Journal of the ACS, Vol. 7, May 2013

A comparative analysis of Techniques for Predicting Academic Performance

Dr. Mohamed El Zeweidy,

El-Shorouk Academy,
Higher Institute for Computers and Information Technology,
Computer Dept.,
El-Shorouk Academy, Cairo, Egypt,
E-mail: melzeweidy1@gmail.com

Dr. Emad Osman,

El-Shorouk Academy,
Higher Institute for Computers and Information Technology,
Computer Dept.,
El-Shorouk Academy, Cairo, Egypt,
E-mail: emad91@hotmail.com

Ass. Prof. Magdy E. Elhennawy
Higher Institute for Computers and Information Technology,
Computer Dept.,
El-Shorouk Academy,
Family Card Project Consultant,
Ministry of State for Administrative Development,
Cairo, Egypt,

Abstract
The main objective of the admission system is to determine candidates who would likely do well in the university. The quality of candidates admitted into any higher institution affects the level of research and training within the institution, and by extension, has an overall effect on the development of the country itself, as these candidates eventually become key players in the affairs of the country in all sectors of the economy.
This article compares the accuracy of various data mining techniques, namely: decision trees, logistic regression, neural network, naive bayes, association rules and clustering for predicting the academic performance of the first semester for the undergraduate engineering students at the Modern Academy for Engineering (MAE) by using the high school grade as the only input, and proposes a method that allows best prediction results from different prediction algorithms to be selected. A set of data has been tried to proof the correctness of the proposed method. According to the obtained results, the data-mining tools were able to achieve levels of accuracy for predicting student performance. The results showed that decision trees, clustering, and naive bayes score was a little more than the other three for the sets {pass, fail} and {excellent, very good, good, pass, fail, very bad, absent} while association rules, came out the last with the least score for both sets.
The results of these case studies give insight into techniques for accurately predicting student performance and compare the accuracy of data mining algorithms.

Keywords: predicting the academic performance, Decision Tree, admission system

1. Introduction

Accurately predicting student performance is useful in many different contexts in universities. For example, identifying exceptional students for scholarships is an essential part of the admissions process in undergraduate and postgraduate institutions, and identifying weak students who are likely to fail is also important for allocating limited tutoring resources as well as strategic programs can be planned in improving or maintaining and assisting students' performance during their period of studies in the institutions.

The remainder of this paper is structured as follows. In section 2 the related work has been surveyed. In section 3 the problem has been defined. In section 4 the basic framework of the used model and selection of the needed data has been presented. In section 5 the related results has been analyzed. The analysis covers the predictive modeling using decision trees, then comparing the results of various algorithms then adding more input attributes to the mining model. In section 6 the prediction for a high school grade for a given faculty success level is described. In section 8 and 9 the conclusion and future work is stated.

2. Related work

Since institutes all over the world wants to be sure they are selecting the cream of the crop, many have tried to work on ways for predicting academic performance for their applicants or students. One of those was the artificial neural networks ANN, which was used to predict the cumulative Grade Point Averages (CGPA) by using ten inputs including: UME score, O level results in mathematics, English language, physics, and chemistry, age of student at admission, time that has elapsed between graduating from secondary school and gaining university admission, parents educational status, zonal location of student's secondary school, type of secondary school attended (privately owned, state or federal government owned), location of university and place of residence, and student's gender [**3**].

Other study compares the accuracy of decision tree and Bayesian network algorithms for predicting the academic performance of undergraduate and postgraduate students at two very different academic institutes [**4**]. They used admissions information, such as academic institute and GPA to predict

GPA at the end of the first year. The data-mining tools were able to achieve similar levels of accuracy for predicting student performance: 73/71% for {fail, fair, good, very good} and 94/93% for {fail, pass} at the two institutes respectively. In 2008, using neural network the CGPA was predicted by the students' demographic profile and the CGPA of the first semester [**5**]. The study compared the accuracy of three predictive models which were artificial neural networks, decision trees and linear regression, and showed that the artificial neural network outperformed the other two with accuracy more than 80%.

Another study examines the relationship between students' overall academic performance (GPA) and matriculation requirements performance in first year courses in the Bachelor of Science and Information Technology (BSCIT) program at UTECH [**7**]. Other researches tried to find if the performance is affected by age, gender, Caribbean Examination Council (CXC) qualification, aptitude test score and experience [**8**].

Another good study showed different ways in which student performance statistics can be used to obtain information which may be used in assessing the individual student, course, program and the department in terms of their performances [**12**]. A number of data warehousing and data mining concepts were applied to obtaining the required results, then the same researchers took it another step further in which, the different ways in which student performance data can be analyzed and presented for academic decision-making are investigated and a software package called the Performance-based Academic Decision-Support System (PADSS) is developed [**11**].

On the other hand many others did not use any artificial intelligence for the prediction but used simple statistics depending on other variables, like [**1**] who used the admission test with the gender, [**2**] used the SAT with all its divisions like writing, verbal, math etc. [**6**] used Graduate Management Admission Test (GMAT) and Undergraduate Grade Point Average (UGPA) for predicting Graduate Student Academic Performance, and [**9**] used Miller Analogies Test to predict the GPA. One interesting result was achieved by [**10**] which found out that SAT predicts performance for male students but not for female students, while Prior related courses did not predict the academic performance.

The Modern Academy for Engineering MAE has had thousands of applicants per year over the last few years. Of this number approximately one thousand is accepted. The academy offers a Bachelor of Engineering in many majors like computer, mechanics, civil, architect …etc.

And since the selection of students solely depend on the high school grade, this study tries to find out how much does the high school grade is suitable or related alone by itself to the academic performance of the first semester at the academy by comparing the results of six algorithms which are: decision trees, logistic regression, neural network, naive Bayes, association rules and clustering algorithms. In the following section the overall methodology of the proposed method will be described. Next, the results of the prediction algorithms will be compared and finally, the conclusions.

3. Problem definition

The research tries to discover the relationship between high school grades from one side and the success level of the student for each subject in a particular faculty on the other side. The objective is to build a data mining system that provides the administration with the information needed to help the students who need academic assistance. Moreover the system can help undergraduate students to choose the best field or branch of study that suits their skills and abilities depending on their high school grade. This can be achieved by the proposed method, which includes three steps.

First, we will use the high school grade to try predicting, the first semester total grade and the grade of each subject by itself, and then these results will be compared to determine which is more likely to be predicted. This step will be done using the Decision Trees algorithm for predicting the academic performance of the first semester for the undergraduate engineering students at the Modern Academy for Engineering (MAE) by using the high school grade as the only input.

After determining which is more likely to be predicted either the total grade or the subjects grades, a comparison between different mining algorithms namely Decision Trees, Logistic Regression, Neural Network, Naive Bayes, Association Rules and Clustering to measure the accuracy of the algorithms for predicting the academic performance. The comparison will be held over two sets of data, the first set is {pass, fail} while the other set is {excellent, very good, good, pass, fail, very bad, absent}.

Finally, another input (the high school type) is added to the model, in addition to the high school grade to check its effect on the performance and comparison results.

This method establishes a pattern that can give insight into techniques for accurately predicting student performance and compare the accuracy of data mining algorithms. Figure 1 shows the hierarchy of the proposed method.

4. Model building and selection

The data obtained for this study were collected for students admitted in 2006/2007 and 2008/2009 and contains 2638 students' data with their grades in all 16 subjects and their high school grades and types which are 12 types. From which, 1537 students' data for the study were obtained after data cleansing process. **Figure 2** shows a sample of the original data collected from MAE. **Figure 3** and **Figure 4** show a distribution of some of the data that should be predicted; the student's actual grades at the end of the 1st semester of undergraduate at MAE. The figures represent the classes {excellent, very good, good, pass, fail, very bad, absent} for the "Physics" & "Math" subjects.

High school types are shown in **Table 1**. High school grades and subjects grades classes are shown in **Tables 2**,and **Tables 3**. Each semester contains eight subjects as shown in **Table 4**. Actual distribution of first semester grades are shown in **Table 5**. Grades Distribution for the {Excellent, Very Good, Good, Pass, Fail, Very Bad, Absent} set grouped by grade and subject are shown in **Figure 5**.

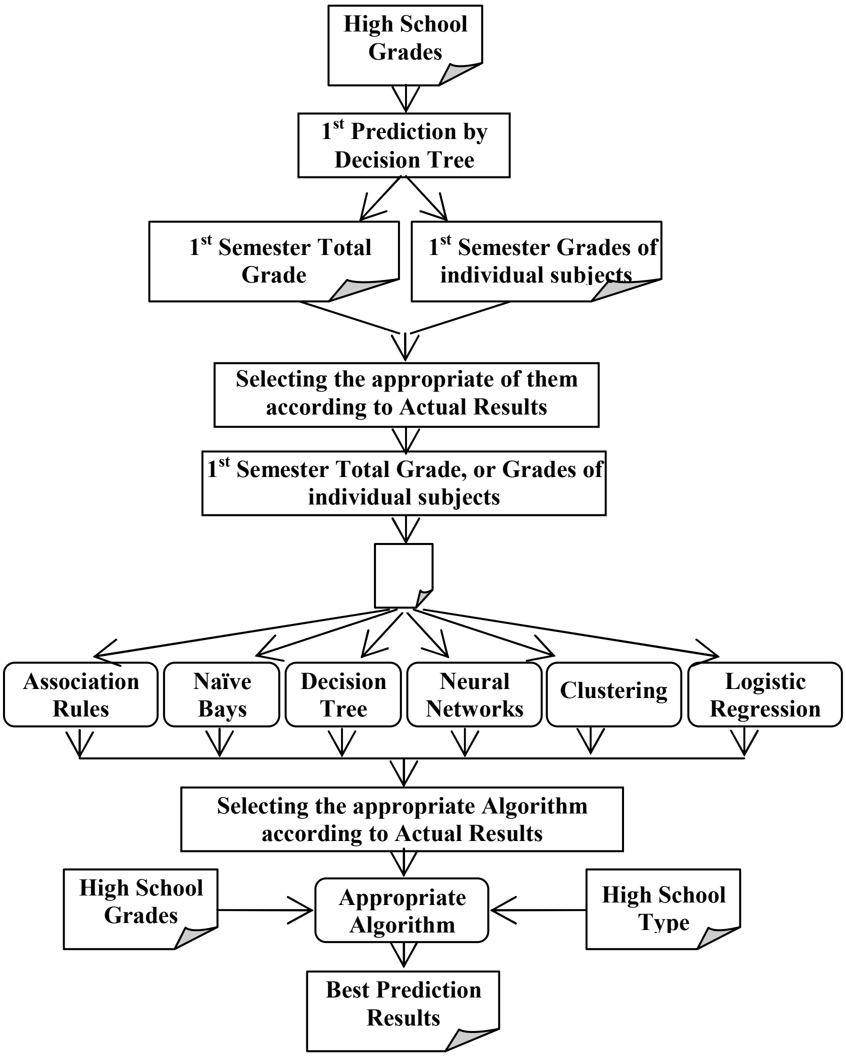


Figure 1: The hierarchy of the proposed method

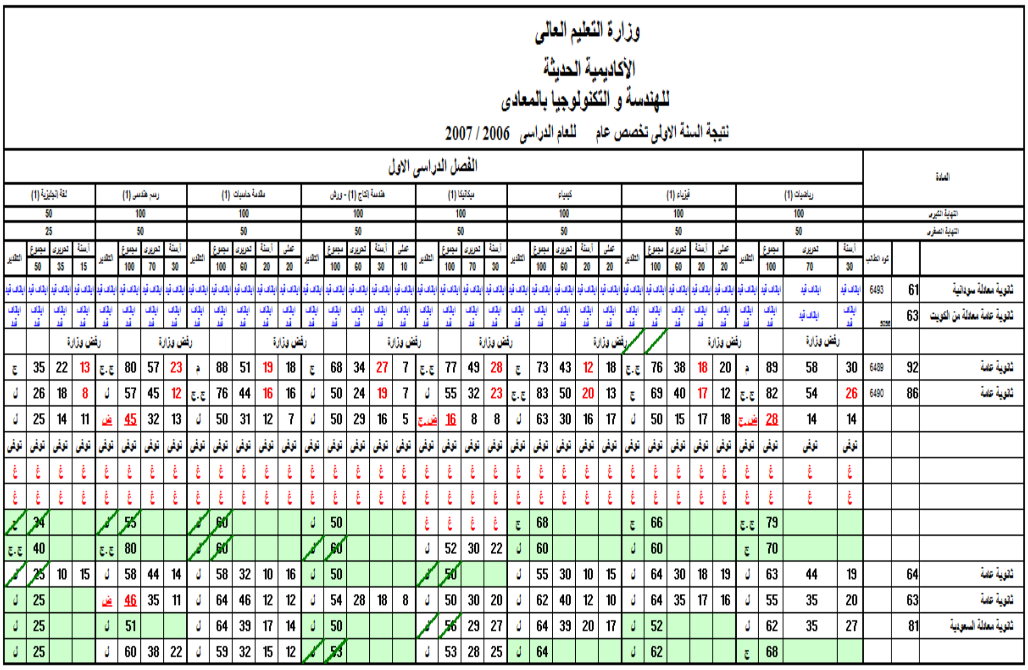


Figure 2: Sample of the Original Data Collected From MAE

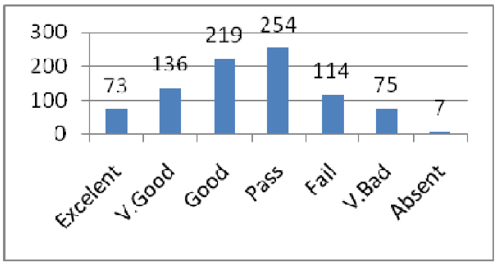


Figure 3: Distribution of actual grades for the “Physics” Subject

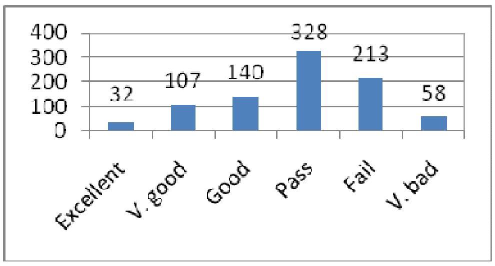


Figure 4: Distribution of actual grades for the “Math” Subject

Table 1: The High School Types

Azhar High School	ثانوي أزهري
General High School	ثانوية عامة
Iraq High School	ثانوية معادلة العراق
UAE High School	ثانوية معادلة الإمارات
KSA High School	ثانوية معادلة السعودية
Oman High School	ثانوية معادلة عمان
Palestine High School	ثانوية معادلة فلسطين
Qatar High School	ثانوية معادلة قطر
Five Years Industrial	دبلوم فني صناعي ٥ سنوات
Sudan High School	ثانوية معادلة السودان
Kuwait High School	ثانوية معادلة الكويت
American Diploma	دبلومه أمريكية

Table 2: High School Grades

From	To	Grade
90+		1
80	<90	2
65	<80	3
50	<65	4

Table 3: Subjects Grades

A	م	1
B	ج ج	2
C	ج	3
D	ل	4
F	ض	5
V. Bad	ض ج	6
Absent	غ	7

Table 4: Subjects Names

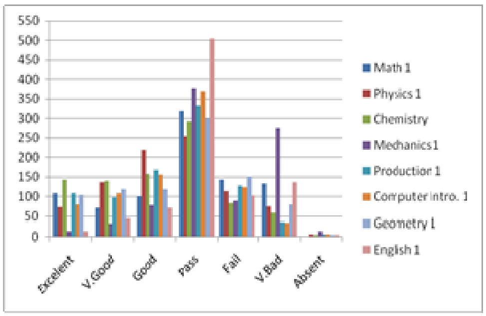
(Subjects Names) أسماء المواد				
First Semester	Math (1)	رياضيات (١)	1	الترم الأول
	Physics (1)	فيزياء (١)	2	
	Chemistry	كيمياء	3	
	Mechanics (1)	ميكانيكا (١)	4	
	Production (1)	هندسة إنتاج (١) - ورش	5	
	Intro. to computers (1)	مقدمة حاسبات (١)	6	
	Geometry (1)	رسم هندسى (١)	7	
	English (1)	لغة إنجليزية (١)	8	

Table 5: Actual Distribution of First Semester Grades

Subject - Grade	Ex.	V.G.	G	P.	F.	V.B.	A
Math							
Physics							
Chemistry							
Mechanics							
Production							
Comp. Intro							
Geometry							
English							

Ex.	Excellent
V.G.	Very Good
G.	Good
P.	Pass
F.	Fail
V.B.	Very Bad
A.	Absent

Grade Legend



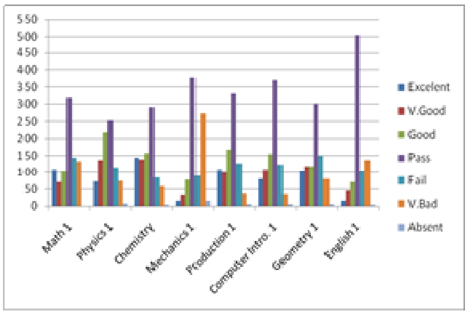


Figure 5: Grades Distribution for the {Ex., V.G., G., P., F., V.B., A.} Set
(A) Grouped by Grade (B) Grouped by Subject

5. Results and Analysis

5.1 Predictive modeling using Decision Trees

First it was essential to determine which is more likely to be predicted from the high school grade, would it be the total grade or the grade of each subject by itself. Decision Trees algorithm was selected for the task of predicting the academic performance of the first semester for the undergraduate engineering students at the MAE by using the high school grade as the only input.

After creating the data mining model, the next step was to train the model. Training is usually the most time-consuming step. The algorithm may iterate over the training dataset a few times to find the hidden patterns. The training

process was done automatically by the used engine which splits the data for training and testing.
After the model is trained, it can be used to do predictions on new datasets. The accuracy score of the prediction using the decision tree algorithm is represented in **Figure 6** and **Figure 7** as for the MATH subject and for the total grade. Table 4 shows that average is 82% while the total grade is 80%.

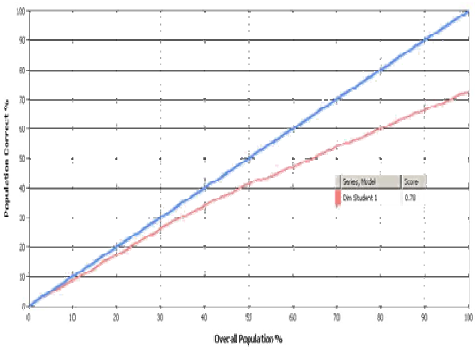


Figure 6: Accuracy Score of "MATH" for the {Pass, Fail} Set Using Decision Trees Algorithm

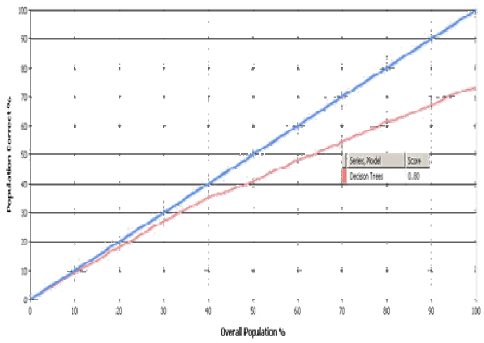


Figure 7: Accuracy Score of "Total Grade" for The {Pass, Fail} Set Using Decision Trees Algorithm

The data-mining tools were able to achieve levels of accuracy for predicting student performance using the selected algorithm is shown in **Tables 6,** and **Table 7**.

Table 6: Decision Trees Accuracy Score For The {Pass, Fail } Set

Math.	Phys.	Chemis.	Mec.	Prod.	Geom.	Comp. Info.	English	Average	Total Grade
78%	85%	89%	72%	85%	87%	77%	81%	82%	80%

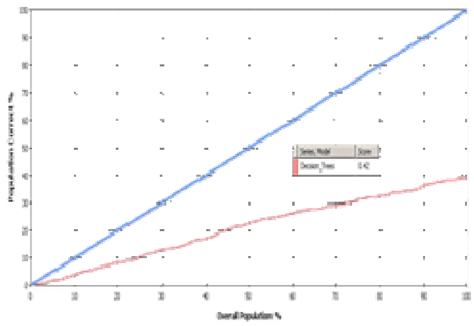


Figure 5: Accuracy Score of "Math" subject for The {Excellent, Very Good, Good, Pass, Fail, Very Bad, Absent} Set Using Decision Trees Algorithm

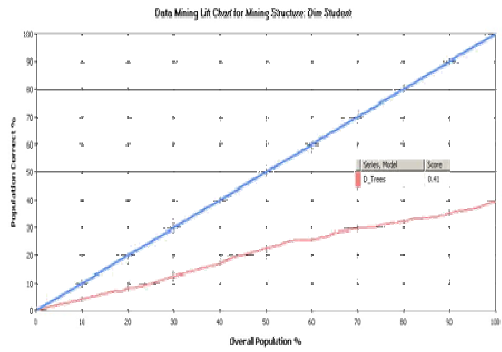


Figure 6: Accuracy Score Of "Total Grade" for The { Excellent, Very Good, Good, Pass, Fail, Very Bad, Absent} Set Using Decision Trees Algorithm

Table 1: Accuracy Score For the {Excellent, Very Good, Good, Pass, Fail, Very Bad, Absent} Set Using Decision Trees Algorithm

Math	Physics	Chemistry	Mechanics	Production	Computer Intro.	Geometry	English	Average	Total Grade
41%	34%	39%	50%	42%	46%	35%	62%	44%	42%

The results show that the predictions for the {pass, fail} set are noticeably more accurate than for the {excellent, very good, good, pass, fail, very bad, absent} set, which is expected, given the much larger number of grades to be predicted. The results also show that the selected algorithm scored an average of 82% for all subjects for the {pass, fail} set, while for the other set the average score was 44%, as shown in **Tables 6** and **Tables 7**.

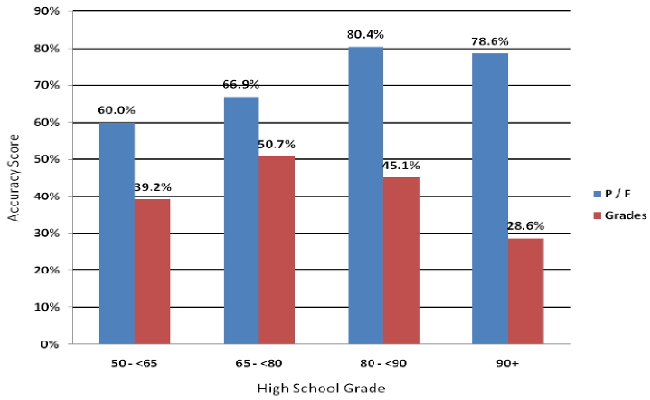


Figure 7: Accuracy Score of "Total Grade" for both Sets Using Decision Trees Algorithm Grouped by High School Grade

5.2 Comparing results of algorithms

This section compares the accuracy of the algorithms used for predicting the academic performance of the first semester for the undergraduate engineering students at the Modern Academy for Engineering (MAE) by using the high school grade as the only input namely; Decision Trees, Logistic Regression, Neural Network, Naive Bayes, Association Rules and Clustering. The accuracy of the results of the MAE predictions using the selected algorithms are shown in the following figures.

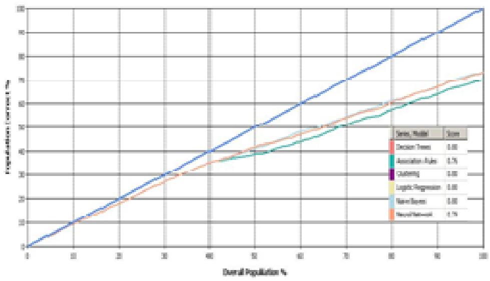


Figure 8: Accuracy Score For The {Pass, Fail} Set

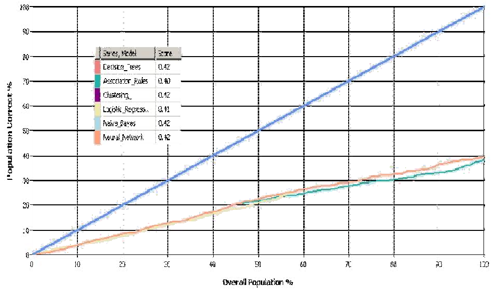


Figure 9: Accuracy Score For The {Ex., V.G., G., P., F., V.B.} Set

Table 2: Accuracy Score for the First Year Grades Using the High School Grade

Algorithm	Decision Trees	Association Rules	Clustering	Logistic Regression	Naive Bayes	Neural Networks
Grades	42%	40%	42%	41%	42%	42%
Pass/Fail	80%	76%	80%	80%	80%	79%

From the shown figures, it is clear that Decision Trees, Clustering and Naïve Bayes score was a little more than the other three for the sets {pass, fail} and {excellent, very good, good, pass, fail, very bad, absent} while Association rules came out the last with the least score for both sets.
The results show that the predictions for the {pass, fail} set are noticeably more accurate than for the {excellent, very good, good, pass, fail, very bad} set, which is expected given the much larger number of grades to be predicted. The results also show that none of the algorithms outperformed the others since all the scores are close to each other in both sets, but the Association Rules algorithm got the least score for both sets as shown in **Table 8**.

5.3 Adding more input attributes to the mining model

In this section we compare the accuracy of the same set of algorithms using two inputs instead of one; namely; the high school type and grade as the only inputs. The predictions accuracy of the selected algorithms are shown in **Figure 10**, and **Figure 11**.

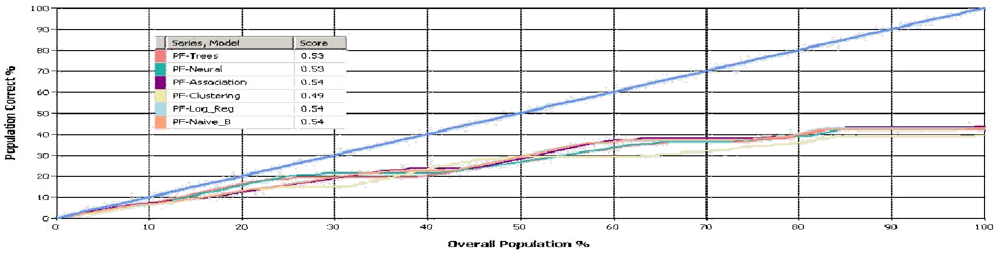


Figure 10: Accuracy Score for the {Pass, Fail} Set

Table 9 shows the results of the classification matrix created for the mining model using the six algorithms. Because there are only two possible values for this predictable attribute, 0 and 1, it is fairly easy to tell how often the model correctly makes a prediction.

The first table "Decision Trees", the first result cell, which contains the value 102, indicates the number of true positives for the value 0. Because 0 indicates that the student failed, the cell contains the value 102 means that in 102 cases the model predicted correctly that the student would not pass.

The cell directly underneath that one, which contains the value 169, tells you the number of false positives, or how many times the model predicted that someone would pass while actually they did not.

The cell that contains the value 59 indicates the number of false positives for the value 1. Because 1 means that the student did pass, this statistic tells you that in 59 cases, the model predicted the student would not pass while in fact they did. Finally, the cell that contains the value 548 indicates the number of true positives for the target value of 1. In other words, in 548 cases the model correctly predicted that the student would pass.

Table 10 shows the Classification Matrix Counts for the {Ex., V.G, G, P, F., V. B.} set using all six algorithms

Table 3: Classification Matrix Counts for the {Pass, Fail} Set using all six algorithms

Decision Trees		
Predicted	0 (Actual)	1 (Actual)
0	102	59
1	169	548

Association Rules		
Predicted	0 (Actual)	1 (Actual)
0	113	52
1	158	555

Clustering		
Predicted	0 (Actual)	1 (Actual)
0	0	0
1	271	607

Logistic Regression		
Predicted	0 (Actual)	1 (Actual)
0	110	58
1	161	549

Naïve Bayes		
Predicted	0 (Actual)	1 (Actual)
0	92	41
1	179	566

Neural Networks		
Predicted	0 (Actual)	1 (Actual)
0	107	56
1	164	551

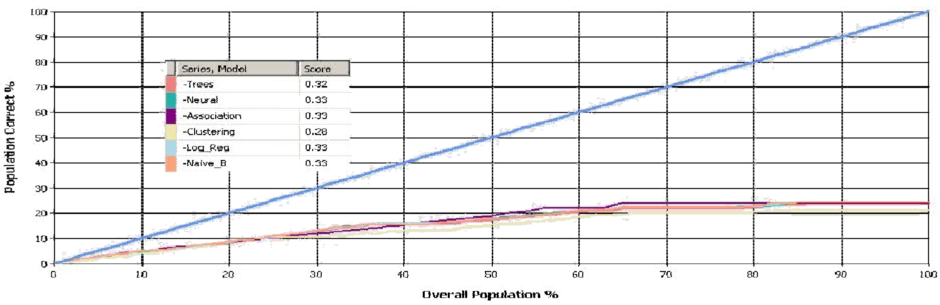


Figure 11: Accuracy Score For the {Ex., V.G., G., P., F., VB} Set

Association Rules, Logistic Regression and Naive Bayes score was a little more than the other algorithms for the sets {pass, fail} and {excellent, very good, good, pass, fail, very bad, absent} while Clustering came out the last with the least score for both sets.

The results show that the predictions for the {pass, fail} set are noticeably more accurate than for the {excellent, very good, good, pass, fail, very bad} set, which is expectedly given the much larger number of grades to be predicted. The results also show that none of the algorithms outperformed the others since all the scores are close to each other in both sets, but the Clustering algorithm got the least score for both sets as shown in Table 9.

Table 4: Classification Matrix Counts for the {Ex., V.G, G, P, F., V. B.} Set using all six algorithms

Decision Trees

Predicted	1 (Actual)	6 (Actual)	5 (Actual)	3 (Actual)	4 (Actual)	2 (Actual)
1	0	0	0	0	0	0
6	0	0	0	0	0	0
5	1	29	61	4	37	0
3	0	0	0	0	0	0
4	18	27	148	125	273	80
2	13	2	4	11	18	27

Association Rules

Predicted	1 (Actual)	6 (Actual)	5 (Actual)	3 (Actual)	4 (Actual)	2 (Actual)
1	0	0	0	0	0	0
6	0	2	1	0	0	0
5	2	31	79	5	42	1
3	0	1	0	3	2	0
4	30	24	133	132	284	106
2	0	0	0	0	0	0

Clustering

Predicted	1 (Actual)	6 (Actual)	5 (Actual)	3 (Actual)	4 (Actual)	2 (Actual)
1	0	0	0	0	0	0
6	0	0	0	0	0	0
5	0	0	0	0	0	0
3	0	0	0	0	0	0
4	32	57	213	137	325	107
2	0	1	0	3	3	0

Logistic Regression

Predicted	1 (Actual)	6 (Actual)	5 (Actual)	3 (Actual)	4 (Actual)	2 (Actual)
1	0	0	2	1	1	0
6	0	1	0	3	3	0
5	2	33	82	7	46	1
3	0	0	2	3	4	2
4	17	24	123	120	260	77
2	13	0	4	6	14	27

Naïve Bayes

Predicted	1 (Actual)	6 (Actual)	5 (Actual)	3 (Actual)	4 (Actual)	2 (Actual)
1	0	0	0	0	0	0
6	0	0	0	0	0	0
5	2	30	74	7	40	1
3	0	1	3	3	4	2
4	17	27	132	124	270	77
2	13	0	4	6	14	27

Neural Network						
Predicted	1 (Actual)	6 (Actual)	5 (Actual)	3 (Actual)	4 (Actual)	2 (Actual)
1	0	0	0	0	0	0
6	0	0	0	0	0	0
5	2	34	84	10	49	1
3	0	0	2	3	4	2
4	18	24	123	121	261	78
2	12	0	4	6	14	26

Table 5: Accuracy Score for the First Year Grades Using the High School Type and Grade

Algorithm	Decision Trees	Neural Networks	Association Rules	Clustering	Logistic Regression	Naive Bayes
Grades	32%	33%	33%	28%	33%	33%
Pass/Fail	53%	53%	54%	49%	54%	54%

6. Predicting High School Grade for a Given Faculty Success Level

This section compares the accuracy of (Decision Trees, Neural Network, Association Rules, Clustering, Logistic Regression and Naive Bayes) algorithms for predicting the High School Grade for a Given Faculty Success Level for the first semester for the MAE students.

Association Rules score was a little more than the other five for both sets while Clustering came out the last with the least score for both sets.

The results show that the predictions for the {pass, fail} set are noticeably less accurate than for the {excellent, very good, good, pass, fail, very bad} set. The results also show that none of the algorithms outperformed the others in either sets since all the scores are close to each other as shown in Table 12.

Algorithm	Decision Trees	Neural Networks	Association Rules	Clustering	Logistic Regression	Naive Bayes
Grades	54%	55%	56%	51%	55%	55%
Pass/Fail	47%	47%	49%	45%	48%	48%

Table 6: Accuracy Score for Predicting High School Grade for a Given Faculty Success Level

7. Scheme Analysis

The high school grades distribution can dynamically changed, from year to year, so the data mining technique that is most adopt to such distribution should be selected accordingly.

The proposed scheme allows the selection of the appropriate data mining technique that is best fit the distribution of the grades of the high school. However, the scheme guarantee the best possible prediction results

8. Conclusion

Six algorithms that handle classification of discrete attributes in data mining are implemented, verified and presented in this work. After determining which is more likely to be predicted between the total grade and the subject's grades, a comparison between the results of the six algorithms has been performed to determine the accuracy for predicting the academic performance of the first semester for the students by using the high school grade as the only input.

The comparison was held over two sets of data; the first set is {pass, fail} while the other set is {excellent, very good, good, pass, fail, very bad, absent}.

From the results shown in tables 6 and 7, high school grade is good for predicting the {pass, fail} for almost all subjects, but to predict the excellent students it did not produce high enough score that can be depended on. So, it is clear that more factors are needed to be taken into consideration.

As for the total grade, it scored a little less than the average score of the subjects where it scored 80% for the {pass, fail} set, and 42% for the {excellent, very good, good, pass, fail, very bad} set, while the average score for subjects was 82%, 44% for the two sets respectively.

Also as seen in tables 8 and 11 it is found that the prediction produced much higher scores for both sets when using the high school grade alone, but when the high school type was added to the prediction process the scores dropped noticeably. This is due to the less number of records to work on and there are some high school types with very little number of students.

The proposed scheme presented in this article assumes only two inputs, the high school grades and the high school type. An individual implementation of each step of the scheme shown in Figure-1is prepared and executed. The related results has been introduced and showed that the scheme is working properly.

9. Future Work

More inputs can be introduced to the proposed scheme to predict the academic performance such as age of student at admission, time that has elapsed between graduating from secondary school and gaining university admission, parents educational status, zonal location of student's secondary school, type of secondary school attended (privately owned, state or federal government owned), location of university and place of residence, and student's gender.

On the other hand, a deeper analysis can be done by testing the grades of each subject in the high school and see if it might give a high prediction score for any subject in the collage which could then result in the selection of students not only by the high school grade but also by demanding a minimum grade in certain subjects to ensure the quality of students.

References

1. Rules and admission requirements at private colleges and institutes . http;//tansik.egypt.gov.eg/tapplication/fanidalil/5-1.htm, ARE, 2011.
2. Rules of admission of students who obtained a high school commercial diploma or equivalents (3 years) http;//tansik.egypt.gov.eg/tapplication/fanidalil/index1.aspx, ARE, 2011.
3. Aavo Luuk, Kersti Luuk. Predicting students' academic performance in Aviation College from their admission test results. 2008.
4. Christopher M. Cornwell, David B. Mustard and Jessica Van Parys. How Does the New SAT Predict Academic Achievement in College. 2008.
5. V.O. Oladokun, A.T. Adebanjo and O.E. Charles-Owaba. Predicting Students' Academic Performance using Artificial Neural Network - A Case Study of an Engineering Course. The Pacific Journal of Science and Technology, Volume 9, Number 1, 2008.
6. Nguyen Thai Nghe, Paul Janecek, and Peter Haddawy. A comparative analysis of techniques for predicting academic performance. 37th ASEE/IEEE Frontiers in Education Conference, 2007.

7. Zaidah Ibrahim and Daliela Rusli. Predicting Students' Academic Performance – Comparing Artificial Neural Network, Decision Tree and Linear Regression, 2007.
8. Nathan R. Kuncel, Marcus Crede, LISA L. Thomas. A Meta-Analysis of the Predictive Validity of the Graduate Management Admission Test (GMAT) and Undergraduate Grade Point Average (UGPA) for Graduate Student Academic Performance. Academy of Management Learning & Education, Vol. 6, No. 1, 2007.
9. Paul Golding and Opal Donaldson. Predicting Academic Performance. 36th ASEE/IEEE Frontiers in Education Conference, 2006.
10. Paul Golding and Sophia McNamarah. Predicting Academic Performance in the School of Computing & Information Technology (SCIT). 35th ASEE/IEEE Frontiers in Education Conference, 2005.
11. Nathan R. Kuncel, Sarah A. Hezlett and Deniz S. Ones. Academic Performance, Career Potential, Creativity, and Job Performance - Can One Construct Predict Them All. Journal of Personality and Social Psychology Vol. 86, No. 1, 2004.
12. S. E. Kruck and Diane Lending. Predicting Academic Performance in an Introductory College-Level IS Course, 2003.
13. DERVIS Z. DENIZ and IBRAHIM ERSAN, An Academic Decision-Support System Based on Academic Performance Evaluation for Student and Program Assessment, Int. J. Engng Ed. Vol. 18, No. 2, 2002.
14. Dervis Z. Deniz and Ibrahim Ersan. Academic DSS for Student, Course and Program Assessment. International Conference on Engineering Education, 2001. Paul Golding and Sophia McNamarah. Predicting Academic Performance in the School of Computing & Information Technology (SCIT). 35th ASEE/IEEE Frontiers in Education Conference, 2005.
15. Paul Golding and Opal Donaldson. Predicting Academic Performance. 36th ASEE/IEEE Frontiers in Education Conference, 2006.
16. Nguyen Thai Nghe, Paul Janecek, and Peter Haddawy. A comparative analysis of techniques for predicting academic performance. 37th ASEE/IEEE Frontiers in Education Conference, 2007.
17. Dervis Z. Deniz and Ibrahim Ersan. Academic DSS for Student, Course and Program Assessment. International Conference on Engineering Education, 2001.

18. V.O. Oladokun, A.T. Adebanjo and O.E. Charles-Owaba. Predicting Students' Academic Performance using Artificial Neural Network - A Case Study of an Engineering Course. The Pacific Journal of Science and Technology, Volume 9, Number 1, 2008.
19. Aavo Luuk, Kersti Luuk. Predicting students' academic performance in Aviation College from their admission test results. 2008
20. Christopher M. Cornwell, David B. Mustard and Jessica Van Parys. How Does the New SAT Predict Academic Achievement in College. 2008.
21. Nathan R. Kuncel, Sarah A. Hezlett and Deniz S. Ones. Academic Performance, Career Potential, Creativity, and Job Performance - Can One Construct Predict Them All. Journal of Personality and Social Psychology Vol. 86, No. 1, 2004.
22. Nathan R. Kuncel, Marcus Crede, LISA L. Thomas. A Meta-Analysis of the Predictive Validity of the Graduate Management Admission Test (GMAT) and Undergraduate Grade Point Average (UGPA) for Graduate Student Academic Performance. Academy of Management Learning & Education, Vol. 6, No. 1, 2007
23. Zaidah Ibrahim and Daliela Rusli, 2007 – Predicting Students' Academic Performance – Comparing Artificial Neural Network, Decision Tree and Linear Regression.
24. Aavo Luuk, Kersti Luuk. Predicting students' academic performance in Aviation College from their admission test results. 2008. [6]
25. Christopher M. Cornwell, David B. Mustard and Jessica Van Parys. How Does the New SAT Predict Academic Achievement in College. 2008. [7]