# Data Mining in Higher Education : University Student Dropout Case Study

**Article** · January 2015

DOI: 10.5121/ijdkp.2015.5102



**2 authors**, including:

Alaa Mustafa El-Halees
Islamic University of Gaza
**42** PUBLICATIONS **560** CITATIONS

SEE PROFILE

# DATA MINING IN HIGHER EDUCATION : UNIVERSITY STUDENT DROPOUT CASE STUDY

Ghadeer S. Abu-Oda and Alaa M. El-Halees

Faculty of Information Technology, Islamic University of Gaza – Palestine

***ABSTRACT***

*In this paper, we apply different data mining approaches for the purpose of examining and predicting students' dropouts through their university programs. For the subject of the study we select a total of 1290 records of computer science students Graduated from ALAQSA University between 2005 and 2011. The collected data included student study history and transcript for courses taught in the first two years of computer science major in addition to student GPA , high school average , and class label of (yes ,No) to indicate whether the student graduated from the chosen major or not. In order to classify and predict dropout students, different classifiers have been trained on our data sets including Decision Tree (DT), Naive Bayes (NB). These methods were tested using 10-fold cross validation. The accuracy of DT, and NlB classifiers were 98.14% and 96.86% respectively. The study also includes discovering hidden relationships between student dropout status and enrolment persistence by mining a frequent cases using FP-growth algorithm.*

## 1. INTRODUCTION

Nowadays, higher learning institutions encounter many problems, which keep them away from achieving their quality objectives. Most of these problems caused from knowledge gap. Knowledge gap is the lack of significant knowledge at the educational main processes such as advising, planning, registration, evaluation and marketing. [1] For example, many learning institutions do not have access to the necessary information to advise students. Therefore, they are not able to give suitable recommendation for them.

Data mining is a powerful technology that can be best defined as the automated process of extracting useful knowledge and information including, patterns, associations [2]. The knowledge discovered by data mining techniques would enable the higher learning institutions in divers ways not limited to making better decisions, having more advanced planning in directing students, predicting individual behaviors with higher accuracy, and enabling the institution to allocate resources and staff more effectively. It results in improving the effectiveness and efficiency of the processes [3] [4] .

One of the biggest challenges that higher education faces today is predicting the academic paths of student. Many higher education systems are unable detecting student population who are likely to drop out because of lack of intelligence method to use information and guidance from the university system. Although, various types of studies exist to analyze student persistence patterns,

DOI : 10.5121/ijdkp.2015.5102 

it remains a challenging task to accurately predict a currently enrolled student's likelihood of returning to university the next term or change his major [5] [6].

Developing learning and teaching initiatives to improve retention and progression in education process is an important academic concern, which mainly depends on monitoring student performance and exploiting student feedback. Student marks and achievement are the main sources to study student feedback and progress, yet university and educational centers can use to predict the student performance, student dropout, and study path.

The main objective of this study is to identify those students who are less likely to return from one semester to other. Having these students identified soon enough for the institution to accommodate its interventions and marketing strategies will greatly enhance the student persistence rate in specific majors. This paper focuses on computer science major of students Graduated from ALAQSA University. In the rest of the paper, we cited related studies that employ data mining approaches in educational domain in section 2. Section 3 expresses in more detail the tasks of preparation and preprocessing of data set for the further analysis. Classification models and their accuracy are written on section 4, where the mined association rules and their analysis are in section 5.

## 2. RELATED WORK

Luan in [7] aims to predict student's persistency by considering the students who either re-enroll or does not re-enroll the following semester. Using the prediction techniques, the likelihood of a student's persistency can be determined. The identification of these students is very useful to universities because if they know those who are less likely to persist, then the university can attempt to increase the persistency rate of these students. By knowing students who are not persistence, the faculty can identify the factors affecting their non-persistency. Therefore, the university's managerial systems have to attempt on improving these factors, which would result in improving the student's persistency rate.

In [8] applied the classification task as one data mining technique to evaluate student' performance, they used decision tree method for classification. They depend on extracting actual student performance indicators by the end of semester examination using the student' previous database including Attendance, Class test, Seminar, and Assignment marks. This study helps earlier in identifying the dropouts and students who need special attention and allow the teacher to provide appropriate advising.

In [9] have employed decision tree to predict to help have the tutor to identify the weak students and improve their Performance before being dropouts. The model aims to classify the students into PASS and FAIL target even before the students take their exams.

Applying modern machine learning methods and cost effective analysis is the main contribute of author in [10], he predicts dropout of new students with satisfying accuracy and thus become a useful tool in an attempt to reduce dropouts. He collects the data set is from the module 'Introduction in Informatics' from the distance education system as an application on eLearning systems. With the support of university freshmen, authors conducted a study to early predict student failure that may lead to drop-out [11] , They used decision tree, Bayesian classifiers,

logistic models, the rule-based learner and random forest to detect/predict first year student drop out .

Authors in [12] go a further step for acquiring data-set in educational domain, they depend on social networks to acquire relevant information that can be used in educational data mining tasks to improve the accuracy of classification models in comparison with usage of only demographic and academic attributes, e.g., students' age, gender, or number of finished semesters. They measure the prediction of dropouts: degree, indegree, outdegree, and betweenness. They conclude that these improve the accuracy of classification models.

## 3. DATA PREPARATION AND PREPROCESSING

The data is collected from ALAQSA university database for bachelor degree students graduated between 2005 and 2011. The attributes selected are indicated in table 3.1. The courses selected for the study are those courses presented in computer science major taken in the first two years of study according to major academic plan, as a suggestion that the student may likely be dropped from the major in the first two years of study. The number of records obtained is about (1290) record representing graduate students and their transcripts. For the first view of data, figure 3.1 shows the large number of students dropped from their original chosen major which is in this case computer science department.

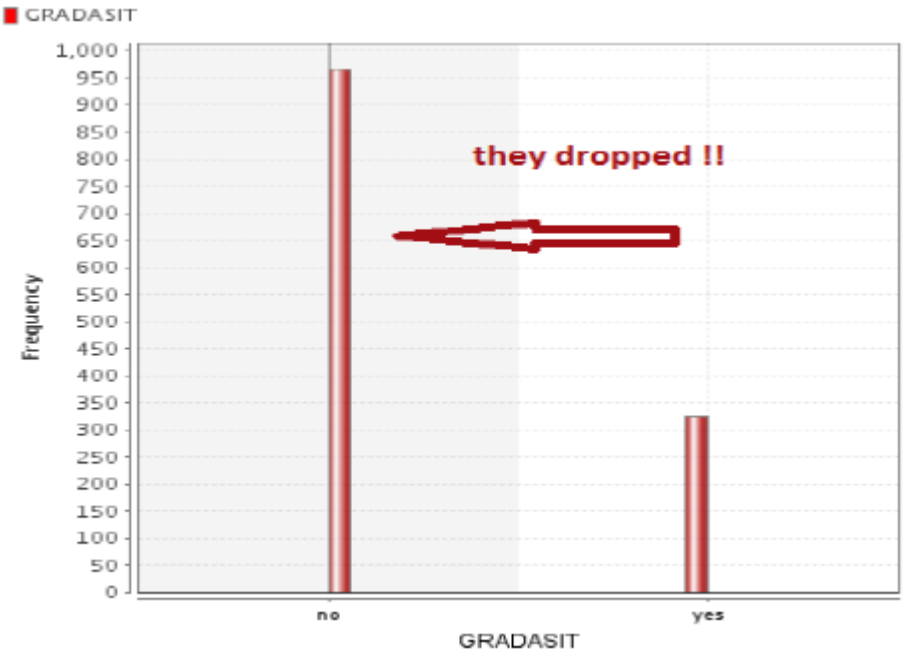


Figure 3-1 Number of dropouts

### 3.1 Data Integration

The goal of integration is to retrieve the records that represents the graduated students history with their associated enrollment data, graduated GPA, major, and their marks in computer

science subjects (subject is interchangeably means course) conforming to the computer science department plan.

According to the relational database schema in the ALAQSA university system, the student enrollment information is held in different relation to the one that holds courses transcripts for each university semester. Figure 2 shows a set of retrieved records by a simple join of the two relations; each student has a record for each course he had registered. However, this kind of combining is not suitable for our processing for instance, (i) there are many repeated records for each student who studied more than one course of the chosen major. (ii) The student mark in each course is scattered in midterm, work term, and final term. (iii) Each record has the major from which the student had graduated including the computer science major; which means the target class has no rigid values but the different university academic programs.

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | STUDENTNO | SUBJECTNO | MARK | AVAREGE | GPA | SEX | MAJOR_NO |
| 2 | 0111237 | COMP1310 | 60 | 68.6 | 75.2391 | 1 | 111 |
| 3 | 0111237 | COMP2313 | 73 | 68.6 | 75.2391 | 1 | 111 |
| 4 | 0111237 | COMP3312 | 70 | 68.6 | 75.2391 | 1 | 111 |
| 5 | 0111237 | COMP3313 | 76 | 68.6 | 75.2391 | 1 | 111 |
| 6 | 0111237 | COMP4310 | 85 | 68.6 | 75.2391 | 1 | 111 |
| 7 | 950096 | COMP1310 | 60 | 57.3 | 68.5913 | 0 | 111 |
| 8 | 950096 | COMP2313 | 63 | 57.3 | 68.5913 | 0 | 111 |
| 9 | 950096 | COMP3312 | 45 | 57.3 | 68.5913 | 0 | 111 |
| 10 | 950096 | COMP4310 | 78 | 57.3 | 68.5913 | 0 | 111 |
| 11 | 950096 | COMP3312 | 47 | 57.3 | 68.5913 | 0 | 111 |
| 12 | 950096 | COMP3313 | 68 | 57.3 | 68.5913 | 0 | 111 |
| 13 | 950096 | COMP3312 | 42 | 57.3 | 68.5913 | 0 | 111 |

Figure 3-2 Integration resulted records

For data preparation we do the following modifications: 1) get the sum of the three attribute (midterm, work term, and final term) can result the student overall mark in each course. 2) We use two values for target class (MAJOR_NO), yes value if student has graduated from the computer science major and no, for others. Thus, we map target class (graduated major) to a computer science major or not, which is more applicable in our case. 3) To solve the repeated records for each student course we prepare SQL-based method to roll student courses into separated columns. Figure 3 shows how the data is after modification. The highlighted two records are for different student records, the courses taken is depicted as columns, and the student mark in the associated course is the column value otherwise the mark will be NULL in case the student did not join that course.

| STUDENTNO | COMP1310 | COMP3313 | COMP3312 | COMP431 | COMP231 | COMP332 | AVAREGE | GPA | GRADASIT |
|---|---|---|---|---|---|---|---|---|---|
| 980290 | 89 | | | | | | 80.9 | 83.3357 | no |
| 980188 | 78 | 92 | 80 | 65 | 79 | | 78.6 | 82.8 | no |
| 970997 | 62 | 75 | 68 | 60 | 77 | | 72 | 71.5214 | no |
| 980176 | 73 | 82 | 75 | 74 | 62 | | 79.2 | 78.9214 | no |
| 980192 | 89 | 91 | 94 | 78 | 85 | | 94.9 | 84.0857 | no |
| 970021 | 69 | 82 | 81 | 83 | 92 | | 83.5 | 85.9286 | no |
| 980200 | 91 | 78 | 90 | 90 | 81 | | 83.5 | 88.5857 | no |
| 980202 | 68 | 82 | 89 | 81 | 83 | | 77.8 | 80.2643 | no |
| 980284 | 83 | 90 | 91 | 80 | 91 | | 88.6 | 87.2071 | no |

Figure 3-3 The resulted modified records

The final attributes selected for each student record and used for further analysis and data-mining tasks is declared in the following table 6.attribute role, type, and description is defined.

Table 3.1 data-set Attributes and description

| Attribute | Role | Type | Description |
|---|---|---|---|
| Graduate AS IT | label | binominal | Indicate if the student graduated from IT or not. values (no, yes) |
| GPA | regular | Real | The overall student average within the whole four years of study. |
| SEX | regular | Binominal | Whether male or female student. Values (0,1) |
| Average | regular | Real | The student's secondary certificate grade before enroll university program |
| COMP3313 | regular | Nominal | Student mark in (database)course |
| COMP1322 | regular | Integer | The student mark in the code subject within the program associated plan (programming I) |
| COMP1310 | regular | Integer | Student mark at (introduction to computer science ) |
| COMP2315 | regular | Integer | Student mark at (algorithm analysis) |
| COMP2316 | regular | Integer | Student mark at (logic design) |
| COMP2312 | regular | Integer | Student mark at (data structure I) |
| COMP1311 | regular | Integer | Student mark at (data structure II) |
| Region_cd | regular | binominal | Student resident region (1,2)(south, north) |

### 3.2 Data preprocessing:

We did the following preprocessing steps:

#### 3.2.1 Replace missing values

As seen in figure (5), it is a student record in the data set collected, notice there are empty values for some columns. It is actually not error in the data, that is because the students did not enroll all the courses. Thus, his mark simply will be nothing.

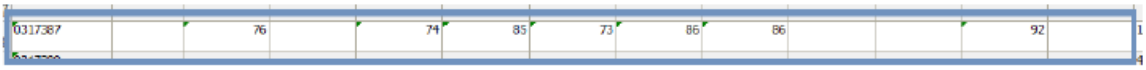


Figure 3-4 Missing Value Example

The missed values replaced by zero to indicate the student did not register the course at all (it does not replaced by the academic zero value, since it may be considered as being failed in the course)

ExampleSet (2910 examples, 1 special attribute, 15 regular attributes) View Filter (2910 / 2910):

| | GRADASIT | COMP1322 | COMP3301 | COMP1310 | COMP3313 | COMP2315 | COMP2312 | COMP2316 | COMP2313 | COMP3328 | COMP1311 | COMP2314 | SEX |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | no | 0 | 0 | 84 | 93 | 0 | 70 | 0 | 0 | 0 | 72 | 0 | 1 |
| 2 | no | 0 | 0 | 84 | 91 | 0 | 75 | 0 | 0 | 0 | 75 | 0 | 1 |

Figure 3-5 Missing values is replaced

### 3.2.2 Over Sampling:

According to the pie chart (see figure (3-7)), there is a large gap between the number of students graduated from computer science major to those graduated from the university in other majors (ratio 0.33, 0.66). I use over sampling technique to balance the number of records in the target class values. The sampling task in this case is done by feeding the data with records belong to the least occurred target class value. The added instances will be the existence records but trying to repeat it many time until balance the peer value of the label.

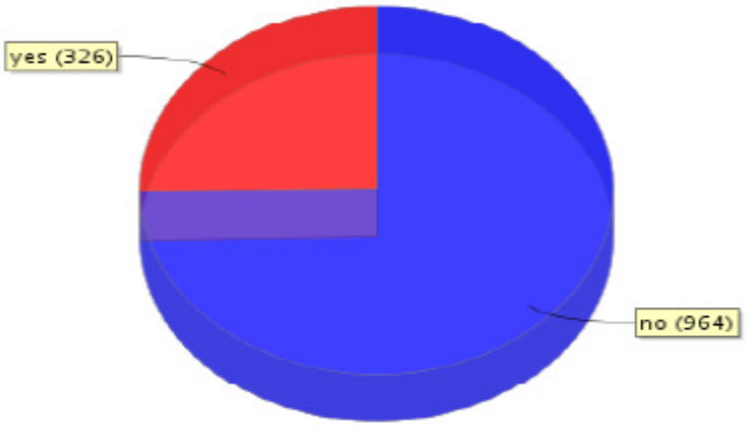


Figure 3-6 Ratio of 'Yes' to 'No' Labels

The resulted records after oversampling are increased to (1895) record because of inserted records with (964) for no label and (931) for yes label (Figure 3-8,3- 9).

ExampleSet (1895 examples, 1 special attribute, 11 regular attributes)

| Role | Name | Type | Range | Missings |
|---|---|---|---|---|
| label | GRADUATE | nominal | no (964), yes (931) | 0 |
| regular | COMP1322 | integer | [0.000 ; 99.000] | 0 |
| regular | COMP1310 | integer | [0.000 ; 97.000] | 0 |
| regular | COMP3313 | integer | [0.000 ; 95.000] | 0 |
| regular | COMP2315 | integer | [0.000 ; 95.000] | 0 |
| regular | COMP2312 | integer | [0.000 ; 97.000] | 0 |
| regular | COMP2316 | integer | [0.000 ; 98.000] | 0 |
| regular | COMP1311 | integer | [0.000 ; 99.000] | 0 |
| regular | SEX | nominal | 1.0 (1275), 0.0 (620) | 0 |
| regular | AVAREGE | real | [0.000 ; 98.400] | 0 |
| regular | GPA | real | [64.590 , 94.664] | 0 |
| regular | REGION CD | nominal | 2.0 (1238), 1.0 (657) | 0 |

Figure 3-7 Oversampling output

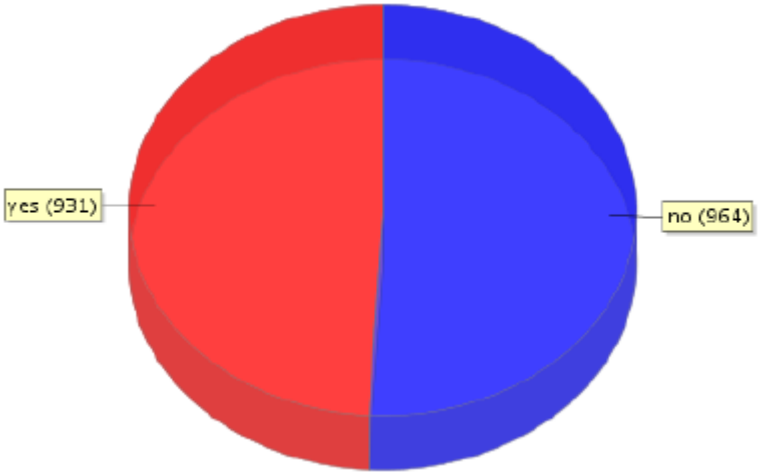


Figure 3-8 Over sampling resulted ratio

### 3.2.3 Discretize attributes values into groups:

Before applying any classification algorithm, the data has specific preprocessing step; it is discretize which is the task of converting the absolute column values into bin or ranges. For example, we need to have the columns of subject's marks, GPA, and average attributes using the ranges in figure (3-10) below. The resulted records will show as that in figure (3-11), notice integer values replaced by F, A, C. . . Categories.

| class names | upper limit |
|---|---|
| A | 100.0 |
| C | 80.0 |
| D | 70.0 |
| E | 60.0 |
| F | 50.0 |

Figure 3-9 Bins ranges

ExampleSet (1290 examples, 1 special attribute, 11 regular attributes) View Filter (1290 / 1290):

| Row... | GRADASIT | COMP1322 | COMP1310 | COMP3313 | COMP2315 | COMP2312 | COMP2... | COM |
|---|---|---|---|---|---|---|---|---|
| 1 | no | F | B | A | F | D | F | C |
| 2 | no | F | B | A | F | C | F | C |
| 3 | no | F | C | E | F | D | F | E |
| 4 | no | F | F | F | F | F | F | F |

Figure 3-10 Discreteize process resulted records

# 4. CLASSIFICATION

Classifying whether the student is dropped or graduated from the major is the core goal of this study; classification task is depicted to do the rules or paths that would lead to each case. At this phase, we train different models for predicting student graduation major. At all classification models the data is split into training of 0.6 (860) of data to 0.3 (430) of data for testing. We use 10-fold cross validation to test the accuracy of each model.

## 4.1 Decision tree

Decision Trees (DTs) are a supervised learning method used for classification and regression [13]. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features. We apply this method on our data set of student records to map observations about each student instance to conclusions about his target major of graduation.

Figure 4-1 depicts the decision tree that resulted from applying the decision-tree classification algorithm on the graduation major as a target class for each student record. As it is seen from the figure, the attribute of course “COMP2315” has a great influence on whether the student will drop out of computer science department or not. In other words, if the student got grade F on algorithm analysis course, the model concludes that he/she would dropped from the major.

The model presented in figure 4-1 has an accuracy of 98.14% as shown in figure 4-2. To interpret the rules in the decision tree, Table 4-1 contains more observed cases based on training set.

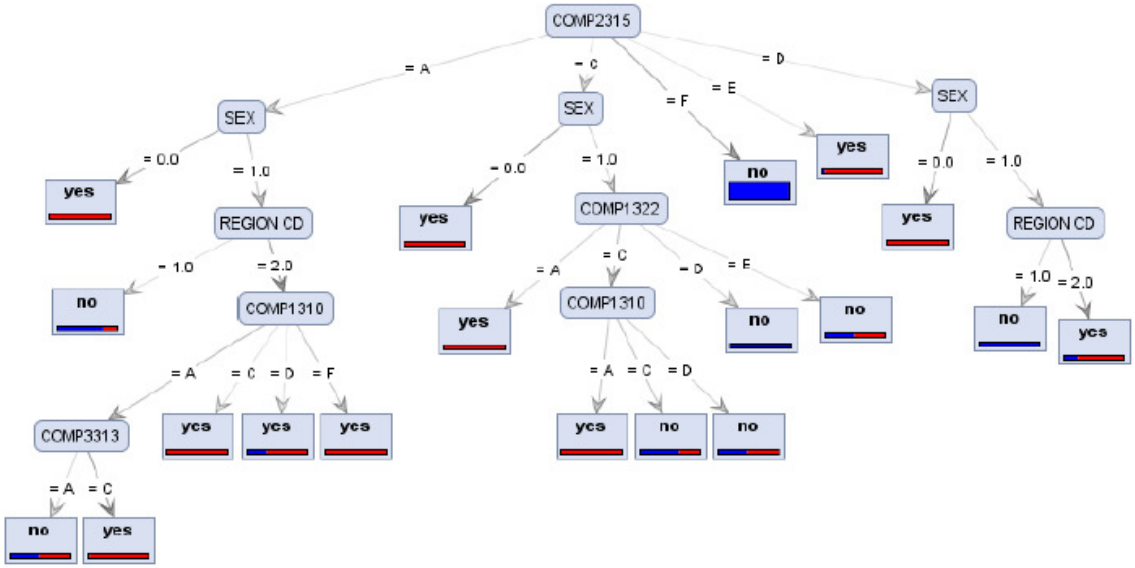


Figure 4-1 Decision Tree classifier output

Table 4-1 Decision Tree observation cases

| Case | Target class |
|---|---|
| If the student get grade A (more than 80%) in algorithm analysis course (COMP2315) and he is a boy --> | He will not be dropped from the computer science department. |
| If the student get grade C (more than 70%) in algorithm analysis course (COMP2315) and he is a boy --> | he will not be dropped from the computer science department |
| If the student get grade C (more than 70%) in algorithm analysis course (COMP2315) and she is a girl and get grade A in programming language course (COMP1322) --> | she will not be dropped from the computer science department |

**accuracy: 98.14%**

| | true no | true yes | class precision |
|---|---|---|---|
| pred. no | 309 | 2 | 99.36% |
| pred. yes | 6 | 113 | 94.96% |
| class recall | 98.10% | 98.26% | |

Figure 4-2 Decision-tree model accuracy

## 4.2 Naïve base algorithm

A Bayes classifier [14] is a simple probabilistic classifier depends on applying Bayes' theorem with strong independence assumptions. In simple terms, a naive Bayes classifier assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature.

By applying the algorithm in classifying around 1800 record in the built model, the chart in figure (4-3) indicates the classification for student college in case he register “digital design” course: if he got grade A , he would classified to be graduated from computer science department. The degree is reducing until it reached the maximum likelihood of being dropped in case he did fail the course and the classifier behavior will change. The accuracy is about 96% (see figure 14).

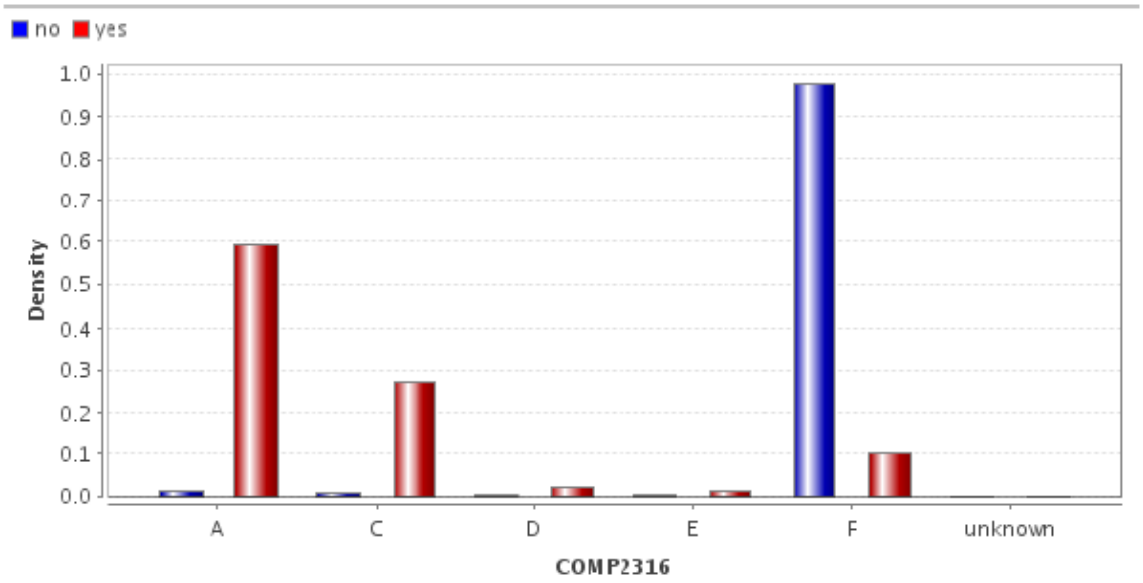


Figure 4-3 classifying record based on (COMP2316) course Grade

**accuracy: 96.98%**

| | true no | true yes | class precision |
|---|---|---|---|
| pred. no | 304 | 2 | 99.35% |
| pred. yes | 11 | 113 | 91.13% |
| class recall | 96.51% | 98.26% | |

Figure 4-4 Naïve base algorithm model accuracy

From this task, it can be concluded that the prediction technique is useful in predicting the likelihood of a student persistency as early as possible. By identifying the students who are less likely to persist, the university interacts by providing them the academic assist and as a result, student retention rate will be increased. This has a positive impact on improving the existence rate through enhancing the student’s assessment process in a higher learning institution.

Table 4-2 Naïve Bayes vs Decision-Tree accuracy

| Classifier | Accuracy (Average) |
|---|---|
| Decision Tree | 98.14% |
| Naïve Bayes | 96.86% |

# 5. ASSOCIATION RULES

In Data Mining, the task of finding frequent pattern in large databases is very important and has been studied in large scale in the past few years. It is useful for discovering interesting relationships hidden in large data sets. The association analysis reveals a set of strong relations between study cases that expressed into a set of association rules. An Association rules are if/then statements that help uncover relationships between seemingly unrelated data in a relational database or other information repository.

In educational domain, such a rule may indeed influence the university strategy and practices in aim to increase student performance, reduce the number of dropped students, or even review courses syllables early.

## 5.1 Frequent-pattern growth (FP-growth):

The FP-Growth Algorithm, proposed by Han in [15] is an efficient and scalable method for mining the complete set of frequent patterns by encode the data set using a compact data structure into frequent-pattern tree called FP-tree and extract frequent item sets directly from this structure

In our study, we apply FP_growth algorithm for a set of (1800) record samples in the hope of discovering un-known relationships between student study path and persistent enrollment in "computer science" major and not being dropout. Figure 15, and 16 show asset of rules obtained for label 'no' and 'yes' respectively.

| | |
|---|---|
| GPA = C, COMP2316 = F | GRADASIT = no |
| GPA = C, COMP2316 = F | GRADASIT = no, COMP2315 = F |
| SEX = 1.0, COMP2316 = F | GRADASIT = no |
| SEX = 1.0, COMP2316 = F | GRADASIT = no, COMP1322 = F |
| SEX = 1.0, COMP2316 = F | GRADASIT = no, COMP2315 = F |
| SEX = 1.0, COMP2316 = F | GRADASIT = no, COMP1322 = F, COMP2315 = F |
| GPA = C, COMP1322 = F | GRADASIT = no |
| GPA = C, COMP1322 = F | COMP2316 = F, GRADASIT = no |
| GPA = C, COMP1322 = F | GRADASIT = no, COMP2315 = F |
| GPA = C, COMP1322 = F | COMP2316 = F, GRADASIT = no, COMP2315 = F |
| COMP1322 = F | GRADASIT = no |
| COMP1322 = F | COMP2316 = F, GRADASIT = no |

Figure 5-1 Rules for label 'NO'

| Premises | Conclusion | Support |
|---|---|---|
| REGION CD = 2.0, COMP1311 = A | GRADASIT = yes, COMP2312 = A | 0.289 |
| COMP1311 = A, COMP2315 = A | REGION CD = 2.0, GRADASIT = yes | 0.286 |
| REGION CD = 2.0, COMP2312 = A | GRADASIT = yes, COMP1311 = A | 0.289 |
| COMP1322 = A | GRADASIT = yes, COMP1311 = A, COMP2315 = A | 0.280 |
| REGION CD = 2.0, COMP1311 = A | GRADASIT = yes | 0.292 |
| COMP1322 = A | GRADASIT = yes, COMP2315 = A | 0.281 |
| COMP1322 = A | GRADASIT = yes, COMP2312 = A, COMP1311 = A | 0.282 |

Figure 5-2 Rules for label 'Yes'

Table 5-1 Example of Association Rules

| Rule # | Rule Statement |
|---|---|
| 1 | [SEX = 1.0, COMP2315 = F]--> [GRADASIT = no, COMP1322 = F] (confidence: 0.994) |
| 2 | [REGION CD = 2.0, COMP2312 = A, COMP1311 = A, COMP2315 = A] --> [GRADASIT = yes] (Confidence: 0.994) |

For table 5-1, we explore two examples of association rules inferred after applying the FP-growth algorithm. Rule 1 argues the following: if a student is a boy and got grade F (less than 50%) in "algorithm analysis" course, he is likely to be dropped from computer science department and subsequently fail in "programming language" course. That is true by 0.95 for boys who fail in this course. In addition, it is 0.994 for all accrued cases.

However, Rule 2 indicates the following: If the case in the hand is for a student who is resident in the south areas of Gaza Strip, and get grade A (more than 80%) in both data structure I and II courses, and "algorithm analysis" course, then the student likely continue in computer science department and will not be dropped. It will be happens for all students has the same conditions by 0.95 and this is true by 0.994 according to accrued cases.

## 6. CONCLUSION

In this paper, we study the student dropout in computer science major in ALAQSA University for the purpose of improving the current teaching procedures and education strategies. Technologically, we do not propose new methods like FB-growth or decision-tree. However, we only use the mature classification and approaches in this study, cannot present more competitive algorithms or improve the existing algorithms. The study finds that mastering "digital design" and "algorithm analysis" courses has a great affect on predicting student persistence in the major and decrease student likelihood of dropout.

## REFERENCES


[1] P. Baepler and C. J. Murdoch, "Academic Analytics and Data Mining in Higher Education," International Journal for the Scholarship of Teaching and Learning, vol. 4, no. 2, pp. 1-9, 2010.

[2] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth and R. Uthurusamy, "Advances in knowledge discovery and data mining," 1996.

[3] R. S. J. D. Baker and K. Yacef, "The State of Educational Data Mining in 2009 : A Review and Future Visions," Journal of Educational Data Mining, vol. 1, no. 1, pp. 3-16, 2009.

[4] A. AL-Malaise, A. Malibari and M. Alkhozae, "STUDENTS' PERFORMANCE PREDICTION SYSTEM USING MULTI AGENT DATA MINING TECHNIQUE," International Journal of Data Mining & Knowledge Management Process (IJDKP) , vol. 4, 2014.

[5] C. Romero and S. Ventura, "Educational Data Mining: A Review of the State of the Art," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 40, no. 6, 2010.

[6] B. R.B., T. S.S and S. A.K, "Importance of Data Mining in Higher Education System," Journal Of Humanities And Social Science (IOSR-JHSS), vol. 6, no. 6, pp. 18-21, 2013.

[7] J. Luan, "Data Mining and Knowledge Management in Higher Education -Potential Applications.," in Processdings ofAIR Forum, Torento , Canada, 2002.

[8] B. Baradwaj and S. Pal, "Mining educational data to analyze student's performance," Internation Journal od Advamced Computer Science and Applications, vol. 2, no. 6, pp. 63-69, 2012.

[9] A. Kumar and Vijayalakshmi, "Implication Of Classification Techniques In Predicting Student's Recital," International Journal of Data Mining & Knowledge Management Process, vol. 1, no. 5, pp. 41-51, 2011.

[10] S. Kotsiantis, "Educational data mining: a case study for predicting dropout-prone students," International Journal of Knowledge Engineering and Soft Data Paradigms, vol. 1, no. 2, p. 101, 2009.

[11] D. G. W, P. Mykola and V. J. M, "Predicting Students Drop Out: A Case Study," International Working Group on Educational Data Mining, 2009.

[12] B. Jaroslav, H. Bydzovská, J. Géryk, T. Obsivac and L. Popelinsky, "Predicting Drop-Out from Social Behaviour of Students," International Educational Data Mining Society, 2012.

[13] L. Rokach, Data mining with decision trees: theory and applications, vol. 69, World scientific, 2008.

[14] S. J. Russell and P. Norvig., Artificial Intelligence: A Modern Approach (AIMA), 3rd ed., Prentice Hall, 2009.

[15] J. a. P. Han and Y. Jian and Yin, "Mining frequent patterns without candidate generation," in ACM SIGMOD Record, 2000.