A Mobile Application for Early Prediction of Student Performance Using Fuzzy Logic and Artificial Neural Networks

https://doi.org/10.3991/ijim.v14i02.10940

Ann Nosseir (✉)
Institute of National Planning (INP), Cairo Egypt
The British University in Egypt (BUE), Cairo, Egypt
ann.nosseir@bue.edu.eg

Yahia Mohamed Fathy
The British University in Egypt (BUE), Cairo, Egypt

Abstract—Identifying students at risk, or potentially excellent students is increasingly important for higher education institutions to meet the needs of the students and to develop efficient learning strategies. Early stage prediction can give an indication of the students’ performance during their study years. This helps to tailor an appropriate learning strategy for weak or excellent students.

This work develops a novel framework for a mobile app to predict student performance before starting university education. The framework has three main components, namely, a neural network model that predicts GPA, a mobile app that tests basic knowledge in different domains, and a fuzzy model that estimates future student performance.

Keywords—Mobile app, Predictive Models, Fuzzy algorithm, Neural Network, Pedagogy

1 Introduction

Work has been done in identifying students’ performance, more specifically students “at risk” [13][10][16][3][17][20][15][26]. These students are at risk of not finishing their degree in the assigned period and may drop out. The reason why it is important to detect and solve this problem earlier is because it has negative consequences on different levels. This is an economic loss for families, universities and the public. These losses can be reduced if, in early stages, these students’ performance is predicted and, accordingly, an appropriate learning strategy is set for them. It is equally important to identify the outstanding ones. Having the right education program designed for them will have a positive impact.

Some universities, based only on students’ entry marks, take precautionary measures to maintain a certain level of education and dropout rate by minimizing the number of students at risk. For example, they could add extra courses such as math or English. This is to help students to engage in the university system as smoothly as possible. For
excellent students, universities encourage and support students to take part in national and international competitions and conferences. They pay for mentors to guide students and help them to gain more skills and succeed. Knowing the exact numbers of weak and excellent students helps universities to plan ahead and provide the needed services to support students to successfully complete their studies.

Students’ performance is not limited to their entry marks only. There are other factors that need to be considered. The work by [5][4][16][9][25] shows other factors that influence their success or failure. For example, factors such as gender and race are reasons for early withdrawal[22]. Additionally, social background, confidence, and motivation[19] greatly affect students’ performance. Checchi and Flabbi[8] and Aina[1] found that well-trained students integrate faster in the University’s system and there is a strong relationship between secondary school choice and parental educational background, cultural and financial factors and their success at university. Güner et.al.[13] classified these factors into the academic background of students, social, economic, and educational levels of students’ families, as well as the physiological and individual characteristics of students.

Dropout commonly happens after the first year of university. Researchers have shown that there is a strong relationship between first year academic success and graduating successfully in time [24]. Tinto[27] has observed that 73% of withdrawals occur within the first two years. Added to that, less than 80% students in early years in Queensland University of Technology managed to graduate within four years [9].

This novel work has a framework of three components which are: a neural network, fuzzy models and a question-based assessment application. The supervised neural network model considers socio-economic factors in addition to other factors.

This paper starts with related work. Then it describes the framework in three sections which are: the neural network model, assessment application, and the fuzzy model. It ends with discussion and conclusions.

2 Related Work

In the literature, there have been many research papers attempting to predict the GPAs of students using data mining models. These use either one model or a hybrid model. Cortz and Silva [7] collected real-world data about student grades, demographic, social and school-related features using school reports and questionnaires. They used k-means clustering, naïve Bayes and C4.5 decision tree algorithm on algorithms to classify the students’ failure in two core classes (Mathematics and Portuguese). Güner et. al. [13] applied three classifiers namely: Support Vector Classification (SVC) approaches, Least-squares support vector classification (LSSVC) and Radial basis function neural networks (RBFNN) to predict students’ university first year GPA. They built their models with 39 inputs that represented their socio-economic status e.g., the distance of hometown to university and other data related to their high school marks such as the quantitative score of the high school in national rankings. Other work [26] proposes a reduced training vector-based support vector machine (RTV-SVM) capable of predicting at-risk and marginal students. They used the marks of university students in
seven courses to develop their model. Djulovic and Li [11] implemented four models i.e., decision trees, Naïve Bayes, Neural Networks, and Rule Induction models to identify first year students who are more likely to drop out of university.

The fuzzy logic technique is used for uncertain facts to support the decision-making process. Jyothi et al.[29] proposed a fuzzy expert system for evaluating teachers’ overall performance. They used student feedback evaluation, teachers’ self-appraisal, assessment by peers, and university exams to predict the assessment results.

Ajiboye [2] developed a model using a fuzzy logic approach to predict the risk status of students based on how basic information correlates with the students’ academic achievement.

Goni [12] used a fuzzy logic-based expert system to predict student academic performance in Adamawa State University Mubi. Their input variables were the score of O’level result, type of secondary school attended and age of the students.

Rathore and Jayanthi [21] and Akansha et.al.[14] implemented a fuzzy inference system to predict student performance. Rathore and Jayanthi’s [21] model inputs are the 10th marks, 12th marks, B.Tech CGPA, M.Tech CGPA, and number of backlogs. The Akansha et.al.[14] model inputs are: exam 1, exam 2 and exam practical. Ulloa-Cazarez [6] built a fuzzy logic model to predict the performance of on-line students and their inputs of their model are the results of the Calif_U1 and Calif_U2.

The current work builds a novel hybrid model that has three components: a neural network and fuzzy models and a question-based assessment application. It uses real data to predict the general performance from the early registered data. It combines both data mining techniques of neural network and fuzzy logic. The following section describes the details of this work.

3 Predictive Student Performance Framework

The objective of this framework is to predict student performance. To achieve this, the framework uses university registration entry data and students’ answers to a set of questions in different domains. The architecture of the framework has three main components. The first component is a neural network model that predicts the expected GPA from the data provided in early stages i.e., before starting undergraduate studies. The second component is the assessment application. This application has a list of questions that examines the students’ knowledge in different domains. The third component is the fuzzy model. It predicts student performances from the predicted GPA and the percentage of correct answers. Figure 1. shows the details of the framework. The following sections describe in detail the three components.
3.1 Neural network model

**Methodology:** To develop the model, a number of steps were followed: data collection, data processing, and building the neural network model.

*Data collection:* To build a supervised model, the model considered data from university records. The data are for students’ graduated between 2009 and 2017. The data record attributes are graduation year, gender, address, country of school, prequalification certificate, prequalification percentage, English language entry mark, prep year average to predict the final GPA (see Fig. 2).

*Data preprocessing:* The data went into two phases of processing: coding and then normalization. In the coding, the data were converted into digital numbers. For example, for the years from 2009-2017, the corresponding numbers are 1-9. For the female-1 and male-0. The students’ addresses were from 65 areas. The digital number was...
converted based on the area, e.g., Gharbia =32. Regarding the students’ country of schooling, the data had 13 countries. The addresses of the schools were converted as follows: Egypt=1 Cameroon=2, United Arab Emirates=3, etc. The Prequalification Certificate was coded as follow Thanawy Amma=1, American High School Diploma=2, and IGCSE / GCSE=3.

In the normalization phase, each value in each attribute is normalized using the minmax normalization. This is equation (1) of the minmax.

\[
Y_i = \frac{(X_i - \min(X))}{(\max(X) - \min(X))}
\]  

(1)

Where \(X_i\) is the i data point and \(\min\) represents the minimum and \(\max\) represents maximum. So \(X_i\) converts to \(Y_i\)

**Network Building:**

*Topology of the network:* The network topology describes the arrangement of the neural network. Choosing the topology of the neural network is a difficult decision [28]. The network topologies available are numerous, each with its inherent advantages and disadvantages. For example, some networks trade off speed for accuracy, while some are capable of handling static variables and not continuous ones. Hence, to arrive at an appropriate network topology, various topologies such as multilayer perceptron, recurrent network, and time-lagged recurrent network were considered. Due to the nature of our case study data, which is static and not sufficiently large to enable the use of complex topologies, the multilayer perceptron was selected.

*Multilayer perceptron:* Multilayer Perceptrons (MLPs) are layered feed-forward networks typically trained with static backpropagation. Fig. 3 shows the topology of this model.

![Fig. 3. Neural network topology](http://www.i-jim.org)

**Learning algorithm:** Learning is actually the process of adapting or modifying the connected weights between neurons as a result of the mismatch between the targeted output and the desired output [7]. There are many techniques commonly used in learning algorithms, which include: Gradient Decent Backpropagation, Radial Basis Func-

---


http://www.i-jim.org
tion and the fastest known technique called Levenberg-Marquardt (LM). The activation or transfer function used in the hidden layers are log sigmoid or “logsig” while in the output layer is mainly pure linear or “purelin”.

Network training: The training will stop when Minimum Square Error or MSE is obtained, which is the minimum error between targeted and predicted. MSE is given.

The network was trained with the number of runs set to 3 and the Epoch set to terminate at 1000. The training performance was then evaluated using the following performance measures:

The Mean Square Error (MSE) in equation (2):

$$\text{MSE} = \frac{\sum_{i=0}^{P} \sum_{j} (d_{ij} - Y_{ij})^2}{NP}$$

where:
P = number of outputs of processing element.
N= number of exemplars in the dataset.
Yij=network output for exemplars i at processing element j, 
dij=desired output for exemplars i at processing element j, 

Model Performance Test: Fig.3 shows the results of the validation performance. The least MSE in the validation step happened at epoch 1000, which has the best validation performance equal to 1.8728e-09.

![Fig. 4. Best validation performance of the model](image)

3.2 Assessment application

The second component is the question-based application. The backbone of this application is the relational database management system (RDBMS). It has 10 tables. Table 1 shows the attribute of each table and Fig. 4 shows the RDB diagram.
Table 1. The database tables description and attributes.

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student</td>
<td>Saves information related to the student (Name – Username – Password – GPA – Performance - Department).</td>
</tr>
<tr>
<td>Category</td>
<td>Category of the Subject (Biology – Science). One subject has several categories containing questions.</td>
</tr>
<tr>
<td>Doctor</td>
<td>Information about Module Leaders (Name – Username – Password – Department).</td>
</tr>
<tr>
<td>Questions</td>
<td>All saved questions in the system (Question – Answer – Difficulty Rate – Category – Type – Module Leader).</td>
</tr>
<tr>
<td>Exams</td>
<td>Information about the saved exam (Exam name – Date – Subject – Answers).</td>
</tr>
<tr>
<td>Sheets</td>
<td>Information about the saved sheet (Sheet name – Date – Subject – Answers).</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Evaluation Exam (Exam number – Date – Student – Subject).</td>
</tr>
<tr>
<td>Subject</td>
<td>All subjects of system (Name – Department).</td>
</tr>
<tr>
<td>Department</td>
<td>Departments in system (Name). Department □ Business.</td>
</tr>
<tr>
<td>Type</td>
<td>Types of Question Example: True or False - MCQ</td>
</tr>
<tr>
<td>Admin</td>
<td>Information for IT individuals (Username – Password).</td>
</tr>
</tbody>
</table>

Fig. 5. Relation database (RDB) diagram

This database is accessed by two main user types i.e. the module leaders and the new students. A module leader has full access to the database. The main responsibility of
the module leader is to add questions, answers and the difficulty rate to test the new students’ background knowledge in different domains.

The second user is the student. He/she uses the mobile app and answers the questions. Fig. 5 shows an example of the interface.

![Example interface](image)

**Fig. 6.** The Interface of the assessment application

### 3.3 Fuzzy model

The third component is the fuzzy model. There are four main steps to build a fuzzy model namely, crisp inputs, fuzzed inputs, fuzzy conclusion and crisp output. It starts with having a crisp input or fuzzy sets and in our case, the sets are both the neural network and the assessment application results. Second is developing the core model. This step begins by processing the fuzzified inputs or the fuzzification where the fuzzy rules are defined. The defuzzification is the following step to get the fuzzy conclusions and produce the crisp output.
3.4 Fuzzy sets (crisp inputs and outputs) and membership function

A fuzzy set is defined by Zayeh [30][1] to be a class of objects that have a continuum of membership grades. The membership function assigns to each object a membership grade with a range between zero and one. The membership function curve defines how each point in the input space is mapped to a membership value between 0 and 1[2]. The x axis represents the universe of discourse, whereas the y axis represents the degrees of membership in the [0,1] interval. The shape of membership function can be Gaussian, triangular, trapezoidal, sigmoid, S-shape or Z-shape[15].

This model has three sets. They are two input sets, which are the GPA and the average correct answers to the questions, and the third is the output set of the students’ performance. The intervals for the sets or crisp inputs and output are as follows. The GPA graph interval carries values between (0 – 4.0). The intervals of the correct answers are between (0-50). For the output of the students’ performance, the intervals are between 0-10.

To develop the membership function, The Gaussian function is selected because the students’ grades and performance set falls under the normal distribution. The Gaussian function is defined by a central value m and a standard deviation k > 0. The smaller k is, the narrower the “bell” is[30] in equation (3).

\[ \mu_A(x) = e^{-\frac{(x-m)^2}{2k^2}} \]  

(3)

This model membership functions are represented in Gaussian curves in Fig. 1, 2, & 3.
3.5 Fuzzy Rule-Based Model (FRBM)

A Fuzzy Rule-Based Model (FRBM) or a fuzzy inference system estimates a set of outputs from a set of input data using the fuzzy sets theory [30][23]. This model predicts
the students’ performance with two inputs, namely the GPA and the average of correct answers. Matlab’s fuzzy inference system was used to generate the rules. These are the rules for the students’ performance.

- If (GPA is Low) and (Average Correct is Low) then (Rate is Low)
- If (GPA is Low) and (Average Correct is Med) then (Rate is Low)
- If (GPA is Low) and (Average Correct is High) then (Rate is Med)
- If (GPA is Med) and (Average Correct is Low) then (Rate is Low)
- If (GPA is Med) and (Average Correct is Med) then (Rate is Med)
- If (GPA is Med) and (Average Correct is High) then (Rate is High)
- If (GPA is High) and (Average Correct is Low) then (Rate is Med)
- If (GPA is High) and (Average Correct is Med) then (Rate is High)
- If (GPA is High) and (Average Correct is High) then (Rate is High)

Fig. 11. Result of rate when GPA=2.95 and average correct answers=18.4 i.e. 60%

3.6 Defuzzification

A defuzzification is the process to produce the fuzzy conclusion. It is called known as fuzzy inference. This defuzzification process means converting fuzzy output back to classical or crisp output to the control objective. The commonly used method is the Mean of Maximum and the Center Of Gravity (COG) Method[18][24].

COG method can be defined by the algebraic expression used for algebraic integration. Fig. 11 represents this method graphically[24]. The weighted average for center of gravity or membership function of the area bounded by function of membership curve is computed then to be the most crisp value of fuzzy quantity. For discrete membership function, the defuzzified value denoted as X using COG is defined as

\[ X = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} \]

http://www.i-jim.org
indicates the sample element, \( \mu(X_i) \) is the membership function, and \( n \) represents the number of elements in the sample (see equation (4)).

\[
X = \frac{\sum_{i=1}^{n} x_i \mu(x_i)}{\sum_{i=1}^{n} \mu(x_i)}
\]  

(4)

Fig. 12. Defuzzification using Center Of Gravity (COG) [24]

Fig. 10 shows the COG results when the GPA is 2.95 and the Average Correct Answers is 18.4. The fuzzy inference system calculates the overall performance or rate of 6. It is equal to 60%.

4 Conclusions and Future Work

This paper presents new work in the area of predicting student performance.

First, data selection was based on the factors influencing drop out or students at risk percentages. The supervised ANN model used student data of the university from 2009 -2017. From this data the main attributes selected were socio-economic factors that affect their performances [13]. In addition, as research shows that the drop out usually happens after the first year [24] and[27], one of the attributes is the prep year mark. Consequently, the attributes are graduation year, gender, address, country of school, prequalification certificate, prequalification percentage, English entry mark, prep year average and to predict the final GPA.

Second, the system tests the students’ basic knowledge in certain area that is related to his specialty or major they select. It develops a mobile app that has a database. It allows the lecturer or tutor to add questions and the system calculates the percentage of correct answers.

Third, based on the real data provided by the university, the system incorporates a neural network that predicts the final GPA of the students. At that point this work is
similar to [7], [13], and [11]. The difference is in the attributes used to predict the performance.

Fourth, the value of this work is the fuzzy model that integrates the students’ current knowledge in the major in which the student wishes to specialize and the predicted final GPA. The model predicts the final performance of students. The major is one of the factors that affects the final performance. This is different to the work of [2], [12], [21], [14] and [6].

Future work will exam whether the system can pick up the exact cause behind low performance. Additionally, we would like to expand this work to add a question-based system that predicts the performance and guides students to improve their performance.

5 References


http://www.i-jim.org


6 Authors

Ann Nosseir is an associate professor at the Institute of National Planning (INP) and the British University in Egypt (BUE). She is the Faculty of Informatics and Computer Science.

Yahia Mohamed Fathy is a Faculty of Informatics and Computer Science (ICS) at British University in Egypt. Yahia is Informatics & Computer Science department graduate in 2019. yahia129151@bue.edu.eg