Implementation of the Analytic Hierarchy Process for Student Profile Analysis

Nawal Sael, Touria Hamim, Faouzia Benabbou
University Hassan II of Casablanca, Casablanca, Morocco
saelnawal@gmail.com

Abstract—Data science is the discipline that allows the exploration and analysis of data in order to extract useful and relevant information for decision making and problem solving. In the educational domain, human experiences need to be synthesized in order to improve the success rate and help the responsible to make the best informed decision. Analytic Hierarchical Process (AHP) is one of the most widely used multi-criteria analysis techniques in decision making. It allows building models for various problems even in the case of insufficient observation data.

This paper aims to benefit from the potentials of AHP technique, to analyze students' profiles. Our objective is to detect and classify the most important factors that increase Moroccan student dropout and failure. We expect that this study is the first one that explores AHP, studying the Moroccan context and describing student profiles depending on variant criteria. It reveals, on the one hand, that Moroccan student failure is strongly related to their family and behavioral characteristics. Indeed, lack of motivation, family instability and lack of responsibility are the top three factors causing failure at the university. On the other hand, student dropout is strongly related to studying context, namely the lack of orientation and repeated failures in modules. These findings will enable the decision makers to develop adequate solutions to overcome these two scourges.

Keywords—Profile Analysis, Analytic Hierarchy Process, student dropout, student failure.

1 Introduction

Decision making is a process that is always encountered in practical life. From the simple choice of a product to the adoption of one strategy or another, the question remains the same: how to make the right choice by taking into account the specificities of the criteria that influence the decision-making process. Nowadays, multi-criteria decision making is present in almost all situations and domains, for example, in industry, business, medicine, etc.

Education is one of the areas where science can facilitate decision-making, whether through artificial intelligence, machine learning or multi-criteria analysis methods. Various researchers used these techniques to analyze student profiles in order to help
Paper—Implementation of the Analytic Hierarchy Process for Student Profile Analysis

in decision-making about students’ problems to improve their outcomes. Student failure and dropout constitute one of the most critical phenomena that educational institutions suffer from. Several factors may be the cause, such as: social factors, psychological and behavioral ones as well as those in relation to student interactions with the courses. Data science techniques exploit the past experiences to analyze and predict the future. In our study, we have focused on the opinion of educational experts and their long experience in this field to analyze students’ data to discover the profiles that may be dropout or fail in their studies. Our objective is to apply the multi-criterion analysis techniques AHP to detect and classify the most important factors that increase Moroccan student dropout and failure and help identify vulnerable ones. The study is focused on Ben M’sik of Science College, which is one of the most important open access institutions belonging to the University Hassan II – Casablanca in Morocco.

This paper is organized as follows: Section 2 describe previous works on student profile analysis, especially for dropout and failure characterization. Section 3 details the Analytic Hierarchy Processes and its background. Section 4 develops the methodology adopted. Section 5 details the experimental results using AHP to classify the student failure factors first, then dropout ones. Finally, section 6 concludes the research and introduces the future work.

2 Profile Analysis in Educational Domain Related Works

The analysis of student profiles offers more and more information about students and helps the decision-makers to develop the most adapted decisions to a specific context. Machine learning [1], data mining [2] and multi-criteria analysis [3] techniques have contributed to the development of several studies in the field, their aim being to better understand students and to improve learning contexts.

The authors in [4] exploited data mining techniques to classify university students according to their academic performance results. They explore the potentials of variants algorithms such as decision tree classifier (J48) and k-NN classifier, and show that decision tree gives the best precision in their case study.

The researchers in [5] used Association rule mining technique to improve the performance of academic planners. Their objective was to find eventual hidden relations between modules and students failure.

The authors in [6] explored the potentials of machine learning techniques to identify students who may have problems in their studying process. They found that, in addition to the academic performance of students, some other factors have a significant influence on student failure such as: living location and teaching ways.

The authors in [7] proposed to classify the alternatives that slow down student absenteeism in engineering schools using Fuzzy AHP. The criteria they used are health, family problems, preparation without teachers, lack of motivation, class environment, psychic factors, evaluation system, distraction, and lack of responsibility.

In [8] the authors analyzed the factors that lead Indian students to feel stress that leads to mental harassment. Indeed, as the authors indicate, stress can be caused by
many complications such as school problems, financial problems, family problems and social problems.

In study [9], the researchers looked at evaluating the quality of innovative education to show whether it can serve as a solution that provides the environment with the facilities necessary for building the future of high-quality talent. This has the potential to motivate students to do better than the conventional mechanism.

In [10], the authors used the AHP method to facilitate the selection of students’ skills and professional expertise in order to orient the latter to enhance their academic ability and achieve success. They wanted to compare three branches of computer science to find out which one would be the best for students. In so doing, they compared these three branches using a test of competence, the interest of students and the content of its three branches.

In [11], the authors used the Fuzzy AHP method to evaluate projects developed by university students as part of a course and to choose the best project. The criteria used by professors who have experience in the field are content, design, technique and presentation. In addition, students’ projects are evaluated using linguistic variables used in ambiguous occasions that are transformed by the following in Fuzzy value using the rule base tables.

Authors in [12] used the AHP method to choose the best students. A preliminary evaluation led them to choose 5 students among 25. Next, they evaluated those students depending on variants criteria describing their academic performance (personal skills, extra-curricular activities…).

In [13], the researchers developed Fuzzy AHP model to evaluate the skills of Chinese teachers depending on certain criteria such as: Theoretical knowledge, pedagogical skills, teaching attitude, personality, etc.

In [14], the authors used the AHP method to identify and rank the factors that have the most effect on the success of online learning. They first grouped together the possible factors and then calculated the weights of these criteria via their own tool.

Authors in [15] combine the AHP method to Delphi method in order to analyze and model User Psychological Experience. The experiment is done at Chinese university.

3 Analytic Hierarchy Process Background

Educational data mining techniques are algorithms that explore data in order to extract hidden, relevant and useful knowledge in educational domain [16][17]. However, in some situations, there are no data to explore. Multi-criteria analysis techniques propose to exploit the expert knowledge about context and domains to extract useful and hidden information about a particular context. AHP is one of the most used ones. It attempts to benefit from mathematic 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 a multi-criteria analysis method 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. It helps to capture the two subjective
and objective aspects of a decision. Based on the available literature, there are multiple qualities of AHP. For instance [18]:

- AHP is simple, flexible, accurate and easy to understand.
- Its process relies on mathematical and logical reasoning to arrive at a decision making.
- AHP is an effective method and superior to other methods in the weight determination of multiple factors in a systematic and logic.
- Its consistency can be measured and controlled even in the presence of contradictory criteria.

The first step on AHP process is the identification of the decision purpose or the objective, as shown in Figure 1. The next step consists of carefully defining the hierarchy structure. It is composed of a set of criteria and sub-criteria that describe the elements that are supposed to influence the objective. At the last level, we define the different alternatives. After objective identification and hierarchy construction, The AHP generates a weight for each criterion according to an evaluation and comparisons did by decision-maker for each pair of criteria using Saaty Scale Table 1.
To calculate the weights for the different criteria, AHP create a set of pair-wise comparison matrices that indicates how often and important or dominant an element is compared to another one. We make a pair wise comparison of the elements according to their impact on the element located above them in the hierarchy.

The matrix $A$ (equation 1) is a real matrix $n \times n$, where $n$ is the number of evaluation criteria considered and $a_{ij}$ the relative importance of the criterion $i$ compared to $j$, assigned to the $(i, j)$ position of the pair wise comparison matrix.

Automatically, the inverse of the assigned number is associated with the position $(j, i)$.

$$
\begin{bmatrix}
1 & \cdots & a_{1n} \\
\vdots & \ddots & \vdots \\
a_{n1} & \cdots & 1
\end{bmatrix}
\text{Where } a_{ij} = 1/a_{ij}
$$

After that, we compute priority vector, which is the normalized Eigen vector of the matrix. It is calculated as:

- Calculate the sum of each column (equation 2)

$$
S_m = \sum_{i=1}^{n} c_{im} \text{ for } m = 1, \ldots, n
$$

- Each element of the matrix is divided by the sum of its column as seen in equation 3.

$$
t_{im} = a_{im}/S_m \text{ Where } i, m = 1, \ldots, n
$$

- At the end, we average the lines.

The weight is calculated by adding the sum of each column and dividing each element of the row by the corresponding column sum. To obtain the weight of each criterion / alternative, the sum of each line is divided by the number of elements in line as formulated in equation 4.

$$
P_j = \frac{1}{p} \sum_{i=1}^{n} a_{im}/S_m \text{ Where } i = 1, \ldots, n
$$

Finally, the AHP combines the weights of criteria and alternative scores, thereby determining an overall score for each option, which is a weighted sum of the scores obtained and the highest score corresponds to the option one has to choose [15].

In order to evaluate the consistency of the comparisons (responses) and verify the logical relationship between data, the author in [15] proposes to calculate a Consisten-

---

<table>
<thead>
<tr>
<th>Intensity (importance)</th>
<th>Priority</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Equal</td>
<td>Same importance</td>
</tr>
<tr>
<td>3</td>
<td>Moderate</td>
<td>Element moderately favored than other</td>
</tr>
<tr>
<td>5</td>
<td>Strong</td>
<td>Element strongly favored than other</td>
</tr>
<tr>
<td>7</td>
<td>Very strong</td>
<td>Element very strongly favored than other</td>
</tr>
<tr>
<td>9</td>
<td>Extreme importance</td>
<td>Element extreme important than other</td>
</tr>
</tbody>
</table>
cy Index parameter (IC). It allows us to represent the level of reliability of our judgment. IC is calculated by the formula detailed in Equation 5.

\[ IC = \frac{\lambda_{\text{max}} - n}{(n - 1)} \]  

\( \lambda_{\text{max}} \): The maximum eigenvalue of the matrix
\( N \): size of the matrix

The consistency ratio is calculated by Equation 6

\[ RC = \frac{IC}{RI} \]  

Where RI is a random consistency index RI [15], which is choosing depending on the matrix dimension. When RC value is smaller than 0.1, we can confirm that the judgments are valid.

4 Research Methodology

The first exploration of the FSBM database reveals that it contains only the students’ academic results, and some basic information about them. However, to have a general view of the factors that strongly influence students’ dropout or failure, we also need students’ social and psychological information, including financial situation, intellectual level of parents, and information on family stability. Taking into account these contexts, we decided to benefit from expert experiences instead, which is why we proposed to use the AHP to analyze students’ profiles to predict which profile is more likely to fail or dropout. To achieve this goal, we first conducted an exhaustive literature review to collect the factors that influence students’ studies. Following this step, paper and web questionnaires were designed and distributed to a group of Ben M’sik of Science college instructors (including sections coordinators and some department officials) in order to synthesize the factors we will work with. A second type of questionnaire was developed and administered to teachers so that the criteria and sub-criteria could be compared with each other using the Saaty scale Fig.2. The last stage involved using the Super Decisions software to calculate the weights of the factors (under the criteria in the hierarchy). The aim was to also classify them to find the most important ones. All those steps were implemented in the analysis of both student failure and dropout.

5 AHP for Decision Making in Moroccan Open Access Colleges

In this part of our study, we explore the potentials of AHP techniques to analyze the profiles of students to detect the most important factors that increase the failure and dropout rates in Moroccan open access colleges. The goal is to identify and classify the criteria that have the greatest impact on these phenomena.
5.1 Determination of candidate factors

After examining the relevant works cited above and looking at other research articles that have explored student failure and dropout, the main factors these studies highlight are as developed in Table 2 and Table 3.

Table 2. Students fail factors from the literature reviews

<table>
<thead>
<tr>
<th>Category</th>
<th>Factors found by the literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Society</td>
<td>Family instability</td>
</tr>
<tr>
<td></td>
<td>Financial problems</td>
</tr>
<tr>
<td></td>
<td>The proximity of the faculty</td>
</tr>
<tr>
<td></td>
<td>Addiction (drugs, social networks), relationship problems</td>
</tr>
<tr>
<td></td>
<td>The intellectual level of parents</td>
</tr>
<tr>
<td>Student</td>
<td>Absenteeism</td>
</tr>
<tr>
<td></td>
<td>Lack of motivation</td>
</tr>
<tr>
<td></td>
<td>Distraction</td>
</tr>
<tr>
<td></td>
<td>Lack of responsibility</td>
</tr>
<tr>
<td></td>
<td>Personal skills</td>
</tr>
<tr>
<td>Teaching system</td>
<td>Overload</td>
</tr>
<tr>
<td></td>
<td>Assessment system</td>
</tr>
<tr>
<td></td>
<td>Insufficient material for practical work, difficulty of courses in French</td>
</tr>
<tr>
<td></td>
<td>Lack of continuous controls</td>
</tr>
<tr>
<td>Teachers</td>
<td>Quality of content</td>
</tr>
<tr>
<td></td>
<td>interaction with the student</td>
</tr>
<tr>
<td></td>
<td>teaching skills</td>
</tr>
<tr>
<td></td>
<td>enthusiasm in teaching</td>
</tr>
<tr>
<td>Post bac level</td>
<td>Series of bac</td>
</tr>
<tr>
<td>(The French equivalent to a high school exit certificate)</td>
<td>Grading of the bac</td>
</tr>
<tr>
<td></td>
<td>Province ofthe bac</td>
</tr>
<tr>
<td></td>
<td>Level of the French language.</td>
</tr>
</tbody>
</table>

Table 3. Dropout factors picked from the literature

<table>
<thead>
<tr>
<th>Category</th>
<th>Factors that lead students to drop out of school</th>
</tr>
</thead>
<tbody>
<tr>
<td>Society</td>
<td>Family financial situation</td>
</tr>
<tr>
<td></td>
<td>Intellectual level of parents</td>
</tr>
<tr>
<td></td>
<td>The proximity of the faculty</td>
</tr>
<tr>
<td>University</td>
<td>Lack of orientation,</td>
</tr>
<tr>
<td></td>
<td>Difficulties related to teacher / learner relationships,</td>
</tr>
<tr>
<td></td>
<td>Repeated failure to validate modules,</td>
</tr>
<tr>
<td></td>
<td>Previous semester marks;</td>
</tr>
<tr>
<td></td>
<td>The complexity of the courses,</td>
</tr>
<tr>
<td></td>
<td>Health problems</td>
</tr>
<tr>
<td>Student</td>
<td>Age, Lack of responsibility,</td>
</tr>
<tr>
<td></td>
<td>Lack of motivation, Learning difficulties, Level of the French language,</td>
</tr>
<tr>
<td></td>
<td>low self-esteem</td>
</tr>
</tbody>
</table>

5.2 The hierarchy construction and weight calculation

In order to determine the factors that have a greater influence on the failure and dropout of students and define the hierarchy, a group of expert was contacted. The Figures 2 and 3 below detail the hierarchy that has been found.
5.3 Comparisons of criteria and factors (sub-criteria)

The next step in AHP process is the pair wise comparison between the criteria and sub-criteria using the Saaty scale. Here, it should be noted that we compared each level with the upper level. The judgment matrix created with pair-wise comparison must have a Consistency Ratio (IC) less than 0.1. In these cases, only the matrix is acceptable; otherwise, the matrix needs to be restructured.
To implement AHP and calculate the weights, we used the Super Decisions software, which allows us to enter the different comparisons and then synthesize the final weights. The tool first allows us to define our goal in the head of the structure (student failure factors detections in the first case and student dropout factors in the second one), next we entered the criteria in the first level and then the sub criteria associated to each criteria in the last level.

After the hierarchy construction, the software allows us to enter the comparison obtained after experts’ response analysis. It offers questionnaire structure for each node (criteria) comparing to its sub-criteria. We choose the appropriate comparison depending on Saaty scale. Figure 4 shows samples of the comparison matrix generated via experts’ questionnaires response analysis and implemented in super decision software. The top side one compares the criteria: student, society, teaching system, teachers and post bac level, between them depending on their impact on student failure. The down side one presents the comparison did for student sub criteria: absenteeism, lack of motivation and lack of responsibility. It compares the effect of each criterion under student factors. In this case, lack of motivation is five times more important than absenteeism and three times more important than lack of responsibility.

**Fig. 4.** Comparison Matrix for student failure analysis, between criteria (top), student node sub-criteria (down)

### 5.4 Weight calculation

Once the comparison was done, AHP generated a weight for each sub criterion via a priority vector. The figure 5 gives us the priority vector approximately as (0.10, 0.63, 0.25) for the sub criteria (absenteeism, lack of motivation, lack of responsibility).
Figure 6 shows a sample of comparison matrix for student dropout analysis (top side) and the priority vector (down side) for student criteria and sub criteria. We deduced the priority vector approximately as (0.51, 0.29, 0.08, 0.11) for the sub criteria (lack of motivation, lack of responsibility, learning difficulties, level of the French language).

In the same way, we entered the different comparisons between the criteria and the sub-criteria of each criterion, and Super Decisions Software calculated the weights for each comparison. At the end, we concluded with a synthesis to obtain the final weights.

5.5 Student failure factors classification

The last step of AHP for failure factors detection is the overall synthesized priority generation. Figure 7 presents the super decision classification (synthesis) of all criteria and sub criteria.
The factors classification obtained is:

- Lack of motivation (0.28)
- Family instability (0.18)
- Lack of responsibility (0.11)
- Interaction with students (0.10)
- Financial problems (0.07)
- Bac serie (0.06)
- Absenteeism (0.04)
- Difficulty of French courses (0.038)
- Teaching skills of teachers (0.035)
- Addiction (0.02)
- The level of the French language (0.01)
- Insufficient material for practical work (0.009)
- Lack of continuous controls (0.003)

Based on these findings, the top three factors that influence student success / failure are lack of motivation, family instability, and lack of responsibility and then the factors mentioned above.

5.6 Student dropout factors classification using AHP process

After analyzing teachers' answers and calculating weight by super decision software, the AHP implementation resulted in the factor classification presented in Figure 8. The factors (criteria and sub-criteria) classification obtained based on AHP process are:
• Lack of orientation
• Repeated failures to validate modules
• Lack of motivation
• Lack of responsibility
• Difficulties related to learner / teacher relationships
• Intellectual level of parents / financial situation of the family
• The level of the French language
• The difficulty of learning subjects

The AHP allowed us to classify the most important factors that influence student dropout. Based on these findings, the three main factors that cause students dropout in Moroccan open access faculties are: lack of orientation, repeated failures to validate modules and lack of motivation.

Unlike what we always think, the difficulties related to the level of the students’ language level is not one of the major factors that cause the university dropout. Other factors directly related to academic orientation and repeated failure have a greater effect on this phenomenon.

5.7 Discussion

As a result of the use of AHP in student profile analysis, especially in relation with social, psychological and behavioral characteristics; we found that student failure depends massively on the lack of motivation, family instability and lack of responsibility parameters. However, this kind of study is strongly related to its context and the results are different from one country to another and even from one institution to another. The analysis shows that in our institution context, motivation is a key factor. Indeed, integrating the faculty of sciences, as open access institution, is not really a
choice if they are not accepted in more prestigious institutions (Engineer school, Medicine Faculty. The social conditions of students also have a significant effect on motivation especially in our geographical context. Lack of motivation may also be due to insufficient interaction with the teacher as in [19] where the authors put the lack of interaction on the top, followed by the content quality and at last the interest in lessons.

For student dropout problem, the raisons depend on the lack of orientation, repeated failures to validate modules and lack of motivation. The orientation importance is justified especially in open access schools where students generally come without any coaching and post baccalaureate orientation; this finding is also confirmed in the report made by the Ministry of Higher Education [20]. Drop out dependent also on repeated failure (academic achievement) as was found in other researches such as [21][22][23] and on student motivation as researchers in [24] claim in their study.

This work allowed us to define the most important factors that affect university students’ dropout and failure. The results can help decision-makers to develop solutions for overcoming those serious problems and help students to achieve more success in their university studies.

The research methodology was very helpful in Moroccan student profile analysis especially in relation to social, psychological and behavioral characteristics. It proves that AHP possess is very important techniques to use especially in the case where the data source is very insufficient to explore with data mining or machine learning techniques. It can also be combined with data mining experiences since it allows us to combine student reel data analysis to expert experiences capitalization. However, even if, our results allow to have a good visibility on students’ failure and drop out. It has some limits: first, due to the absence of this kind of research in our context, we could not compare it to others. Next, the AHP process doesn’t allow to integrate uncertainty into the assessment of criteria weights.

6 Conclusion and Future Research

This study reveals that the AHP as a multi-criteria technique is an interesting palliative solution in the case of insufficient data. Based on the expert opinion, we could identify and classify the factors that contribute to an increase in the rate of failure and dropout of college student. This finding was very important in this case because it provides the profile of the student with a high risk of failure or dropout. Based on this system, the decision-maker can propose appropriate solutions. In our next work, we aim to benefit from the Fuzzy AHP implementation to integrate uncertainty to better improve the weight of each factor’s effect [25] and to confirm or regulate our obtained factors classification. Furthermore, we also aim to build a decision support system based on similarity calculations to detect the risk factors for a student to leave university or fail their final exam in their first semester.
7 Acknowledgement

Great acknowledgement to Sara Guebba, student of the SII Master of Ben M’sik of Science Faculty who has contributed to the realization of this work.

8 References

Nawal Sael is a professor of Computer Science and member of Computer Science and Information Processing laboratory at faculty of science Ben M’sik (Casablanca, Morocco). She received her Ph.D. in Computer Science from the Faculty of Sciences, University Hassan II Casablanca, Morocco, 2013 and her engineer degree in software engineering from ENSIAS, Morocco, in 2002. Her research interests include data mining, educational data mining, machine learning and Internet of things.

Touria Hamim received the engineer degree in information and communication systems engineering from ENSAJ, Morocco, in 2017. Currently, she is preparing her
PhD in computer science in faculty of Science Ben M’sik. Her research interests student profile modeling using machine learning.

**Faouzia Benabbou** is a professor of Computer Science and member of Computer Science and Information Processing laboratory. She is Head of the team "Cloud Computing, Network and Systems Engineering (CCNSE)". She received his Ph.D. in Computer Science from the Faculty of Sciences, University Mohamed V, Morocco, 1997. His research areas include cloud computing, data mining, machine learning, and Natural Language Processing. She has published several scientific articles and book chapters in these areas.