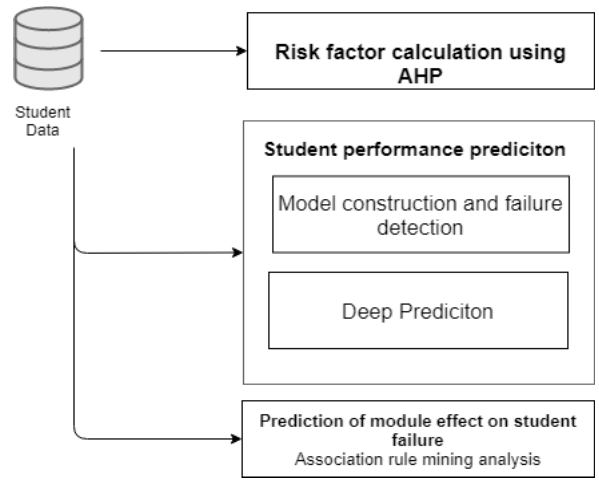

Fig. 1 System process

4.2 Student Performance Prediction Process Description

Figure 2 shows the prediction steps of our process. After dataset exploration, we carried out the preprocessing step on which we defined a set of rules to solve the problem of missing data. Next, we performed a feature selection based on information gain to reduce the size of the dataset and increase the prediction performance. This allowed as to collect the most relevant attributes to our study. In the third step, we elaborated on the prediction model to detect the failing students.
The prediction is made using three machine learning techniques to compare their performance and find the most relevant one. The techniques adopted are: C4.5 decision tree, SVM and Naïve bye algorithm. The prediction is made based on two classes (failure and success) and the performance is calculated following the parameter optimization step. To best perform the comparison we applied these techniques on variant dataset size and calculate their relevance via four coefficients (accuracy, precision, recall and F-score).
To gain more detail on students' profiles and offer a deep prediction, we realized a second prediction to define the level of vulnerability of the students. In this case we separate failing students into two classes that are (failure, at-risk) followed by a parameter optimization similar to the one in the previous step.

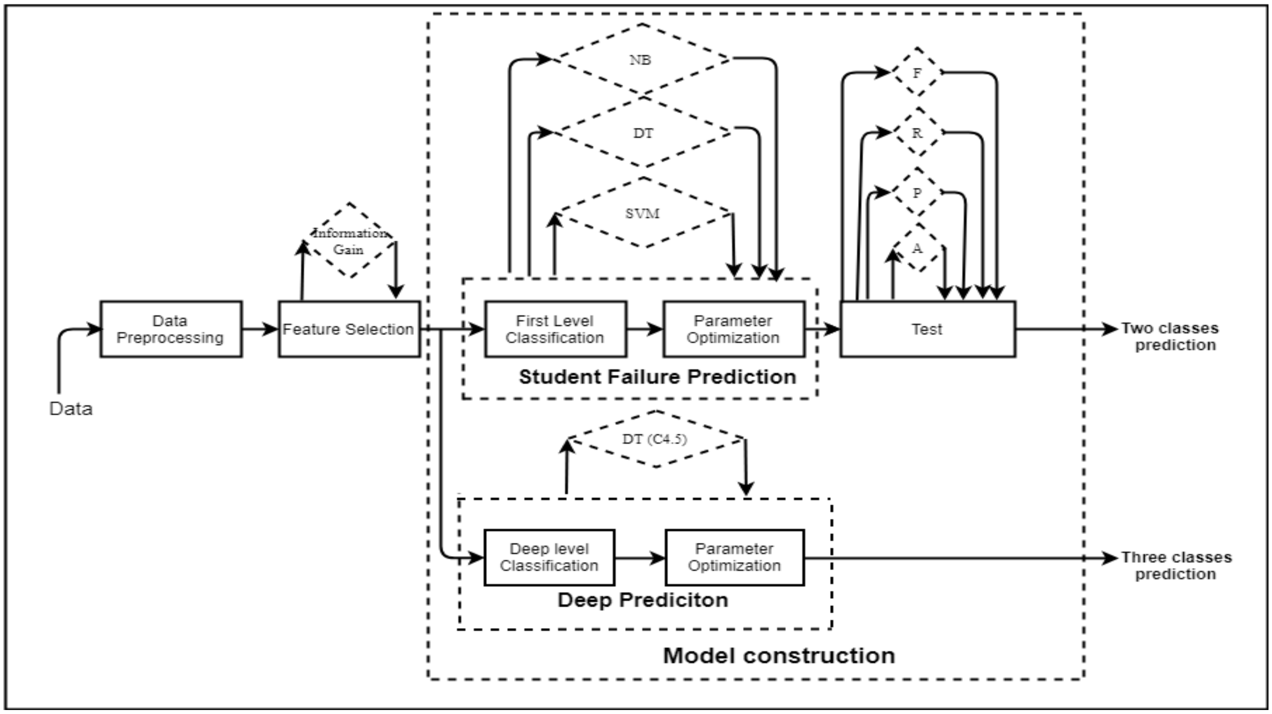

Fig. 2 Prediction steps process

5. Experiment

The objective of this part is to identify the most efficient machine learning techniques. The comparison is accomplished by taking into consideration different dataset sizes and different performance metrics

5.1 Dataset

The dataset used is extracted from the Apogee database of FSBM which is an integrated management software package used for the management of registrations and student files. Our dataset contains the characteristics of more than 3000 students. These characteristics regroup:

1) Personal information that contains student identity such as their surname, first name, address, date of birth, social information such as parent functions, family status…
2) Administrative registration, which relays to the annual inscription in major steps.
3) Pedagogic information which contains information on modules by year.
4) Results, which contains results of modules (mark, validation, session...), grades and results in semesters, and diplomas validations…

The number of students is 3911 and each student is characterized by 20 academic features. Table 1 detail these features.

Table 1: Student related variables

Variable	Description	Possible Values
LIB_BAC	Baccalaureate series	{Physic and Chemistry Sciences, Mathematical Sciences, life and Earth Sciences, Other }
LIB_MNB	Baccalaureate certificate	{very good, good, fairly well, fair}
Major	Major (discipline) of university study	Mathematics(MA), Chemistry(MC), Mathematics Computer Science(MI), Physic(MP), Earth Science(TU), Life Science(VI)
100	Step: validation result of the year ADM: admitted NO_ADM: not admitted	{ ADM,NonADM }
110	semester I validation result V: Validated NV: Not validated	{ V,NV }
120	SemesterII validation result	{ V,NV }
111,112,113, 114,115,116, 117,121,122, 123,124,125, 126,127	Validation results of modules ABSN :unjustified absence	{ V,NV,ABSN }

5.2 Data Preprocessing

In the first exploration of the database, the number of students was 5864 with 26.04% of missing values within all features of all the students. Our preliminary preprocessing was based on the application of some automatic rules in order to replace some missing values. These rules help to give performance decisions based on the number of absenteeism, validation and non-validation of modules for each student. After this preprocessing step, we were able to reduce the missing values to 14.3%, and detect and eliminate 1953 students who dropped out of the university and who will not be part of our study that focuses on failure. The student number has become 3911, and, after removing the remaining students with missing data (data cleaning), the final student numbers used in this study is 2497 students.
In this study, we are interested in the prediction of the final marks of students with different majors. Table 2 shows the number of students for each major, and due to the reason of the small number of students majoring in TU (Earth Science), the study focused on the remaining five majors (MP, MC, VI, MA, MI).

Table 2: Number of Students per Major Visualization

Major	*MP*	*MC*	*VI*	*MA*	*MI*	*TU*
Nbr of students	842	705	665	152	126	7

5.3 Feature Selection

In our study, 20 attributes are chosen for the prediction in the first step. The optimal number of variables is not known a priori. The use of a rule to control the selection-elimination of variables will allow us to minimize the processing time and reduce the number of attributes used in the study to use only the most relevant attributes in the end.
Table 3 shows the result of attributes Reduction by Weight by Information Gain for each major and for the two classes Semester II and Step (120 and 100) label. The table shows that, for both classes, the LIB_BAC feature is excluded, which shows that the appropriate subset of the features consists of the remaining nine attributes (LIB_MNB, 110 111, 112, 123, 114, 115, 116 and 117).

Table 3: Feature Selection by Weight information gain

	Class: Semester II									
	111	**112**	**113**	**114**	**115**	**116**	**117**	**110**	**LIB_BAC**	**LIB_MNB**
MP	0.452	1	0.306	0.654	0.886	0.726	0.148	0.766	0	0.694
MC	0.661	0.742	0.497	1	0.553	0.879	0.359	0.579	0	0.586
VI	0.802	0.879	0.447	0.632	0.597	0.654	0.509	0.683	0	1
MA	1	0.905	0.759	0.398	0.603	0.083	0.284	0.319	0	0.183
MI	0.095	0.625	0.691	0.471	0.699	0.100	0.164	0.015	0	1
	Class: Step									
MP	0.539	1	0.310	0.839	0.859	0.718	0.124	0.555	0	0.451
MC	0.651	0.842	0.489	1	0.746	0.733	0.165	0.673	0	0.504
VI	0.729	0.591	0.377	0.765	0.620	0.744	0.341	1	0	0.934
MA	0.601	0.715	0.608	0.346	1	0.079	0.026	0.357	0	0.121
MI	0.315	1	0.682	0.596	0.489	0.095	0.171	0.070	0	0.731

5.4 First Level Classification.

In this first level we attempt to predict whether the students will fail or pass their first year. The classification algorithms study was done for the three periods (Step, Semester I and Semester II), with the knowledge that the input attributes were only the modules of semester 1. This was done to be able to predict students' success based on their learning results in their first semester. This therefore allows the decision-maker to propose scenarios of solutions that can help these students succeed their year with an effective prediction model.
A Split Data (two partitions with ratios 0.7 and 0.3) operator is used and compared in order to predict student performance, and the effectiveness of the performance classification was measured by Accuracy, Precision, Recall and F-measure.
To improve our prediction results, we used the parameter optimization (Grid) for each algorithm. Below are the parameters we optimized for each algorithm:

- SVM: Kernel type, Gamma kernel, kernel degree, C
- DT: Criterion, Minimal gain, Maximum depth, Minimal leaf size, Minimal size for split
- NB: Laplace correction

Fig.3 and Fig.4 show the prediction results after parameter optimization for each algorithm. Fig.3 shows the prediction results for the MA and MI majors. DT outperformed the other algorithms for the Step class for both streams, while the SVM was good for the Semester II class. Fig.4 shows that DT has surpassed the SVM and the NB for almost all the performance metrics for the three majors (MP, MC and VI). The effect of the parameter optimization is very remarkable, and the results show a value increase for all metrics.
The results show that the decision trees have given good results for most sectors and classes.

Fig. 3 Prediction results before and after parameter optimization for MA and MImajors for both classes Semester II and Step

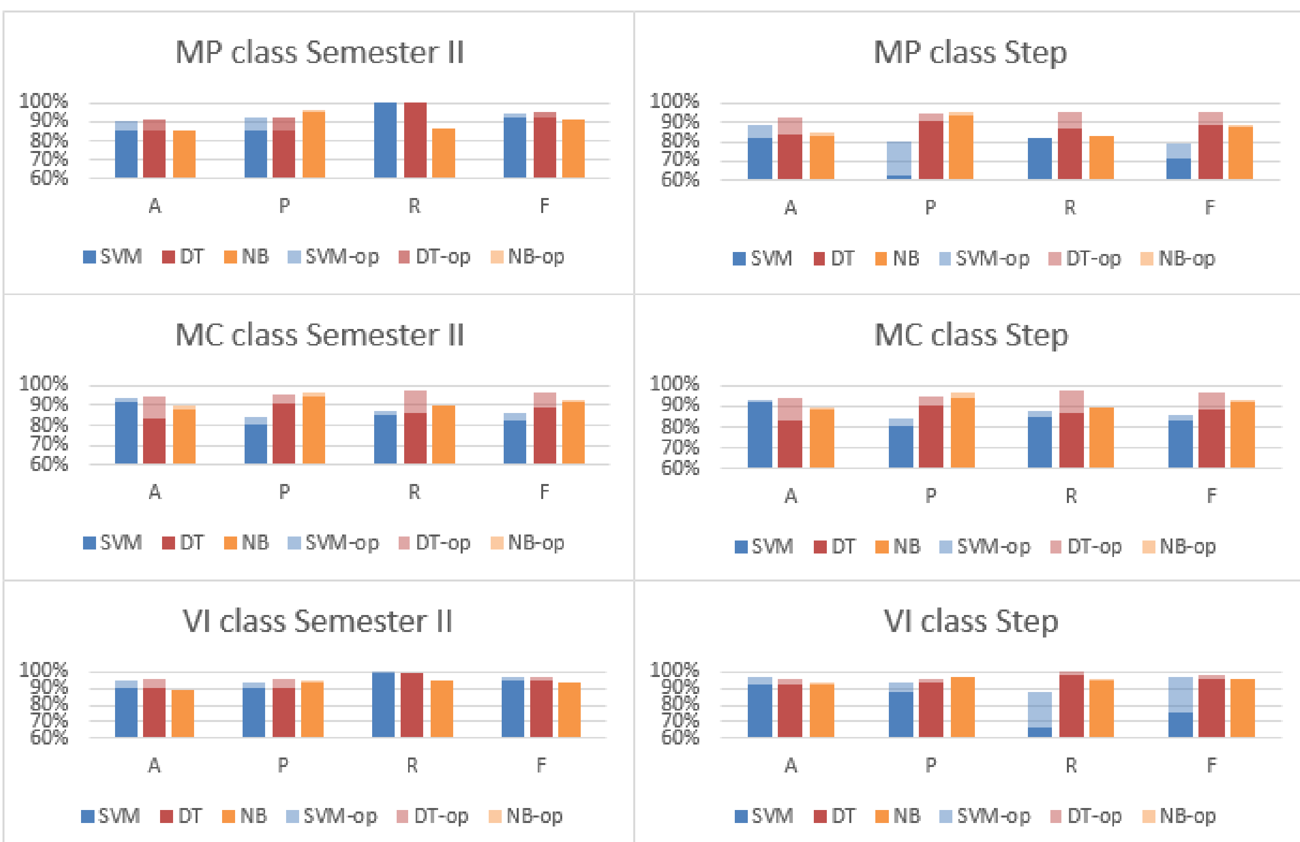

Fig. 4 Prediction results before and after parameter optimization for MP, MC and VI majors for both classes Semester II and Step

5.5 Deep Prediction Level

In this step we predict the vulnerability level of students. In fact, we define a non-validation scale that splits the failing students into two groups:

- At risk: students who have a grade strictly lower than 10 and greater than or equal to 9 in the overall average of the Step.
- Fail: students who have a grade strictly lower than 9 in the overall average of the Stage.

This division will allow us to improve decision makers' intervention and allow them to define the tutoring process according to students risk level. In this step, our main goal is to decide on the appropriate form of intervention according to the situation (class, label) of the student.

Table 4 shows that the deep prediction performed well especially when it came to the majors carrying the largest number of students (MP, MC and VI). For the two other majors (MA and MI) the accuracy is between 85.21% and 85.71% after optimization.

Table 4: Deep prediction results

	StepWithtwo classes Class « At Risk » Class « NonADM »	
	A	A-op
MP	76,59%	88.82%
MC	100%	100%
VI	80.31%	90.57%
MA	79.07%	86.21%
MI	67.57%	85.71%

5.6 Association Rule Mining

Association rule mining is one of the unsupervised data mining algorithms. These algorithms are very useful in discovering interesting relationships and hidden information in data. For the same purpose, we used them to find relationships between the attributes we have, in order to extract the modules that have a greater effect on the success of students for a given semester. These techniques were applied for each major in the three periods (step1, semester 1 and semester 2).

In order to analyze students' results and find eventual relationships between their success and module validation, we had to study the student data for each major separately.

Table 5 provides details regarding an instance of the rules generated for each major and each period (Semester I, Semester II, and Step) using Apriori algorithm. The rules generated show that for the Semester I class, students who could not validate modules 112 and 114 are most likely to fail in this class with support of 0.521 and 0.518 respectively. For the semester 2 class, the non-validation of the module 126 leads to the non-validation of this class with a support 0.771 followed by the modules 121 and 124. For the class Step, the rule with the most important support 0.717 show that the failure in the year is strongly linked to the failure in the second semester, followed by modules 126,121,124 and semester I whose non-validation influence on the non-validation of the step.

Table 5: Rules extracted from the datasets of student academic performance for each period

Association rules			Support	Confidence
Semester I				
112=NV	=>	110=NV	0.521	0.889
114=NV	=>	110=NV	0.518	0.914
111=NV	=>	110=NV	0.483	0.850
112=NV,114=NV	=>	110=NV	0.429	0.976
116=NV	=>	110=NV	0.416	0.914
115=NV	=>	110=NV	0.412	0.918
Semester II				
126=NV	=>	120=NV	0.771	0.928
121=NV	=>	120=NV	0.673	0.940
124=NV	=>	120=NV	0.667	0.951
126=NV,124=NV	=>	120=NV	0.632	0.966
126=NV,121=NV	=>	120=NV	0.629	0.969
125=NV	=>	120=NV	0.580	0.982
Step				
120=NV	=>	100=NonADM	0.717	0.842
126=NV	=>	100=NonADM	0.684	0.824
120=NV,126=NV	=>	100=NonADM	0.678	0.880
121=NV	=>	100=NonADM	0.619	0.864
120=NV,121=NV	=>	100=NonADM	0.615	0.914
124=NV	=>	100=NonADM	0.612	0.871

6. Discussion

The prediction of students' failure is one of the objectives of the FSBM in order to reduce the critical percentage of failure. In this study, we were able to recognize the modules that have more effects on the success and failure of students in each major. We were also able to predict the step and the semester 2 validations based on the first semester results. The two results can help the decision-maker enormously:

- Defining the modules that must have more review sessions
- Identifying the students that may have problems in succeeding, in order to give them more attention.

The first attribute we analyzed was the students' modules results, which was very important in order to be able to solve student failure problems in the learning level. The results achieved by applying the three different classification algorithms reveal that Decision Tree provides good precision in predicting failure for the majority of majors. For the rules extracted using Apriori algorithm, we notice that, for each period, there are modules which influence the validation of the student, and that the non-validation of the Step class is strongly linked to semester 2 and its modules.

7. Conclusion and Future Works

Our project aims to analyze students' profiles in order to improve their success rate and reduce students' dropout rate. In this research, we focus our study on first year data, because it is the one that suffers from the highest failure and dropout rate. We first exploit classification algorithms to predict students' final marks depending on their first semester marks for each major using different data mining algorithms. The results are then compared using different performance metrics. This allowed us to define the most fitting techniques for our data analysis and will help the educational responsible to detect the students that have learning problems in order to propose targeted tutoring processes. In the second step, we benefited from association rule mining technique to analyze the students' success depending on their final marks in modules; we were able to detect modules that had a greater effect on students' final results. This finding can help deciders with defining, accompanying and upgrading strategy to help students overcome any challenges faced with specific modules (several are not specialty modules).

These results are the first part of our project at the FSBM student profile analysis. In the next step, we aim to analyze students' profiles in the five years of the last accreditation to have a more complete review of student success. This analysis will be extended to take into account the eventual effect of other factors -such as social and personal factors - on student failure or dropout at the Faculty of Sciences Ben M'sik.

References

[1] S. Rajkaran et K. J. Mammen, « Identifying Key Performance Indicators for Academic Departments in a Comprehensive University through a Consensus-based Approach: A South African Case Study », Journal of Sociology and Social Anthropology, vol. 5, no 3, p. 283-294, sept. 2014.

[2] J. Y. Chung et S. Lee, « Dropout early warning systems for high school students using machine learning », Children and Youth Services Review, vol. 96, p. 346-353, janv. 2019.

[3] A.-S. Hoffait et M. Schyns, « Early detection of university students with potential difficulties », Decision Support Systems, vol. 101, p. 1-11, sept. 2017.

[4] S. M. Redha, « The Prediction of the Rate of the Dropout of the Primary Schools Students by Using the Genetic Algorithm », vol. 9, no 2, p. 17, 2018.

[5] J. Kong, J. Han, J. Ding, H. Xia, et X. Han, « Analysis of students' learning and psychological features by contrast frequent patterns mining on academic performance », Neural Computing and Applications, oct. 2018.

[6] C. Romero et S. Ventura, « Educational data mining: A survey from 1995 to 2005 », Expert Systems with Applications, vol. 33, no 1, p. 135-146, juill. 2007.

[7] P. Guleria et M. Sood, « Data Mining in Education : A Review on the Knowledge Discovery Perspective », International Journal of Data Mining & Knowledge Management Process, vol. 4, no 5, p. 47-60, sept. 2014.

[8] D. Kabakchieva, « Predicting Student Performance by Using Data Mining Methods for Classification », Cybernetics and Information Technologies, vol. 13, mars 2013.

[9] A. M. Shahiri, W. Husain, et N. A. Rashid, « A Review on Predicting Student's Performance Using Data Mining Techniques », Procedia Computer Science, vol. 72, p. 414-422, janv. 2015.

[10] E. Chandra et K. Nandhini, Knowledge Mining from Student Data. .

[11] B. Kumar et S. Pal, « Mining Educational Data to Analyze Students Performance », International Journal of Advanced Computer Science and Applications, vol. 2, no 6, 2011.

[12] B. K. Bhardwaj et S. Pal, « Data Mining: A prediction for performance improvement using classification », vol. 9, no 4, p. 5, 2011.

[13] S. Imane et E. B. Omar, « Prediction Approach for Improving Students Orientation in University: Case of Sidi Mohammed Ben Abdelah University (SMBAU) », International Journal of Multimedia and Ubiquitous Engineering, vol. 10, no 6, p. 9-16, juin 2015.

[14] J. F. Superby, « Determination of factors influencing the achievement of the first-year university students using data mining methods », p. 8.

[15] G. Fatima-zahra et D. Khadija, « Modernization of a domain E-orientation metamodel », in 2016 11th International Conference on Intelligent Systems: Theories and Applications (SITA), Mohammedia, Morocco, 2016, p. 1-6.

[16] C. Ayca et K. Hasan, « An application of fuzzy analytic hierarchy process (FAHP) for evaluating students project », Educational Research and Reviews, vol. 12, no 3, p. 120-132, févr. 2017.

[17] Wang A.P., « Application of fuzzy synthetic evaluation model to the assessment of competency of chinese teachers », Chemical Engineering Transactions, vol. 46, p. 415-420, déc. 2015.

[18] S. Mondal et J. Mukherjee, « A Fuzzy Analytical Hierarchy Process in Innovative Education - An Approach »,

International Journal of Computer Applications, vol. 150, no 9, p. 47-52, sept. 2016.
[19] M. Aitdaoud et al., « Standardized modeling learners ontology «SLMonto» to support interactive learning environment », in 2015 5th International Conference on Information & Communication Technology and Accessibility (ICTA), Marrakech, 2015, p. 1-6.
[20] A. Yani et L. Darmawan, « Competence Making on Computer Engineering Program by Using Analytical Hierarchy Process (AHP) », International Journal of Advanced Computer Science and Applications, vol. 6, no 9, 2015.
[21] N. A. Baki, N. I. Norddin, et W. A. W. Azaman, « Application of Analytic Hierarchy Process for Selecting Best Student », p. 5, 2017.
[22] M. Mehregan, « Application of Fuzzy Analytic Hierarchy Process in Ranking Modern Educational Systems' Success Criteria », International Journal of e-Education, e-Business, e-Management and e-Learning, 2011.
[23] A. Topîrceanu et G. Grosseck, « Decision tree learning used for the classification of student archetypes in online courses », Procedia Computer Science, vol. 112, p. 51-60, 2017.
[24] N. F. Kirby et E. R. Dempster, « Using Decision Tree Analysis to Understand Foundation Science Student Performance. Insight Gained at One South African University », International Journal of Science Education, vol. 36, no 17, p. 2825-2847, nov. 2014.
[25] S. V. K. Kumar et S. Padmapriya, « An Efficient Recommender System for Predicting Study Track to Students Using Data Mining Techniques », vol. 3, no 9, p. 3, 2014.
[26] C. Vialardi Sacín, J. Bravo, L. Shafti, et A. Ortigosa, « Recommendation in Higher Education Using Data Mining Techniques. », 2009, p. 191-199.
[27] Y. Zhang, S. Oussena, T. Clark, et H. Kim, « Use Data Mining to Improve Student Retention in Higher Education - A Case Study. », 2010, p. 190-197.
[28] U. Pujianto, E. N. Azizah, et A. S. Damayanti, « Naive Bayes using to predict students' academic performance at faculty of literature », in 2017 5th International Conference on Electrical, Electronics and Information Engineering (ICEEIE), Malang, 2017, p. 163-169.
[29] G. Gray, C. McGuinness, et P. Owende, « An application of classification models to predict learner progression in tertiary education », in 2014 IEEE International Advance Computing Conference (IACC), Gurgaon, India, 2014, p. 549-554.
[30] E. B. Costa, B. Fonseca, M. A. Santana, F. F. de Araújo, et J. Rego, « Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses », Computers in Human Behavior, vol. 73, p. 247-256, août 2017.
[31] Z. Abdullah, T. Herawan, A. Noraziah, et M. M. Deris, « Mining Least Association Rules of Degree Level Programs Selected by Students », International Journal of Multimedia and Ubiquitous Engineering, vol. 9, no 1, p. 241-254, janv. 2014.
[32] D. M. D. Angeline, « Association Rule Generation for Student Performance Analysis using Apriori Algorithm », vol. 1, no 1, p. 5, 2013.
[33] K. T. Chui, D. C. L. Fung, M. D. Lytras, et T. M. Lam, « Predicting at-risk university students in a virtual learning environment via a machine learning algorithm », Computers in Human Behavior, juin 2018.
[34] R. K. Jena, « Predicting students' learning style using learning analytics: a case study of business management students from India », Behaviour & Information Technology, vol. 37, no 10-11, p. 978-992, nov. 2018.
[35] S. Helal et al., « Predicting academic performance by considering student heterogeneity », Knowledge-Based Systems, vol. 161, p. 134-146, déc. 2018.
[36] S. Nawal, H. Touria, et B. Faouzia,«Implementation of the Analytic Hierarchy Process for Student Profile Analysis. International Journal of Emerging Technologies in Learning», vol. 14, no 15.2019.