Applying Deep Learning Neural Networks in Predicting Students' Cumulative Grades

Elmasry, Mohamed Abbas

Abstract: *The distinguished universities aim to provide quality education to their students. One way to achieve the highest quality in university studies is to discover knowledge to predict student performance and grades in courses etc. Recently, the amount of data stored in educational databases is accumulating very quickly, as these databases contain indirect information that can be used to improve student performance. Academic performance is affected by many factors, so it is necessary to predict student performance to determine the difference between students who are excelling in studies and students who need to exert more effort to improve their performance and their level of achievement. Hidden or Indirect knowledge is part of the educational data set and can be extracted using various means, such as data mining techniques and the use of classification, and deep learning through neural networks. This paper has been designed to extract knowledge describing students' performance in the courses required for graduation, in a way that helps academic advisors in providing academic advice and guidance to students to improve their cumulative grades.*

Keywords: *Deer Learning, Neural Networks, Classification, Data Mining, Academic Performance, CGPA.*

I. INTRODUCTION

Predicting student performance is very important on the academic level. There are a number of methods that can be used to predict student grades, the most prominent of which is data mining. Data mining models can be used to run on a large amount of data to discover the different relationships of importance in guiding students towards improving their academic performance. Classification is one of the most popular data mining methods used in educational datasets, and classification is a predictive method used in data extraction, in which data values are predicted using known results found through processing of students' results data [2]. The ability to predict strong academic performance of students helps to discover students with underachievement early and then take measures to help improve their level of achievement, and on the other hand, to identify students who are expected to excel based on currently available data on their results, which helps to enhance their competence and abilities, Thus achieving the highest score. The findings of this prediction are meant to be available to the academic staff and faculty members, to use in guiding students towards improving their grades. Data mining reveals useful information from large sets of data, where multiple algorithms are applied with the aim of determining the patterns of the data being processed [24]. As the mining of data and discovering the relationships and patterns associated with them is of great importance in making various decisions, it has been met with great interest in the scientific research community. As the means of data mining have been used in various fields, the mining of educational data has gained increasing importance in recent times. It relies on the use of methods that discover knowledge represented in the relationships and patterns extracted from educational data [J. Han and M. Kamber]. It is possible to use a number of data mining methods with educational data such as classification, the use of a decision tree, the use of deep learning based on neural networks, and many other methods. Using data mining techniques, it is possible to discover abnormal values in student results, predict student grades, etc. This paper aims to use data mining methodologies to study student performance on courses. Data mining provides several tasks that can be used to study student performance. There are many methods used to classify the data. In this paper, the decision tree method and deep learning using neural networks are used to predict the academic performance of the student. Detailed information on student results has been collected, including course work, mid-term exam results, final exams results, and data processed to predict students' cumulative grades. This paper is being applied to students of the Faculty of Business Administration at the National Egyptian E-Learning University, where the final grades of the students depend on the final exams by 50%, the mid-term exams by 20%, the coursework by 30% distributed between assignments, quizzes and participation. The students' cumulative grade is calculated According to the relative weight of the courses based on the number of credit hours for each course. It is expected that the results of predicting students' grades through the means of data mining used to provide a successful way to reform the educational path of students, and to direct faculty members towards improving the effectiveness of academic performance, and it should be noted that obtaining student results without processing them through the means of data mining for the purpose of discovery Patterns and relationships will be useless and will not allow any real conclusion to be used for prediction purposes. This paper examines the accuracy of decision tree techniques and deep learning using neural networks with the aim of predicting students'

Elmasry*, National Egyptian E-Leaning University (EELU), Giza, Egypt. E-mail.id malmasry@eelu.edu.eg
Mohamed Abbas, National Egyptian E-Leaning University (EELU), Giza, Egypt. E-mail.id malmasry@eelu.edu.eg

Retrieval Number: 100.1/ijeat.B20841210220
DOI:10.35940/ijeat.B2084.1210220
Journal Website: www.ijeat.org

293

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication
© Copyright: All rights reserved.

cumulative grades.

II. RELATED WORKS

Data mining in undergraduate education is a promising field of research due to its enormous potential that can achieve tangible results that contribute to benefiting from the accumulated data of university students' results.

Referring to the literature, it has been found that some studies have used educational data mining (EDM) through a student's GPA to predict their final grades or to predict the high probability of failure.

Asif et al. [5] prepared a study to predict the performance of students at the undergraduate level, as the study conducted on more than two hundred university students concluded that the grades of students in secondary education (pre-university) and their grades in the courses at the first and second level play a major role in the accuracy of the prediction of their cumulative grade point average upon graduation.

Ibrahim, Z., & Rusli, [11] found that the most important way to predict the academic achievement of students is the GPA, as he found that there is a correlation rate of no less than 87% to predict student success using the GPA compared to other data used in prediction.

Kabakchieva [15] analyzed the data of more than ten thousand university students, whose academic grades varied significantly, and by using the educational data mining, it was found that the number of times of students' failure in first-level courses has a prominent role in the accuracy of predicting the final results of the students.

As for Alija [1], he studied the relationship between attendance and students' final results using regression analysis, and the study concluded that there is a positive relationship between attendance in classes and achieving success.

Nahid Hajizadeh, Marzieh Ahmadzadeh [19] In this research, a variety of analysis were made using educational data mining, one of these analysis included the courses that are presented by more than one professor separately, where the results of the students enrolled in the same course varied according to the professor responsible for the course, the results of the analysis Indicates that if a certain professor teaches that course, it is possible to predict that most students will successfully pass the course in the first attempt.

Lukkarinen et al [18] studied the relationship between attendance in and academic achievement, as the results of students were analyzed in one of the courses in which attendance was not required, and the study concluded that attendance has a major role in evaluating students' academic performance, and it has a close relationship with accurately predicting students 'grades. nAs for Aulck et al. [4] they found that scores earned on quizzes and classwork significantly affect the predictive accuracy of students' final grades. For Huang Wu Fang [10] it was found that the score students obtain on the mid-term examination is the most important factor affecting the accuracy of predicting final results. As for Pradeep, A., & Thomas, J. [22] they found that the results of students at the first level play a major role in predicting the high probability of failure, as it represents an early warning bell that should be taken seriously by academic officials.

III. DATASET

We obtained data from the results of 340 graduate students from the Faculty of Business Administration at the National Egyptian E-Learning University (EELU) for the period from 2012 to 2018. These data were collected through the university's management information system (UMS), where the results of the students officially stored. We encountered some difficulties in the data collection phase and preparing student results for processing through data mining. Among the most prominent of these difficulties are the following:

• Data of students transferred from other universities, where some courses were transferred to them with the grade they obtained after passing successfully, without information on the number of previous attempts until achieving success, and without information about any other courses they had previously studied but were unable to obtain the equivalence because EELU academic charter requires at least a C grade to transfer any course.

• There is a deficiency in the results of some students, as the results do not include all the courses required to be successfully completed for technical reasons related to the databases and how the results are presented.

• Students transcripts were obtained for the purpose of verifying that all students have successfully completed the 52 courses required for graduation and grades related to each course.

IV. METHODOLOGY

Classification is an easy-to-use tool for extracting data, it is very useful for decision-making purposes.

The classification process goes through two stages:

• The first stage: developing a training model.

• The second stage: evaluation of the model using the dataset. There are many algorithms that can be used in dataset classification, some of which can be illustrated by the following figure:
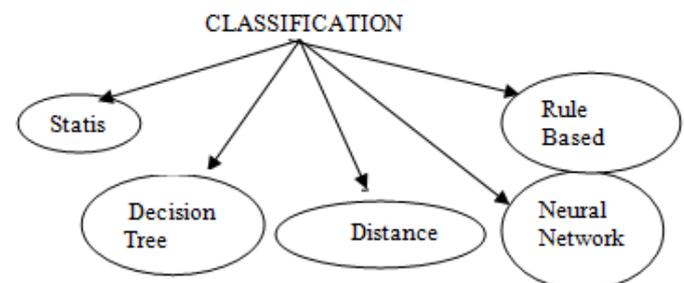

Figure 1. Different Classification Algorithms

In this paper, we investigated two classifiers: decision trees and deep neural networks. Decision tree is a learning technique mostly used for classification purposes. It is a classifier organized in the form of a tree, where the internal nodes contain the properties of the dataset, while the branches represent the decision rules, in this model, each leaf of these nodes represents the result reached. In the decision tree,

the decision node is used for decision-making purposes, and it has many branches, while the leaf nodes are considered the output of decisions and do not include any branches otherwise. The second investigated classifier in this model is: deep neural networks. Neural network with a certain level of complexity consisting of two layers or more is considered to qualify as a deep neural network. Deep networks process data in intricate ways by using complex mathematical modeling.

The learning part of creating models led to the development of artificial neural networks. Artificial neural networks use a hidden layer as a storage space and evaluate how inputs are of significance to outputs. The hidden layer stores the information related to the importance of the inputs.

The one hidden layer is an artificial neural network. Deep neural networks take advantage of the Artificial Neural Network component, because every node in the hidden layer makes both correlations and scores significance of the input to determine the output.

We used the **RapidMiner**® tool for all implementation steps. RapidMiner is considered as a software platform for data science, it is characterized by providing a suitable environment for data preparation, machine learning, deep learning, accurate predictions and so on.

In this model, we have compared the performance of these two classifiers with the same dataset.

V. PRACTICAL DATA MINING

Final Grade Prediction

In this task, we are trying to benefit from the historical records of the students in the Faculty of Business at EELU to build a supervised machine learning model for predicting the final grade for the student based on his/her performance in all the taken courses.

For more details, we collected the students' grades in all the courses that they passed in all the semester in addition to two extra features which are Minor (the student's graduation minor) and ChangeMinor (flag indicating if the student changes his/her minor before graduation or not). The target variable for this task is the cumulative grade at graduation with four possible values:

- Pass
- Good
- Very Good
- Excellent.

In total the collected dataset contained about 340 records for all the graduated students with 69 features as described earlier.

In the literature, there are a huge list of classification algorithms and methods that were used in similar tasks. In our model, we investigated two classifiers: Decision Trees and Deep Leaning Neural Networks. For the implementation, we used **RapidMiner**® tool for all the implementation steps two compare the performance of these two classifiers with the same dataset as shown in Figure 2.

The workflow started with loading the dataset of all the graduate business students and replicating it to be used in both classifiers. All the features in the dataset contain nominal values (i.e. have categorical values), which need to be preprocessed as the deep learning neural networks can work only on numerical data. This requires passing the dataset into the "One-Hot Encoding" operator to remove all the

categorical attributes and convert them into numerical ones based on the number of categories in each attribute. This step is not needed for the Decision Tree classifier as it can handle such values correctly.
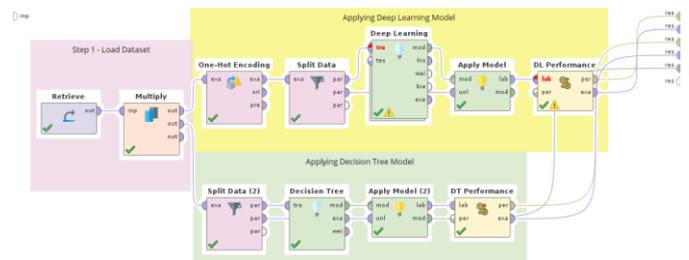

Figure 2. Final Grade Prediction Model Implementation in RapidMiner

The next step is to split the dataset into two parts: training and testing parts. The purpose of this step is to train the model using the training part which holds about 80% of the records and hold back 20% of the records to measure the model's performance on a totally unseen data. This ensures the generalization of the model and its ability to work on the newly added datasets. This step is performed for both classifiers as shown in Figure 2.

For the "Deep Learning" operator, it contains 1 hidden layer with 10 neurons and applied ReLU activation function with dropout rate of 20%. Followed by the output layer with 4 outputs (one for each target label value) and applied Softmax activation function to normalize the outputs of the deep learning model. On the other hand, we used the "Decision Tree" operator for training the decision tree classifier with information gain as split criteria, maximum depth of 15 and confidence threshold value of 0.2.

The trained models were applied on the testing part to generate the predicted outputs of both models. There predictions were passed to the "Performance" operator to measure the performance of these classifiers on the unseen testing records. The deep learning model overcomes the decision tree with a very huge margin achieving accuracy of **91.04%**, while the decision tree only obtained **70.15%** accuracy.

For a deeper look on the performance, we calculated the confusion matrix to compare the true and predicted labels by both classifiers. This will give us another perspective about the model's performance for each target label value as summarized in **Error! Reference source not found.**.

Table- I: Confusion Matrix –Final Grade Prediction

True \ Predicted	Deep Leaning Model				Decision Tree Model			
	Pass	Good	Very Good	Excellent	Pass	Good	Very Good	Excellent
Pass	19	1	0	0	13	4	0	0
Good	0	23	1	0	6	18	2	0
Very Good	0	1	12	1	0	3	9	1
Excellent	0	0	2	7	0	0	4	7
Class Recall	100%	92%	80%	87.5%	68.42%	72%	60%	87.5%

The trained deep learning model have a high accuracy of 91.04%. This was reflected on the confusion matrix shown in **Error! Reference source not found.**.

The deep learning model can predict the correct final grade labels with 100%, 92%, 80% and 87.5% for the Pass, Good, Very Good and Excellent grades respectively. The lowest class recall was for the "Very Good" grade because the model misclassified 3 records out of the total 15 records in this class. In addition, the model was able to correctly predict all the records for the "Pass" grade with 100% class recall. We tried to tune the decision tree parameters with many values, but the best recorded performance has accuracy of 70.15% only. As shown in **Error! Reference source not found.**. The model has weak performance in the "Pass", "Good", and "Very Good" labels with class recall of 68.42%, 72% and 60% respectively. The model predictions for the "Excellent" grade has the highest class recall of 87.5% which matches the same predictions of the deep learning model. In general, we were able to build a supervised machine learning model using the deep learning neural networks classifier to correctly predict the students' final grade based on their performance on the courses. This model has a high accuracy of 91.04% which can help the academic advisors in the faculty to guide the students to improve their performance to achieve better final grades. With more new data coming from the graduated students every year, the deep learning model can be retrained to achieve higher accuracy.

VI. CONCLUSION AND FUTURE WORK

One of the important areas of benefit from information technology is educational data mining. Predicting students grades is of multiple benefits and uses, as it can be used by academic officials to improve students' achievement level, and it can also be used administratively in predicting outstanding students who are eligible for discounts or scholarships according to the university's financial regulations. In this study, we extracted data from a database of students of the Faculty of Business Administration at the National Egyptian E-learning University and processed the dataset using decision tree and deep neural networks algorithm in order to predict the most accurate results for the students cumulative grades. Through our use of the Deep Learning Neural Network Classifier, we were able to build a supervised machine learning model that is able to accurately predict final grades for students based on their previous course performance. Through this paper we attempted to adjust the decision tree parameters using several values, with the best recorded performance having an accuracy of only 70.15%. The model also showed poor performance in the ratings of "success", "good" and "very good" with category calls of 68.42%, 72% and 60%, respectively. The model predictions of an "excellent" score have the highest class recall of 87.5% which matches the same predictions of the deep learning model. The use of a deep learning model led to more accurate results, the accuracy of this model reached 91.04%, which could help the academic officers to guide students to improve their academic performance and achieve better final grades. We hope that the results obtained through this research paper will be useful in improving the academic performance of students, and we also hope that it will contribute to improving the academic advising service that students receive. This work may guide students to improve their final grades; Reducing the dropout rate from university education; This is done by making timely directive decisions to improve students' achievement at the university level and improve the quality of education. For future work, we hope that more classifier algorithms will be tested in the field of educational data mining to achieve the best possible results.

REFERENCES

1. Alija, S. (2013), "How attendance affects the general success of the student", International Journal of Academic Research in Business and Social Sciences, 3 (1), 168.
2. AI-Radaideh,Q. A., AI-Shawakfa, E.M., and AI-Najjar, M. I, "Mining Student Data using Decision Trees", International Arab Conference on Information Technology(ACIT'2006), Yarmouk University, Jordan, 2006.
3. A.Mohamed Shahiria,, W. Husaina , N. Abdul Rashida, (2015). "A Review on Predicting Student's Performance using Data Mining Techniques" Procedia Computer Science 72 ,414 – 420.
4. Aulck, L., Velagapudi, N., Blumenstock, J., & West, J. (2017). "Predicting student dropout in higher education". Retrieved 12 January, 2019 from https://arxiv.org/pdf/1606.06364.pdf.
5. Asif, R., Merceron, A., Abbas Ali, S., & Ghani Haider, N. E. D. (2017). Analyzing undergraduate students' performance using educational data mining. Computers & Education, 113, 177–190.
6. Badr, G., Algobail, A., Almutairi, H., & Almutery, M. (2016). "Predicting students' performance in university courses: A Case Study And Tool in KSU mathematics department". Procedia Computer Science, 80–85.
7. Bhutto, S., Siddiqui, I. F., Arain, Q. A., & Anwar, M. (2020). "Predicting students' academic performance through supervised machine learning". ICISCT 2020: 2nd international conference on information science and communication technology. Institute of Electrical & Electronics Engineers.
8. Dutt, A., Ismail, M. A., & Herawan, T. (2017). "A systematic review on educational data mining". IEEE Access. Institute of Electrical & Electronics Engineers.
9. Goni, U., YaganaWali, S., & Waziri, B. M., Kaltum, A. H. (2015). "Gender difference in students' academic performance in colleges of education in Borno State", Nigeria: Implications for counselling. Journal of Education and Practice, 6 (32), 109-112
10. Huang, S., & Fang, N. (2013). "Predicting student academic performance in an engineering dynamics course: A comparison of four types of predictive mathematical models", Computers & Education, Vol 61, Pages 133-145
11. Ibrahim, Z., & Rusli, D. (2007). Predicting students' academic performance: Comparing artificial neural network, decision tree and linear regression. In 21st Annual SAS Malaysia forum.
12. I. Milos, S. Petar, V. Mladen and A. Wejdan, (2016). " Students' success prediction using Weka tool", INFOTEH-JAHORINA Vol. 15 ..
13. Jain, A., Arora, A., Rawat, A., & Dhami, N. (2017). "Analysis of various decision tree algorithms for classification in data mining". International Journal of Computer Applications, 163 (8).
14. J. Han and M. Kamber, "Data Mining: Concepts and Techniques," Morgan Kaufmann, 2000.
15. Kabakchieva, D. (2013). "Predicting student performance by using data mining methods for classification. Cybernetics and Information Technologies", 13 (1), 61–70.
16. Kemper, L. (2018). "Predicting student dropout: A machine learning approach". Karlsruhe Institute of Technology, Internal Report.
17. Kumar, M., Handa, D, & Singh, A. J. (2017). "Literature survey on student's performance prediction in education using data mining techniques". International Journal of Education & Management Engineering, 7 (6), 42–48.
18. Lukkarinen, A., Koivukangas, P., & Seppälä, T. (2016). "Relationship between class attendance and student performance. Procedia: Social and Behavioral Sciences", 228 (July), 341–346
19. Nahid Hajizadeh, Marzieh Ahmadzadeh, (2014),"Analysis of factors that affect students' academic performance - Data Mining Approach", International Journal of advanced studies in Computer Science and Engineering IJASCSE, Volume 3, Issue 8.

20. Nabizadeh, S., Hajian, S., Sheikhan, Z., & Rafiei, F. (2019). "Prediction of academic achievement based on learning strategies and outcome expectations among medical students". BMC Medical Education, 19 (1).
21. Nithya, P., Umamaheswari, B., & Umadevi, A. (2016). "A survey on educational data mining in field of education". International Journal of Advanced Research in Computer Engineering and Technology (IJARCET), 5 (1), 1–12.
22. Pradeep, A., & Thomas, J. (2015). "Predicting college students dropout using EDM techniques". International Journal of Computer Applications, 123 (5), 27-30.
23. Slater, S., Joksimović, S., Baker, R. S., Kovanovic, V., & Gasevic, D. (2017). "Tools for educational data mining: A review". Journal of Educational and Behavioral Statistics, 42 (1), 88–105.
24. U. Fayadd, Piatesky, G. Shapiro, and P. Smyth, (1996), "From data mining to knowledge discovery in databases", AAAI Press / The MIT Press, Massachusetts Institute Of Technology. ISBN 0–262 56097–6.

AUTHOR PROFILE

Author Dr. **Mohamad Elmasry** is the Director of Business Administration program at the Egyptian E-Learning University (EELU). He is responsible for the operational management and implementation of the academic aspects of the program He is specialized in Accounting and Information Systems. He got a Ph.D. in Accounting Information Systems from Alazhar University, Cairo, Egypt. Moreover, he is a member of scientific associations including the American Association of E-commerce (AAE), USA, the institute of E-commerce consultants (Certified E-Commerce Consultant ™), USA, the Egyptian Association for Educational Technology (EAET) and the Egyptian Association for curricula and teaching methods (ECCI).