

# Factors Influencing Students Decisions to Enrollment in Sudanese Higher Education Institutions

### Zainab M. Sedahmed<sup>1</sup>, Noureldien A. Noureldien<sup>2</sup>

<sup>1</sup>Future University, Khartoum, Sudan <sup>2</sup>University of Science and Technology, Omdurman, Sudan Email: Zeinab\_Sedahmed@hotmail.com, Noureldien@hotmail.com

How to cite this paper: Sedahmed, Z.M. and Noureldien, N.A. (2019) Factors Influencing Students Decisions to Enrollment in Sudanese Higher Education Institutions. *Intelligent Information Management*, **11**, 61-76.

https://doi.org/10.4236/iim.2019.114006

**Received:** May 28, 2019 **Accepted:** July 27, 2019 **Published:** July 30, 2019

Copyright © 2019 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

http://creativecommons.org/licenses/by/4.0/

## Abstract

There is a growing body of literature that recognizes the importance of data mining in educational systems. This recognition makes educational data mining a new growing research community. One way to achieve the highest level of quality in a higher education system is by discovering knowledge from educational data such as students' enrollment data. Many mining tools that aim to discover exciting correlations, frequent patterns, associations, or casual structures among sets of items in educational data sets have been proposed. One of the widely used tools is association rules. In this paper, the Apriori algorithm is used to generate association rules to discover the importance and correlation between factors that influence student's decision to enroll in higher education institutions in Sudan. The algorithm is applied using a student's enrollment data set that was created using a questionnaire and 800 students enrolled in governmental and private sector universities as a sample. This paper classifies factors that influence enrollment into: student's demographic factors and four categories of enrollment related factors (Student and Society, Educational Institution, Admission, and Employment related factors), and determines the most influential factors in determining student's decision to enroll in Sudanese universities. The analysis result shows that the Educational Institution related factors (50%) and Admission related factors (40%) are strongly influencing students' enrollment decision, while the Employment related factors (10%) and Student and Society related factors (0%) have weak influence. The factors out of the 14 Educational Institution related factors that have a high impact are: reputation, diversity of study, quality of education, education facilities, and feasibility.

### **Keywords**

Enrolment Related Factors, Educational Data Mining, Association Rules,

Student and Society Related Factors, Educational Institution Related Factors, Admission Related Factors, Employment-Related Factors

#### **1. Introduction**

Enrollment management is one of the most important education process phases that refer back to the late nineteenth century when Harvard University founded the board of freshman advisors in 1889. The board's purpose was to establish orientation, provide advising and counseling, and develop social events for freshmen [1]. Since then, enrollment management is increasingly getting the attention of scholars in educational data mining, particularly in computer-related disciplines.

Data mining is a field of computer science that focuses on the detection of patterns and hidden knowledge in enormous data and gives the information in logical form. Artificial intelligence, machine learning, statistics are some techniques applied in data mining [2].

One of the emerging interdisciplinary study areas is the educational data mining, which is a new growing research community [3]. Educational data mining concerns the developing and using methods that can explore or extract interesting information from educational data [4]. One of these methods is association rules.

An association rule is an unsupervised learning method that used for pattern discovery, which in turn may reveal new knowledge and exciting discovery. Association rules are considered as an important method for decision making if it has a support and confidence that is at least equal to some minimal support and confidence thresholds defined by the user. Association rules mining was first encountered by Agrawal *et al.* [5].

There are various association rule mining algorithms such as AIS, SETM, Apriori, Aprioritid, Apriorihybrid, FP-growth [6]. The Apriori algorithm is the most well-known association algorithm used for finding frequent itemsets with candidate generation [7].

Using association rules to discover new knowledge in educational data is a common method [4] [8] [9], more specifically, many studies apply association rules algorithms to enrollment data sets [10] [11] [12].

This paper applies the Apriori algorithm to a student's enrollment data set that was created using a questionnaire and a sample of students who are enrolled in governmental and private sector Sudanese Universities. The Apriori algorithm is selected since it is the most frequently used association rule algorithm. The aim is to discover the most influential category among enrollment related factor categories, and further to discover the essential factors within each influential category. Also, the paper shows the correlation between factors within the different categories that influence the student's decision to enroll in Sudanese Universities.

The expected extracted knowledge can be of great value and offer a helpful and constructive recommendation to the academic planners to improve enrollment to their higher institutions. The rest of the paper is organized as follows: Section 2 discusses related work that applies association rules as a mining tool to enrollment data sets. Section 3 reviews association rules and the Apriori algorithm. Section 4 explains how the enrollment data set for students enrolled in the Sudanese Universities were created. Section 5 presents how association rule mining is applied to Sudanese Universities enrollment dataset. Section 6 is results and discussion and the conclusion was drawn in Section 7.

#### 2. Related Work

Many researchers study enrollment related factors that influence student's selection decisions to enroll in higher education institutions. Recently, there has been more interest in extract various relations from educational data by using association rules as the data mining tool.

Many researchers examined the influence of the student and society related factors in enrollment decision. Researchers prove that student educational aspirations have positively related with Higher Educational Institution (HEI) selection decisions [8] [9] [10]. Some studies compare aptitude and the ability of the students' as factors in the selection of HEI [11] [12]. Others recognize the critical role played by students' guardians, family and friends factors in directing student decide to enroll in specific HEI's [12] [13] [14]. Students' interests, motivations, and occupational plans, the class level, socioeconomic level, ethnic of the students, and institutional characteristics are examined to see their effects in institution's total enrollment [15] [16] [17].

Student's selection of colleges depends on several criteria, including academic quality, facilities, campus surroundings, and personal characteristics [12] [18]. Also, income affects the choice of students along the public-private education divide. The college location is a significant predictor of HEI selection, and a visit to the university campus was found out as an important factor [19].

The reputation of HEI's has a significant influence on the student's enrollment decision, which was examined by many researchers who found that engaging in international partnerships attracts larger numbers of international students [20] [21].

Recent evidence suggests that there are needs of increasing diversification at the programs level for adopting more general programs with based on the diversity of students' sample population, multiple regional, social, and economic needs [22]. Also, the institutional image emphasized the significance of building positive emotions in achieving educational institution enrollment goals and the availability of postgraduate studies in institutions are affecting student's selection of HEI [23] [24] [25].

A considerable amount of literature was published on factors related to Ad-

mission and examines the influences of these factors on the decision of the student to enrolment. The recognition of the academic degree or the program nationally or internationally, the degree flexibility, the diversity of courses, and the flexible of entry requirements have influenced enrollment [18]. The marketing mix, marketing efforts, channels, and advertisement is found to be important factors that influence student's college selection [25] [26], also, financial aid induces more enrollments in colleges than other factors [27] [28].

Finding a job had become a central issue for students and their family recently. The decision of enrollment to specific education institution was affected by many factors related to the employment, and employment opportunities are a stronger predictor of enrollment decisions [19] [29].

Association rules have been used to discover relations between; admission system attributes in King Abdul-Aziz University (KAU) [30], the preliminary students knowledge [31], the specialties and student's interests [32], the factors that affect postgraduate study and assessment [33], and the courses and the failed students [34].

Association rules as a data mining technique are used to investigate the correlation between different enrollment factors. Some studies investigated the Apriori algorithm on enrollment to extract the behavior of low- and high-income students [35], the quality of talent training and enhance the overall competitiveness of colleges and universities [36], student performance for a certain outcome (Pass or Fail) [37].

This paper gets benefits from previous studies to determine the different factors influencing the enrollment decision. Then the paper provides a new categorization to these factors. To determine the most influential factors and the correlations between these factors the paper uses association rules in a similar way to [4] [38] [39]. The paper deals with enrollment dataset as in [30] [40] [41].

The paper is unique in that, it defines categorization to enrollment factors and uses the Apriori algorithm to extract association rules that determine which category has more roles in student's enrollment decisions, and what factors within this category are most important.

#### 3. Association Rules and the Apriori Algorithm

Association rule mining is a descriptive data mining technique for finding patterns, associations, and correlations among sets of items in a database. A standard association rule is a rule of the form  $X \rightarrow Y$  which means that if X is true for instance in a database, then Y is true for the same instance, with a certain level of significance [5] [13].

Typically, an association rule is called strong if it satisfies both a minimum support (min-supp) threshold and a minimum confidence (min-conf) threshold that is determined by the user [2].

The minimum support is defined as the minimum percentage of occurrences of the item/itemset in the database, while minimum confidence defined as the minimum certainty or trustworthiness associated with each discovered pattern [6].

Let *A* and *B* be itemsets in the database *D*. An association rule between item sets *A* and *B* is an implication of the form  $A \rightarrow B$ , where  $A \cap B = \emptyset$ . The support for an association rule  $A \rightarrow B$  denoted as sup  $(A \rightarrow B)$ , is defined as the number of transactions in D that contains  $A \cup B$  [9]. The item sets *A* and *B* are called antecedent and consequent, respectively.

The support of an itemset A, supp(A) is the proportion of transaction in the database in which the item A appears. It signifies the popularity of an itemset.

support = 
$$P(A, B)/N$$

The confidence determines how frequently items in B appear in the transactions that contain A; it is the ratio of the number of transactions that include all items in the association rule.

Confidence = 
$$P(B|A) = \operatorname{supp}(A, B)/\operatorname{supp}(A)$$

Moreover, one of the simplest correlation measures is Lift. Lift is important to measure the interestingness of a rule. Lift measure tells us whether the LHS (left-hand sides) of a correlation influences the RHS (right-hand sides) positively or negatively. The lift between the occurrence of item set A and B can define as [42]:

$$\operatorname{lift}(A,B) = \frac{\operatorname{conf}(A \to B)}{\operatorname{supp}(B)}$$

The strength of correlation was measured from the lift value as follows [42]:

- If Lift(A, B) = 1 or P(B/A) = P(B) (or P(A/B) = P(B)), then B and A are independent and there is no correlation between them.
- If *Lift(A, B) > 1 or P(B/A > P(B) (or P(A/B) > P(A))*, then *A* and *B* are positively correlated, meaning the occurrence of one implies the occurrence of the other.
- If *Lift(A, B) < 1 or P(B/A) < P(B) (or P(A/B) < P(B))*, then *A* and *B* are negatively correlated, meaning the occurrence of one discourage the occurrence of the other.

The Apriori algorithm was proposed in 1994 [10]. The algorithm identifies the frequent items in the database and extending them to larger and larger item sets as long as those itemsets appear adequately often in the database. The Apriori algorithm defines confidence level for association rules using two parameters; the minimum support threshold, and the minimum confidence threshold. The frequent itemsets defined by Apriori can be used to determine association rules which highlight general trends in the database [15].

Apriori algorithm uses a level-wise search, where k-itemsets (An itemset which contains k items is known as k-itemset) are used to explore (k + 1)-itemsets, to mine frequent itemsets from the transactional database for Boolean association rules. In this algorithm, frequent subsets are extended one item at a time, and

this step is known as the candidate generation process. Then groups of candidates are tested against the data [6].

To count candidate item sets efficiently, Apriori uses breadth-first search method and a hash tree structure. It identifies the frequent individual items in the database and extends them to larger and larger item sets as long as those itemsets appear sufficiently often in the database [6].

Apriori algorithm defines frequent item sets that can be used to determine association rules which highlight general trends in the database. Mining of association rules from a database consists of finding all rules that meet the user-specified threshold support and confidence. The problem of mining association rules can be as stated in the Algorithm below.

#### Algorithm:

1) Find all sets of items which arise with a frequency that is greater than or equal to the user-specified threshold support, *s*.

2) Generate the anticipated rules using the large itemsets, which have user-specified threshold confidence,  $\alpha$ .

Apriori algorithm starts by scanning the data set to find all the items and count their support as candidates of size1 (*i.e.*, C1) and removes infrequent items (count < min-supp).

## 4. The Enrollment Dataset

We classify the enrollment factors into four categories that contain 32 factors; **Table 1** below shows these categories and the factors constitute each category.

Based on categorization in **Table 1**, the enrollment data collected via a questionnaire that has been subject to a rigorous assessment by experts and professionals. The questionnaire has two sections; the first section is about demographic factors related to the enrolled student, namely; the university name the student enrolled in, the type of institution (governmental/private), gender (male/female), age (from 18 to 20), academic year class (1/2/3/4/5), and academic level (B.Sc./Diploma), and the second section is about the enrollment-related factors.

The sample of students is enrolled in eight Sudanese universities in Khartoum state in computer-related studies. From each of the 8 selected universities, 125 students are selected randomly from different classes. Then the 1000 sample has been reviewed and organized, making analysis easy, each type of data in the questionnaire is encoded.

The demographic information is encoded as follows; universities take codes  $\{1, 2, 3, 4, 5, 6, 7, 8\}$ , institution type as  $\{1, 2\}$ , gender is encoded as 1 or 2, age as  $\{1, 2, 3\}$ , academic class as  $\{1, 2, 3, 4, 5\}$ , and academic level as  $\{1, 2\}$ .

For the enrollment-related factor, the questionnaire uses the Likert fifth scale (Strongly agree, agree, Na, disagree, and strongly disagree), which are encoded as  $\{5, 4, 3, 2, 1\}$ .

The questionnaire data *n* transformed into tables as a data set. Statistical ad-

justments applied to data that requires scale transformations. Validation tests were used to evaluate the questionnaire scale and contents as follows:

- Cronbach alpha (*a*) was utilized for estimating the reliability coefficient for the questionnaire scale. Reliability coefficient means obtaining the same values when re-using the measuring instrument with the same circumstance, and give thus results [43]. Cronbach Alpha value for most factors is above 70%.
- Moreover, the percentage of the main categories of enrollment related factors is 92%. These show that the enrollment factors are the highly reliable and consistent measuring tool.

		Table 1.	Enrollment	factors.
Table 1. Enronment factors.	Table I. Enronment factors.			
Laure I. Emonnent lactors.	Laure 1. Emonment lactors.			
Laure I. Emonnent lactors.	Laure 1. Emonment lactors.			
Laure L. Linonnent lactors.	aute 1. Emonnent factors.			
Lable I. Linonment lactors.	Lable L. Lintonnicht lactors.			
Lubic L. Emonitement factors.	Lable I. Emonitement factors.			
Laore Li Lini onnient factoro.	Lucie II Elli onnient factoro.			

Category	No of factors	Factors Names
		• Aspiration,
		• Family/society motivation,
		<ul> <li>Parent's occupation/qualification,</li> </ul>
Student and Society	7	• Family income,
		• Family Social class,
		• Proudness/academic prestige,
		• Ethnic/religion
		• Location,
		• Reputation,
		• Education facilities,
		• International partnership,
		• Diversity of study,
		• Quality of education,
Educational Institution	n 14	• Image/campus,
Educational Institution		• Activity facilities,
		• Social life,
		• Postgraduate and institutional research,
		• Systematic and organized,
		• Academic assistance,
		• Modern/feasibility,
		• Campus visiting
		• Aptitude test,
		• Education/tuition cost,
		• Financial aid,
. 1	-	• Approved degree,
Admission	/	• Admission requirements,
		• Educational institution representative and
		advertising,
		• Retention
		• Good job opportunity,
		• High income,
Employment	4	• Good position,
		• Employed promotion and planning

• The Kaiser-Meyer-Olkin (KMO) and Bartlett's tests were used to check for adequacy of sample contents. KMO is statistic values that demonstrate the proportion of variance in the variables that might be affected by underlying factors [44]. KMO returns values between 0 and 1. Here in this sample test, KMO values are between (0.8 and 1), which indicates that the sampling is adequate, and a factor analysis may be useful with this dataset.

## 5. Applying Association Rule Mining to the Sudanese Universities Enrollment Dataset

To determine what enrollment factor category is the most influential, we apply the Priori association rule algorithm, with a suitable minimum support value, to find all the frequent itemsets in the dataset with that minimum support threshold, and then extract the association rules.

Factors within the extracted rules, define the set of most critical factors and thus the most influential category. Besides, the extracted rules determine the correlations between these factors.

Two types of relationships are of interest to this study, the first is the relationship between the demographic factors and the enrollment-related factors, and the second is the relationship between enrollment related factors themselves.

In extracting association rules, three steps have to be followed:

1) Determine the minimum support: A fundamental problem in Apriori is how to choose a minimum support (min-supp) value to find interesting patterns. There is not an easy way to determine the best min-supp threshold. In this paper, the minimum support is determined by trial and error.

2) Create the itemset list: In association rules, a set of the item also defined as a large item set. The item set is said to be frequent if it occurs more than the predefined minimum support.

3) Extract the association rules:

a) Generate the standard association rules from the frequent items by using the Apriori algorithm.

b) An association rule measures. The correlation rule is measured based on the minimum support, minimum confidence, and interest correlation between itemsets A and B by using lift measures.

### 5.1. Extraction of Association Rules between Demographic Factors and Enrollment Related Factors

To determine the suitable minimum support value, we try different values of minimum support and minimum confidence, and record the number of association rules that can be generated. Table 2 shows the number of association rules for the range of values 0.1 to 1 for minimum support and minimum confidence Since the number of generated rules is similar for support from 0.1 to 0.3 and confidence from 0.1 to 0.9, we select as appropriate values min-supp = 0.3 and min-conf = 0.9 as thresholds.

To create the candidate set, called large itemset L(1), the set is generated for demographic factors and enrollment related factors. The size of L1 is 19 items, as shown in **Table 3**, and the size of large itemsets L(2) is 14 items, as shown in **Table 4** below.

<b>0</b> /	Confidence									
Support	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.1	10	10	10	10	10	10	10	10	10	10
0.2	10	10	10	10	10	10	10	10	10	0
0.3	10	10	10	10	10	10	10	10	10	0
0.4	10	10	10	10	10	10	10	8	8	0
0.5	2	2	2	2	2	1	1	1	1	0
0.6	0	0	0	0	0	0	0	0	0	0
0.7	0	0	0	0	0	0	0	0	0	0
0.8	0	0	0	0	0	0	0	0	0	0
0.9	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0

 Table 2. The number of association rules extracted for support and confidence values (from 0.1 to 1).

**Table 3.** Large itemsets L(1).

No	Itemset	Support (%)
1	Academic level = 1	93.7
2	Gender = 1	54.4
3	Age = 2	52
4	Type of institution = 1	50.4
5	Type of institution = 2	48.9
6	Admission requirements = 4	46.5
7	Gender = 2	44.9
8	Family income = 4	43.7
9	Diversity of study = 4	42.8
10	Approved degree = 4	42.6
11	Age = 3	40.6
12	Aptitude test = 4	38.9
13	Aspiration = 5	37.9
14	Reputation = 4	37.7
15	Good job opportunity = 4	37.6
16	Approved degree = 5	37.4
17	Quality of education = 4	35.9
18	Postgraduate/institutional research = 4	35.3
19	Aspiration $= 4$	35.1

No	Itemset	Support (%)
1	{Gender = 1, Academic level = 1}	51.7
2	{Age = 2, Academic level = 1}	48.5
3	{Academic level = 1, Type of institution = 1}	47.5
4	{Academic level = 1, Type of institution = 2}	46.2
5	{Academic level = 1, Admission requirements = 4}	43.4
6	{Gender = 2, Academic level = 1}	42
7	{Academic level = 1, Family income = 4}	40.9
8	{Academic level = 1, Approved degree = 4}	40
9	{Academic level = 1, Diversity of study = 4}	39.6
10	{Age = 3, Academic level = 1}	39.1
11	{Academic level = 1, Aptitude test = 4}	36.1
12	{Academic level = 1, Approved degree = 5}	35.3
13	{Academic level = 1, Aspiration = 5}	35.2
14	{Academic level = 1, Reputation = 4}	35

Table 4. Large itemsets L(2).

The generated common Apriori association rules of demographic factors and enrollment related factors using min-supp is (0.3), min-conf is (0.9), and using lift measures are shown in Table 5.

**Table 5** shows 10 association rules. Six rules correlated demographic factors and four rules were correlated demographic factors with enrollment factors. Numbers within each rule shows the number of itemsets comes together, for example, the in rule 1: Age =  $3406 \Rightarrow$  Academic levels = 1391, 406 is the total number of records with Age = 3391 of them come with academic level = 1.

The column of Lift was used to measure the new rules and specified the type of association rules. Rule 1and rules 2 are higher than one; this means the factors are positively correlated. In positive correlation, both antecedent and consequent factors move in the same direction. While rules 3, 4, 5, and 6 are equal to one, this means they have an independent correlation. In independent correlation, the probability of occurrence of the antecedent and that of the consequent factors are independent of each other. While rules 7, 8, 9, and 10 are less than one, this means they negatively correlated with high confidences. A negative correlation is an inverse relationship between the two factors

## 5.2. Extraction of Association Rules between Enrollment Related Factors

To determining the minimum support value; association rules generated by trying different values of min-supp, and min-conf and the number of rules generated as shown in **Table 6**. The most appropriate values found are min-supp is (0.2), min-conf is (0.6).

No	Rules	Confidence	Lift
1	Age = 3406  → Academic level = 1391	0.96	1.02
2	Gender = 1544 ➔ Academic level = 1517	0.95	1.01
3	Type of institution = 2489  → Academic level = 1462	0.94	1
4	Approved degree = 5374  → Academic level = 1353	0.94	1
5	Type of institution = 1504  → Academic level = 1475	0.94	1
6	Approved degree = $4426 \Rightarrow$ Academic level = 1	0.94	1
7	Family income = $4437 \Rightarrow$ Academic level = $1409$	0.94	0.99
8	Gender = 2449 ➔ Academic level = 1420	0.94	0.99
9	Admission requirements = $4465 \Rightarrow$ Academic level = $1434$	0.93	0.99
10	Age = 2520 $\rightarrow$ Academic level = 1485	0.93	0.99

Table 5. Association rules of demographic factors and enrollment factors.

**Table 6.** The number of association rules extracted for support and confidence values (from 0.1 to 1).

Sunnart	Confidence									
Support -	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.1	10	10	10	10	10	10	10	10	0	0
0.2	10	10	10	10	10	8	0	0	0	0
0.3	0	0	0	0	0	0	0	0	0	0
0.4	0	0	0	0	0	0	0	0	0	0
0.5	0	0	0	0	0	0	0	0	0	0
0.6	0	0	0	0	0	0	0	0	0	0
0.7	0	0	0	0	0	0	0	0	0	0
0.8	0	0	0	0	0	0	0	0	0	0
0.9	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0

In the extract the association rules, the candidate set called large itemsets, (L) is generated of enrollment related factors. The size of L(1) is 73 items, and the size of large itemsets L(2) is 29 items.

The common Apriori association rules of enrollment related factors were extracted using min-supp = 0.2, min-conf = 0.6 and lift > 1, as shown in **Table 7**.

## 6. Results and Discussion

**Table 5** shows that there are only 7 factors that influence student's enrollments,4 of them related to the demographic factors, and 3 factors of them are enrollment related factors. These factors are shown in Table 8.

Table 7 shows that the Student and society related factors category did not

appear in all association rules, Educational Institution related factors appeared in five rules, Admission related factors appeared in four rules, and employment-related factors appeared in one rules. These factors are shown in **Table 9**.

In addition to that, we found out there is the ministry approves a positive influence from the two factors reputation of universities and education facilities with a degree. Besides, there is a positive correlation between factors reputation of universities and education facilities with a degree are approved by the ministry, a correlation between the factors; good position, retention, the degree is approved by the ministry, aptitude test and feasibility and the factor admission requirements.

Table 7. Association rules found between enrollments related factors.

No	Rules	Confidence	Lift
1	Reputation = 5340 ==> Approved degree = 5208	0.61	1.62
2	Quality of education = 4359 ==> Diversity of study = 4224	0.62	1.45
3	Education facilities = 4337 ==> Diversity of study = 4207	0.61	1.43
4	Good position = 4314 ==> Ministry requirements = 4200	0.64	1.36
5	Retention = 4341 ==> Ministry requirements = 4213	0.62	1.34
6	Approved Degree = 4426 ==> Ministry requirements = 4267	0.63	1.33
7	Aptitude test = 4389 ==> Ministry requirements = 4240	0.62	1.32
8	Modern /feasibility = 4341 ==> Ministry requirements = 4208	0.61	1.3

 Table 8. The demographic and enrollment related factors influencing student's decision to enrollment.

Category	Demographic Factors	Students and Society-Related Factors	Educational Institution Related factors	Admission- Related factors	Employment- Related Factors
Factors	Age, Gender, Type of institution, Academic level	Family income		Approved degree, Admission requirements	
Total	4	1	0	2	0

Table 9. The enrollment related factors influencing student's decision.

Categorie	Students and <sup>s</sup> Society-Related Factors	Educational Institution Related factors	Admission Related Factors	l Employment- Related Factors
		Reputation	Approved degree	
		Diversity of Study	Admission requirements	
Factors		Quality of education	Retention	Good position
		Education facilities	A	
		Modern/feasibility	Aptitude test	
Total	0	5	4	1

## 7. Conclusions

This paper shows that only 4 out of the 6 student's demographic factors, namely: age, gender, type of institution, and academic level have a strong influence on student's enrollment. Moreover, the most important categories of factors that affect the Sudanese student's decision to enroll in Higher Education Institutions are the Educational Institution and Admission categories of enrollment related factors.

The findings can be used by higher education institutions as a guideline in offering students the appropriate knowledge about enrollment related factors.

## **Conflicts of Interest**

The authors declare no conflicts of interest regarding the publication of this paper.

#### References

- [1] Upcraft, M.L. and John N. (1989) The Freshman Year Experience: Helping Students Survive and Succeed in College. Jossey-Bass Publishers, San Francisco, CA.
- [2] Han, J., Pei, J. and Kamber, M. (2012) Data Mining: Concepts and Techniques. 3rd Edition, Elsevier, Amsterdam.
- [3] Romero, C., Ventura, S. and García, E. (2008) Data Mining in Course Management Systems: Moodle Case Study and Tutorial. *Computers & Education*, 51, 368-384. <u>https://doi.org/10.1016/j.compedu.2007.05.016</u>
- [4] Abdullah, Z., Herawan, T., Ahmad, N. and Deris, M.M. (2011) Mining Significant Association Rules from Educational Data Using Critical Relative Support Approach. *Procedia-Social and Behavioral Sciences*, 28, 97-101. https://doi.org/10.1016/j.sbspro.2011.11.020
- [5] Agrawal, R., Mannila, H., Srikant, R., Toivonen, H. and Verkamo, A.I (1996) Fast Discovery of Association Rules. *Advances in Knowledge Discovery and Data Mining*, 12, 307-328.
- [6] Chung, W., Chen, H. and Reid, E. (2009) Business Stakeholder Analyzer: An Experiment of Classifying Stakeholders on the Web. *Journal of the American Society for Information Science and Technology*, **60**, 59-74. <u>https://doi.org/10.1002/asi.20948</u>
- [7] Kumbhare, T.A. and Chobe, S. (2014) An Overview of Association Rule Mining Algorithms. *International Journal of Computer Science and Information Technologies*, 5, 927-930.
- [8] Dunham, M. (2002) Data Mining: Introductory and Advanced Topics. Pearson Education, London.
- [9] Yahya, O., Hegazy, O. and Ezat, E. (2012) An Efficient Implementation of A-Priori algorithm Based on Hadoop-Mapreduce Model. *International Journal of Reviews in Computing*, 12, 59-67.
- [10] Agrawal, R. and Srikant, R. (1994) Fast Algorithms for Mining Association Rules in Large Databases. *Proceedings of the* 20th *International Conference on Very Large Data Bases*, Santiago de Chile, 12-15 September 1994, 487-499.
- [11] Light, A. and Strayer, W. (2000) Determinants of College Completion: School Quality or Student Ability? *Journal of Human Resources*, **35**, 299-332.

https://doi.org/10.2307/146327

- [12] Mehboob, F., Shah, S. and Bhutto, N. (2012) Factors Influencing Student's Enrollment Decisions in Selection of Higher Education Institutions (HEI's). *Interdisciplinary Journal of Contemporary Research in Business*, 4, 558-568.
- [13] Akerhielm, K., Berger, J., Hooker, M. and Wise, D. (1998) Factors Related to College Enrollment. Final Report.
- [14] Esters, L. (2007) FactorsIinfluencing Post Secondary Education Enrollment Behaviors of Urban Agricultural Education Students. *Career and Technical Education Research*, **32**, 79-98. <u>https://doi.org/10.5328/CTER32.2.79</u>
- [15] Braxton, J.M., Hirschy, A.S. and McClendon, S.A. (2011) Understanding and Reducing College Student Departure: ASHE-ERIC Higher Education Report. Jossey-Bass, San Francisco, CA.
- [16] Smyth, F.L. and McArdle, J.J. (2004) Ethnic and Gender Differences in Science Graduation at Selective Colleges with Implications for Admission Policy and College Choice. *Research in Higher Education*, **45**, 353-381. https://doi.org/10.1023/B:RIHE.0000027391.05986.79
- [17] George-Jackson, C. (2013) High School Students Who Intend to Major in Math: Differences by Gender, Race, Ethnicity, and Confidence Levels. Math Alliance Research Study.
- [18] Sidin, S.M., Hussin, SR and Soon, TH. (2003) An Exploratory Study of Factors Influencing the College Choice Decision of Undergraduate Students in Malaysia. *Asia Pacific Management Review*, 8, 259-280.
- [19] Jain, V. (2014) The Impact of Choice Factors Affecting Student Selection of Higher Education Institutions: An Empirical Study. *Management Research*, 1, 55-67.
- [20] Ming, J.S.K. (2010) Institutional Factors Influencing Students' College Choice Decision in Malaysia: A Conceptual Framework. *International Journal of Business and Social Science*, 1, 53-58.
- [21] Altbach, P.G., Reisberg, L. and Rumbley, L.E. (2010) Tracking a Global Academic Revolution. *Change: The Magazine of Higher Learning*, **42**, 30-39. <u>https://doi.org/10.1080/00091381003590845</u>
- [22] Teixeira, P.N., Rocha, V., Biscaia, R. and Cardoso, M.F. (2012) Competition and Diversity in Higher Education: An Empirical Approach to Specialization Patterns of Portuguese Institutions. *Higher Education*, **63**, 337-352. https://doi.org/10.1007/s10734-011-9444-9
- [23] Gray, R., Vitak, J., Easton, E.W. and Ellison, N.B. (2013) Examining Social Adjustment to College in the Age of Social Media: Factors Influencing Successful Transitions and Persistence. *Computers & Education*, 67, 193-207. https://doi.org/10.1016/j.compedu.2013.02.021
- [24] Kusumawati, A., Yanamandram, V.K. and Perera, N. (2010) University Marketing and Consumer Behavior Concerns: The Shifting Preference of University Selection Criteria in Indonesia. 2010 Asian Studies Association of Australia 18th Biennial Conference, Adelaide, South Australia, 5-8 July 2010, 1-16.
- [25] Maringe, F. and Gibbs, P. (2008) Marketing Higher Education. Theory and Practice. McGraw-Hill Education, New York.
- [26] Ruswahida, R., Ku Halim, A., Aminul, I., Noor, I. and Zaidi, M. (2014) Determinants Students' Selection of Higher Education Institutions in Malaysia. *Advances in Environmental Biology*, 8, 406-416.

- [27] Levitz, N. (2012) Why Did They Enroll? The Factors Influencing College Choice. <u>https://docplayer.net/9559359-Why-did-they-enroll-the-factors-influencing-college</u> <u>-choice.html</u>
- [28] Stewart, S., Lim, D.H. and Kim, J. (2015) Factors Influencing College Persistence for First-Time Students. *Journal of Developmental Education*, 38, 12-20.
- [29] Robb, C.A., Moody, B. and Abdel-Ghany, M. (2012) College Student Persistence to Degree: The Burden of Debt. *Journal of College Student Retention: Research, Theory & Practice*, 13, 431-456. https://doi.org/10.2190/CS.13.4.b
- [30] Abdul Fattah, M., Mohamed, F., Philp, S. and Tarek, F. (2013) Discovery of Association Rules from University Admission System Data. *International Journal of Mod*ern Education and Computer Science, 4, 1-7. <u>https://doi.org/10.5815/ijmecs.2013.04.01</u>
- [31] Encheva, S. and Tumin, S. (2006) Application of Association Rules in Education. In: Huang, D.S., Li, K. and Irwin, G.W., Eds., *Intelligent Control and Automation. Lecture Notes in Control and Information Sciences*, Springer, Berlin, Heidelberg, 834-838. <u>https://doi.org/10.1007/978-3-540-37256-1\_105</u>
- [32] Feng, S., Zhou, S. and Liu, Y. (2011) Research on Data Mining in University Admissions Decision-Making. *International Journal of Advancements in Computing Technology*, 6, 176-186.
- [33] Cantabella, M., Martínez-España, R., Ayuso, B., Yáñez, J.A. and Muñoz, A. (2019) Analysis of Student Behavior in Learning Management Systems through a Big Data Framework. *Future Generation Computer Systems*, 90, 262-272. <u>https://doi.org/10.1016/j.future.2018.08.003</u>
- [34] Adekitan, A.I. and Noma-Osaghae, E. (2019) Data Mining Approach to Predicting the Performance of First Year Student in a University Using the Admission Requirements. *Education and Information Technologies*, 24, 1527-1543. https://doi.org/10.1007/s10639-018-9839-7
- [35] Cortes, K.E. and Lincove, J.A. (2019) Match or Mismatch? Automatic Admissions and College Preferences of Low-and High-Income Student. *Educational Evaluation* and Policy Analysis, 41, 98-123. <u>https://doi.org/10.3102/0162373718813360</u>
- [36] Wu, H., Enhe, M. and Wang, J. (2018). Research on Application of Data Mining Based on Improved APRIORI Algorithm in Enrollment Management in Colleges and Universities. In: Hung, J., Yen, N. and Hui, L., Eds., *Frontier Computing. FC* 2018. *Lecture Notes in Electrical Engineering*, Springer, Singapore, 1077-1082. https://doi.org/10.1007/978-981-13-3648-5\_136
- [37] Helal, S., Li, J., Liu, L., Ebrahimie, E., Dawson, S. and Murray, D.J. (2019) Identifying Key Factors of Student Academic Performance by Subgroup Discovery. *International Journal of Data Science and Analytics*, 7, 227-245. https://doi.org/10.1007/s41060-018-0141-y
- [38] Merceron, A. and Yacef, K (2008) Interestingness Measures for Association Rules in Educational Data. *Proceedings of the 1st International Conference on Educational Data Mining*, Montreal, 20-21 June 2008, 57-66.
- [39] Baradwaj, B.K. and Pal, S. (2011) Mining Educational Data to Analyze Students' Performance. *International Journal of Advanced Computer Science and Applications*, 2, 63-69. <u>https://doi.org/10.14569/IJACSA.2011.020609</u>
- [40] Maltese, A.V. and Tai, R.H. (2011) Pipeline Persistence: Examining the Association of Educational Experiences with Earned Degrees in STEM among US Students. *Science Education*, 95, 877-907. <u>https://doi.org/10.1002/sce.20441</u>

- [41] Klemettinen, M., Mannila, H., Ronkainen, P., Toivonen, H. and Verkamo, A.I. (1994) Finding Interesting Rules from Large Sets of Discovered Association Rules. In: Adam, N.R., Bhargava, B.K. and Yesha, Y., Eds., *Proceedings of the Third International Conference on Information and Knowledge Management*, ACM, New York, 401-407. <u>https://doi.org/10.1145/191246.191314</u>
- [42] Hussein, N., Alashqur, A. and Sowan, B. (2015) Using the Interestingness Measure Lift to Generate Association Rules. *Journal of Advanced Computer Science & Technology*, 4, 156. <u>https://doi.org/10.14419/jacst.v4i1.4398</u>
- [43] Golafshani, N. (2003) Understanding Reliability and Validity in Qualitative Research. *The Qualitative Report*, **8**, 597-606.
- [44] Worthington, R.L., and Whittaker, T.A. (2006) Scale Development Research: A Content Analysis and Recommendations for Best Practices. *The Counseling Psychologist*, 34, 806-838. <u>https://doi.org/10.1177/0011000006288127</u>