كلية الدراسات العليا


Sudan University of Science and Technology

College of Graduate Studies


# Improving Students Academic Performance Using Hybrid Recommendation Techniques

# تحسين أداء الطلاب الأكاديمي باستخدام طرق توصية هجين

A dissertation Submitted in Partial Fulfillment of the Requirements

For MSc Degree in Information Technology

**By**


Mohamed Osman Bashar


**Supervised by**

**Dr.** Wafaa Faisal Mukhtar

November 2018

# Introductive

قال الله تعالى:

{فَتَعَٰلَى ٱللَّهُ ٱلۡمَلِكُ ٱلۡحَقُّۗ وَلَا تَعۡجَلۡ بِٱلۡقُرۡءَانِ مِن قَبۡلِ أَن يُقۡضَىٰٓ إِلَيۡكَ وَحۡيُهُۥۖ وَقُل رَّبِّ زِدۡنِي عِلۡمٗا ١١٤}[1]

(1) سورة طه: [114]

# Dedication

To my:

-Parents

-Brothers

- Sisters

- Scientists

- Relatives

-Colleagues

- Friends

With great respect and gratitude

I offer this effort

Mohammed

## Acknowledgement

Praise be to the Allah, the All Mighty, who granted me health and power to accomplish this work. It gives me great pleasure to express my deep thanks, sincere gratitude and appreciation to my supervisor: *Dr. Wafaa Faisal Mukhtar,* I would like to express appreciation her guidance for this work. I am highly indebted to my parents who taught me the right things, encouraged and gave me the hope and unconditional love. I would like to express my sincere thanks to all the staff of faculty of Computer Science and Information Technology, Sudan University of Science and Technology , for their valuable help, encouragement, follow up and assistance. Also the kind devoted help of those who made the study possible, special and deep thanks are extended to Dr. Mohammed Hamoda, Mohammed Osman Hassan, Hassan Elnoor, *Ustaz Mohammed Idris, Mohammed Adam and Muntasir Ahmed for* their kind help, assistance and support me. Finally my deepest thanks are extended to my colleagues of the batch (8) in Master degree, and all the other individuals who helped me during the study period.

## Abstract

Development of students' performance is significant in educational environments because it plays an essential role in making the best quality graduates and post-graduates who will become great leaders in the future and sources of the workforce for the country. A recommendation system is an intelligent system that proposes different suggestions to students, based on the previous actions from other students who faces the same environment, such as academic performance. One of the major problems today the high rate of students failure is a worry for many universities. This study proposed recommendation system to identify weak academic students as soon as possible to help them in a suitable time, encourage students to study hard when they know that they are at risk and to plan their workload carefully. The study is applied the hybrid recommendation system that is one approach for recommendation system. This approach is executed by used each of clustering algorithms and association rules algorithms on the nature of data which have been collected from the University of Kordofan, Faculty of Computer Studies and Statistics. The clustering algorithms results were evaluated regarding to high accuracy for each cluster and then applied Association rules algorithms in particular. The obtained results are generated strong rules that appear which courses are effectiveness posative or nagative on accumulative GPA.

## المستخلص

تطوير اداء الطلاب امر مهم في البيئة التعليمية لأنه يعلب دور اساسي في صنع خريجين افضل نوعية من حملة البكالريوس و الدراسات العليا الذين يعتبروا مستقبل و مصادر القوة البشريةِ للبلادِ. نظام التوصية من الانظمة الذكية التي تقترح مقترحات مختلفة للطلاب ، بناءاً علي احداث سابقة لطلاب اخرين يواجهوا نفس البيئة مثل : المستوى الاكاديمي. واحدة من المشاكل الرئيسة التي احدثت قلق لكثير من الجامعات هي النسبة العالية لإستبقاء الطلاب. نظام التوصية المقترح سوف يعرف الطلاب الضعفين اكاديميا بأسرع ما يمكن و مساعدتهم في الوقت المناسب ، تشجيع الطلاب للدراسة بشدة عندما يعلموا بأنهم في مخاطرة و يخططوا لواجباتهم باهتمام. هذه الدراسة طبقت طرق هجين كطريقة من نظام التوصية. هذه الطريقة طبقت باستخدام كل من الخوارزميات العنقودية و خوارزميات قواعد الارتباط علي طبيعة البيانات التي جمعت من كلية دراسات الحاسوب و الاحصاء بجامعة كردفان. نتائج الخوارزميات العنقودية التي تم تقييمها بشأن الدقة العالية لكل عنقود ومن ثم طبق عليهاخوارزميات قواعد الارتباط علي وجه الخصوص. النتائج المتحصلة من الدراسة ولدت قواعد قوية توضح المواد التي لها تأثير ايجابي و سلبي علي المعدل التراكمي.

## Table of Contents

## List of tables

**Table of figures**

.

## List of Abbreviations

| | |
|---|---|
| GPA | Grade Point Average |
| Cumulative GPA | Cumulative Grade Point Average |
| WEKA | Waikato Environment for Knowledge Analysis |
| CSV | Comma-separated values (Comma-delimited) |
| RS | Recommendation Systems |

# CHAPTER ONE

# INTRODUCTION

## 1.1 Research background

Education is an essential element for the betterment and progress of a country. It enables the people of a country civilized and well mannered. Today the important challenge that higher education faces is reaching a stage to facilitate the universities for having more efficient, effective and accurate educational processes. To date, higher educational administrations are placed in a very highly competitive environment and are targeted to get more competitive advantages over the other competitors. To remain competitiveness among educational field, these administrations need deep and enough knowledge for a better assessment, evaluation, planning, and decision-making (Bambrah *et al.*, 2014). The required knowledge cannot be gained from the tailor-made software used nowadays. Major problem that the academic institutions are facing worldwide is poor performance of students in academics. Which causes high attrition rate and resulted in a loss to students, parents and institutions. Student's performance prediction can help to reduce attrition rate as it raises an early alarm for students not performing well and are likely to leave the institution. In accordance with this institution of thought, though studies have been carried out at one time or the other showing that various predictors at various time and different locations contribute to the outcome of students. There exist some evidence of students' background information that contribute immensely to the early prediction of student success. Though none of the studies directly shows how family background factors relate to student performance, it is necessary to construct a model to capture students' success at the first-year level. Moreover, most researchers using recommender systems are doing this only for a particular course however, Giving a good prediction for student academic performances for a particular course is no guarantee that it will give a good prediction for another course. Therefore it is necessary to design a recommender that can give a global prediction at the end of each student session. This recommender system not only helps to predict the grades of students, it also helps to recommend the courses to the students by considering their timetable. Educational Recommender systems can use data mining techniques such as classification techniques, clustering, generation of association rules, to get recommendation rules from huge amount of data to create their recommendations using information learned from the academic actions and attributes of learners and learning system.

### 1.2 Research problem

Most university students obtain low marks in their courses and sometimes fail in the examinations which result getting poor GPA. Due to insufficient experiences of the student to enhance the academic performance and lack of intellectual abilities required for success in each course. There is a lack of faculty academic guidance that implements modern approaches to assist the students in promoting their academic performances.

### 1.3 Objectives

The main objective of this research is to propose an approach based on the recommendation systems techniques to improve the students' results through the following points:

- Collecting dataset of the students.
- Preprocessing and analysis the datasets.
- Extracting recommendation rules that informing the students before joint each year study to know which courses are affecting negatively on their GPA.

### 1.4 Methodology

The method suggested in this research to improve students' academic performance is belongs to the architecture of recommendation systems using Data Mining technique. There are four main stages in this method, Data collection; preprocessing, clustering based collaborative filtering, and applying association rules. Data collection is gathering all information available on students considering factors affect student performance. Pre-processing data is a necessary step for preparing the dataset before applying clustering based collaborative filtering. Clustering based collaborative system is the process of grouping the data into classes or clusters. Then generate association rules based on the frequent item set.

### 1.5 Scope of the research

This research is focusing on generating strong rules for recommendation system of dataset collected at the Department of Information Technology, Faculty of Computer Studies and Statistics, at the University of KORDOFAN.

### 1.6 Thesis organization

This research has five chapters organized as follows:
Chapter I contains introduction. Chapter II discusses the literature review and related work. Chapter III describes the research methodology and the implementation of the techniques used. Chapter IV presents the results and their discussion. Lastly, Chapter V concluded and presents the Future work.

# CHAPTER TWO

# LITERATURE REVIEW

## 2.1 Introduction

This chapter mainly describes the state of the art articles in analyzing the student performance using different techniques. Relevant information sources and related publications are mentioned. The first part presents data mining techniques and especially that was implemented in the educational section. The second part presents the theory about recommender systems in order to familiarize the readers with the techniques utilized in the area. The third part presents the related work which applies recommendation system approaches in the educational section**.** The last summary of the literature review is mentioned in the educational section.

## 2.2 Data mining

Data mining is a logical process that is used to search through a large amount of data in order to find useful knowledge. Data mining goal is to find patterns that were previously unknown. These patterns are further used to make certain decisions for the development of businesses. Many people treat data mining as a synonym for another popularly used term, Knowledge Discovery from Data, or KDD. Knowledge discovery as a process is depicted in Figure 2-1 (Jiawei Han, 2006).

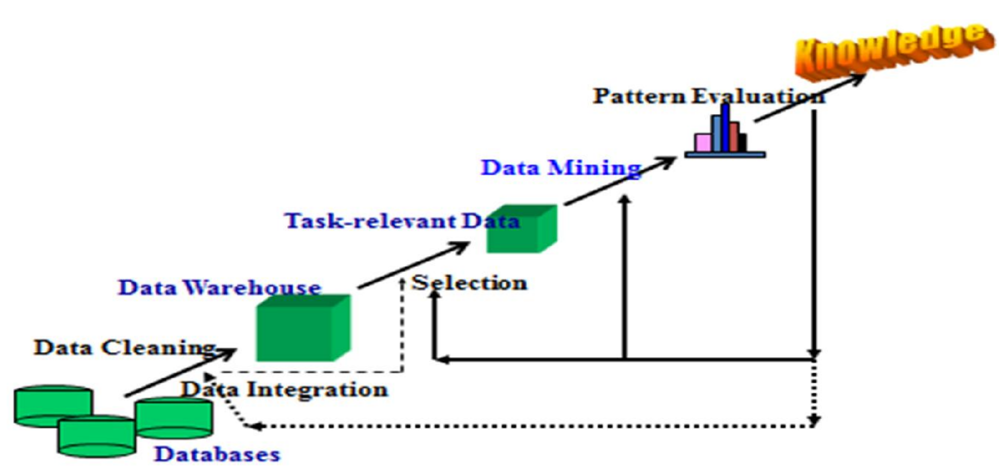



Figure 2-1: Data mining as a step in the process of knowledge discovery

Data mining techniques can be classified broadly as predictive and descriptive. Predictive techniques are Classification, Regression, Time-series Analysis, and Prediction.

Classification is a data mining task that maps the data into predefined groups & classes (Aher and L.M.R.J., 2012). Classification algorithms are especially focusing on four rule induction algorithms One R, Zero R, JRIP and PART, and four decision algorithms J48, Random tree, REP tree and Decision stump (Mobasher, Shawish and Ibrahim, 2017).

The regression using known data formats like linear or logistic and assume the future data format will fall into the data structure. It then tries to predict the value by applying some mathematical algorithms on the data set.

With time series analysis, every attribute value determines by the different time interval (Al-Badarenah and Alsakran, 2016). Prediction is related with time series but not time bound. It is used to predict value based on past data and current data. Descriptive techniques are clustering, summarization, association rule algorithm and sequence discovery. Clustering is finding groups of objects such that the objects in one group are dissimilarity from another objects group (Al-Badarenah and Alsakran, 2016). Summarization is associating the sample subset with a small description.

Association rules are used to show the relationship between data items (Aher and L.M.R.J., 2012). Sequence discovery is finding a sequence of an activity. Such as in a shop, people may often buy toothpaste after toothbrush. It is all about what sequence user buying the product and based on the shop owner can arrange the items nearby each other.

### 2.3 Data mining in an educational section

Educational institutes are gaining popularity from the availabilities of potentialities in areas of research. Data mining can be used in the educational field to enhance the understanding of the learning process to focus on identifying, extracting and evaluating variables related to the learning process of students. Some researchers used data mining in educational data to improve the learning process and enhance the academic performance of the students as illustrated the flowing points.

Many classifications techniques were applied to predict the final grade. (Sumitha and Vinothkumar, 2016), (Al-barrak and Al-razgan, 2016) applied decision tree J48. (Raut, 2017) was used ID3 and decision tree C4.5. (Badr, Din and Elaraby, 2014), (Priya, 2013), (Badr, Din and Elaraby, 2014) applied classifier decision tree ID3. While (Yadav, 2012) compared between ID3, decision tree C4.5, and decision tree CART. The outcome of his results indicated that the C4.5 and ID3 had a better prediction than a CART. ID3 was also used in (Yadav, 2012) to predict if the student enrolling in the specific course will continue or not. He compared different types of decision trees; ID3, C4.5 and ADT. The outcome of their results, the ADT was effective predictive than C4.5 and ID3.

Also, other researchers compared between neural networks and different classifier algorithms. (Osofisan, Adeyemo and Oluwasusi, 2014) compared between neural networks and decision tree. Their result indicated that the prediction analysis of neural network was clearly outperforming decision tree, hence the generated rule from decision tree was more understandable than in the neural network. (Osmanbegović and Suljić, 2012), (Mueen, 2016) compared between neural networks and Bayesian classifier, decision trees. Naïve Bayes classifier was more accurate in prediction than decision tree and neural network.

(Gadhavi and Patel, 2017) predict the final grade of student's using linear regression. The model was applied in the prediction of the final grade of student's. They obtained result from the model would help a student in knowing final grade in a particular subject. Moreover (Kadambande *et al.*, 2017) applied support vector machine to predict the final grade of the student, the size of data of 181 records. Their obtained results generated rules to help both the lower students as well as the topper students.

(R.B, 2013) used k-means cluster and divided the data of the students into three classes. Firstly; high class which contains 13 students. Secondly; medium class which contains 87 students. Thirdly; Low class which contains 20 students. The obtained result from the cluster will help the teacher to reduce the drop-out ratio to a significant level and improve the performance of students.

### 2.4 Recommendation system

Recommender systems (RS) are a type of information filtering system that gives advice on products, information, or services that a user may be interested in. RS assist users with the decision making process when choosing items with multiple alternatives. Recommender systems are popular due to their e-commerce application purposes. In the e-commerce world, recommendations are increasing business by providing aid to customers and help them find what they may be looking for. In addition, it can be used as a tool to predict the user's behavior.

There are two basic entities that appear in any recommender system; the user and the item of interest. The user can be a customer in an e-commerce platform or a book reader looking for a recommendation for the next book they should read. The users provide their ratings on items and are used to aid other users with their recommendations. The item is the second piece of a recommender system. Users give items ratings and the algorithm outputs recommended items based on new user queries.

The widely utilized recommendation techniques can be divided into four approaches: Collaborative filtering recommendation, Content-based recommendation, Knowledge-based recommendation and hybrid systems (Aher and L.M.R.J., 2012).

#### 2.4.1 Collaborative filtering recommendation

Collaborative Filtering is the process of filtering or evaluating items using the opinions of other people. This filtering is done by using profiles. Collaborative filtering techniques collect and establish profiles, and determine the relationships among the data according to similarity models. Collaborative filtering algorithms can be divided into two main categories Memory based (user-based) and Model-based (item-based) algorithms (Hameed, 2012).

##### 2.4.1.1 Memory based recommendation

In memory-based learning, users are divided into groups based on their interest. When a new user comes into the system well be determine neighbors of users to make predictions for him. Memory-based recommendation uses entire or sample of the user-item database to make predictions (Hameed, 2012).

##### 2.4.1.2 Model-based collaborative filtering

Model-based collaborative filtering is a two-stage process for recommendations. In the first stage, the model is learned offline. In the second stage, a recommendation is generated for a new user based on the learned model. Model-based collaborative filtering can be represented as a Bayesian Belief network collaborative filtering and Simple naïve Bayesian collaborative filtering (Hameed, 2012) .

Bayesian Belief network collaborative filtering is a directed acyclic graph (DAG).

Naive Bayes strategy is used to make predictions for simple Bayesian collaborative filtering algorithms. In Naive Bayes classifier, we assume the features are independent of a given class. The probability is calculated by taking all features; the class with the highest probability will be classified as the predicted class.

In clustering collaborative filtering algorithms, a cluster is a collection of data objects that have high intraclass similarity and low interclass similarity. The similarity is measured using Minkowski distance and Pearson correlation for two data object *X=(x1, x2, x3……x n)* and *Y= (y1, y2, y3…..y n*) Minkowski distance is defined as equation (1).

$$d(X,Y)=\sqrt[q]{\sum_{i=1}^{n} | x_i - y_i |^q}$$

Equation (1)

Where $n$ is the dimension number of objects and $x_i$, $y\ i$ are the values of $i$ dimension of object $x$ and $y$ respectively. Is a positive integer If $q$=1, then d is called as manhattans distance If $q$=2, then d is called as the Euclidian distance (Hameed, 2012).

### 2.4.2 Content-based recommendation

In content-based recommender systems, a user marks some interesting items and the system offers the most similar items to the user's favorite items. These Systems need a lot of details about items in the database to be able to recommend to the similar users (Pazzani and Billsus, 2007). Thus content-based filtering is to compare the content consumed items up to that time. A list of new items that can be potentially recommended to the user to find items that are already similar (Felfernig *et al.*, 2014).

### 2.4.3 Knowledge-based recommendation

Knowledge-based recommendation attempts to suggest objects based on logical inferences.

A knowledge-based recommendation when compared to the approaches of collaborative filtering and content-based filter does not primarily depend on item ratings and textual item descriptions but on deep knowledge about the offered items (Felfernig *et al.*, 2014).

### 2.4.4 Hybrid systems

Hybrid recommender systems combine two or more recommendation techniques to gain better performance with fewer of the drawbacks of any distinct technique (Felfernig *et al.*, 2014).

## 2.5 Recommendation system in the educational section

Recommender systems (RS) can be used in different fields including the educational environment. RS is mainly focused on providing an educational section and tries to enhance the process of teaching and learning. Nowadays, researchers also try to improve the learning process and enhance the academic performance of the students as the flowing points.

(Bydžovská, 2013) used 67 temples on 138 courses from the Faculty of Informatics, Masaryk University between the years of 2010 and 2013. They proposed recommender system was conducted to predict the final grade of student's using classification and regression algorithms. Then were used collaborative filtering to recommender students. The outcome of their results designed a model for recommending students in the courses enrolling. They are presented final by green color was presented excellent, very good was presented by yellow color, good was presented by red color.

## 2.6 Recommendation with data mining in an educational section

There are some researchers used data mining with the recommendation system in educational data. To improve the learning process and enhance the academic performance of the students as the flowing points**.**

(Kumar and Padmapriya, 2014) applied Fuzzy C-Means clustering to divided data of students according to their marks. Then were applied C4.5 classifications on clusters data. The outcome of their results generated rules to recommender students when enrolling in the specific course is suited or not.

(Vialardi *et al.*, 2009) were applied decision tree c4.5 and collaborative filtering. The outcome of their results generated rules to recommender students for taking a decision on their academic program. While (Aher and L.M.R.J., 2012) were applied a combination of data mining algorithms such as decision Tree in Classification with association rule algorithm and K-means in clustering with Apriori association rule algorithm. Also, they have combined the Clustering with Classification algorithm into Association Rule algorithm and Classification with clustering algorithm into Association Rule algorithm. The outcome of their results indicated that the clustering combined with classification into association rule algorithm had a better combination for recommending the courses of students in E-learning.

(Priya, 2013) were collected data with a size of 2500 student. They are applied classifiers OneR, ZeroR and random tree to build an intelligent recommender system for predicting the performance of the students. The outcome of their results indicated that the random tree was high accuracy comparing to other classifiers OneR and ZeroR.

(Mobasher, Shawish and Ibrahim, 2017) collected data from three major factors demographic data educationally related attributes and psychological characteristics with a size of 200 records. There are applied eight classification algorithms especially focusing on four rule induction algorithms One R, Zero R, JRIP, and PART. Four decision algorithms J48, Random tree, REP tree and Decision stump. For recommending the students are improving their academic performance. The outcome of their results indicated that the classifier REP Tree was better prediction accuracy than other classifiers.

(Thangavel and Learn, 2017) collected data with a size of 2205 tuples for training dataset and testing dataset had a size of 289 tuples. They are applied Logistic Regression, Classification classifier, decision tree and Metbagging Classifier and Naïve Bayes to proposed recommendation system was conducted to predict the placement status of the students. The outcome of their results proved that the decision tree classifier stands out with 0.01 seconds of running time and 84.42% accuracy comparing to another classifier.

### 2.7 Summary of Literatures

Table 2-1: Summary of Literatures

| **Author** | **Techniques** | **Details of the data** | **Results** |
|---|---|---|---|
| reviewed papers of data mining in an educational section | | | |
| (Osofisan, Adeyemo and Oluwasusi, 2014) | neural network decision tree | none | prediction |
| (Al-barrak and Al-razgan, 2016) | decision tree J48 | Size of data 236 record | prediction |
| (Badr, Din and Elaraby, 2014) | decision tree ID3 | Size of data 1548 record | prediction |
| (El-Halees, 2008) | Apriori rules decision tree EM-clustered | Size of data 151 record | prediction |
| (R.B, 2013) | k-means cluster | Size of data 87 record | divided data into three class |

| | | | |
|---|---|---|---|
| (Yadav, 2012) | decision tree ID3 decision tree C4.5 | Size of data 432 record | prediction |
| (Gadhavi and Patel, 2017) | linear regression | Size of data 181 record | prediction |
| (Gamulin, Gamulin and Kermek, 2014) | Neural Networks Random Forests Support Vector Machine K-nearest neighbor | Size of data 232 record | prediction |
| (Sorour *et al.*, 2014) | Neural Network Classifier k-means cluster | Size of data 1400 record | divided data into 5 clusters Rule to prediction |
| (Yadav, 2012) | decision tree ID3, C4.5 and CART | Size of data 90 record | prediction |
| (Kadambande,Thakur, Mohol, & Ingole, 2017) | support vector machine | none | prediction |
| (Raut, 2017) | decision tree C4.5 | none | prediction |
| (Mueen, 2016) | Naïve Bayes Neural Network Decision Tree | Size of data 60 record | prediction |
| (Sumitha and Vinothkumar, 2016) | decision tree J48 | Size of data 250 for training data And 50 record for testing data | prediction |
| reviewed papers of the recommendation system in educational section | | | |
| (Bydžovská, 2013) | classification and regression algorithms | Size of data 67 temples on 138 courses | recommendation |
| Recommendation with data mining in **an** educational section | | | |
| (Kumar and Padmapriya, 2014) | Fuzzy C-Means decision tree C4.5 | none | recommendation |
| (Vialardi *et al.*, 2009) | decision tree c4.5 | Size of data 100274 record | recommendation |
| (Aher and L.M.R.J., 2012) | Classification, K-means clustering Apriori association rule | Their data content 13-course category about 82 subjects | recommendation |
| (Priya, 2013) | Classifiers OneR and ZeroR | Size of data 2500 record | recommendation |
| (Mobasher, Shawish and | Classifiers J48, | Size of data 200 | recommendation |

| | | | |
|---|---|---|---|
| Ibrahim, 2017) | Random tree, REP tree and Decision stump, One R, Zero R, JRIP and PART | record | |
| (Thangavel and Learn, 2017) | Logistic Regression Classification classifier decision tree Metbagging Classifier Naïve Bayes | Size of data 2205 record for training data and 289 for testing data. | Divided data into 5 class values and rule for recommendation |

## 2.8 Summary

The recommendation system is widely used in E-learning systems. E-learning systems used for recommending such as new books, optional courses and mandatory and learning objects. In this study reviewed the existing surveys and journal papers about educational recommendation system, educational data mining as well as recommendation system based on data mining techniques. Therefore depending on reviewed papers mentioned. This research adopts methodology according to recommendation system based on data mining techniques which is explaining broadly in next chapter.

# CHAPTER THREE
# METHODOLOGY

## 3.1 Introduction

This chapter will present the structured methodology used in this research. The methodology was carried on four phases in order to achieve the objective of this research. The first phase explains the data preprocessing techniques applied to the dataset under examination. The second phase reconsiders the algorithm implemented by the researcher based on recommendation system as discussed in previous chapter. The third phase goes through the evaluation clusters and last phase implements the association rule algorithm. Figure 3-1 explains the four steps to implement the mentioned methodology.

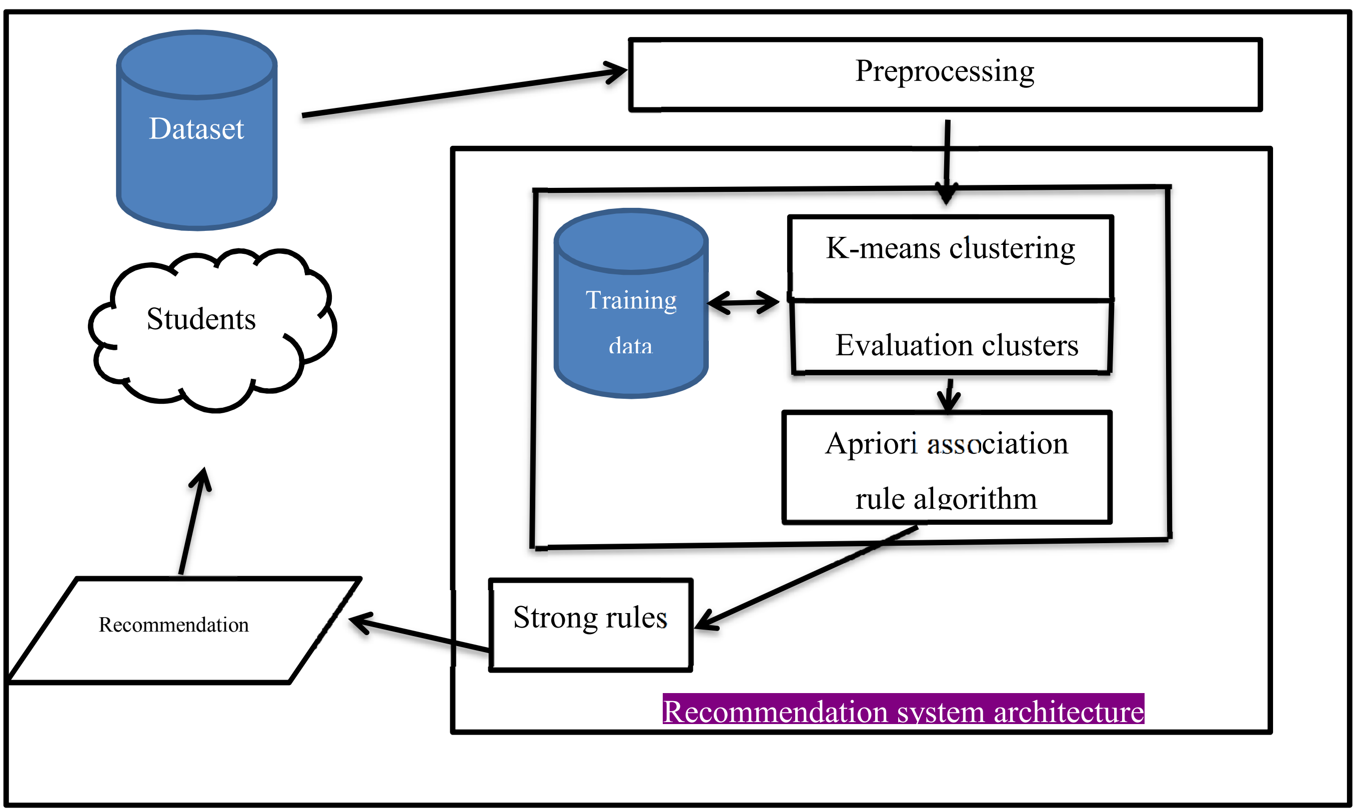



Figure 3-1: Architecture of the methodology

### 3.2 Data preprocessing

The academic data has been collected from the University of KORDOFAN, faculty of computer studies and statistics. The data contains 1620 records corresponding to students enrolled through the year's study 2008 to 2014. The structure of study in faculty has been distributed into four year study equivalent to eight semesters. Each year study included two semesters, the number of courses in faculty equals sixty courses affording to all the semesters. The dataset contains the number of the index, the name of the student, number of student, courses, and average of courses. Since hidden the name and number of student due to privacy, figure 3-2 explains the sample of data to one semester marks.

| L | J | I | H | G | F | E | D | C | B | A |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | تخصص : تقانة المعلومات | | | | | | | |
| المتوسط | تحليلية | م.اقتصاد | ا.ملفات | أ.البرمجة | أ.تطبيقي | م.الالكترونيات | لغة انجليزيةIII | رقم الجلوس | الاسم | م |
| 20 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | الساعات المعتمدة | | |
| 67 | 60 | 85 | 65 | 50 | 87 | 73 | 52 | 00-BT00 | | 1 |
| 72 | 71 | 76 | 71 | 66 | 94 | 76 | 50 | 00-BT00 | | 2 |
| 29 | 58 | 60 | غ غ | 45 | غ غ | غ غ | 40 | 00-BT00 | | 3 |
| 60 | 60 | 67 | 57 | 62 | 56 | 56 | غ غ | 00-BT00 | | 4 |
| 69 | 61 | 78 | 74 | 70 | 75 | 71 | 56 | 00-BT00 | | 5 |
| 66 | 52 | 81 | 63 | 67 | 89 | 52 | 55 | 00-BT00 | | 6 |
| 63 | 50 | 70 | 52 | 52 | 80 | 86 | 52 | 00-BT00 | | 7 |

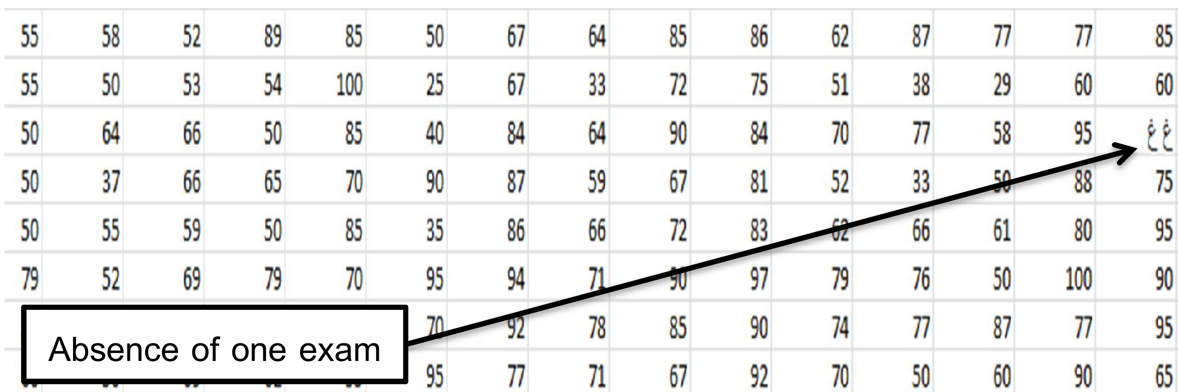
Figure 3-2: sample from the dataset

Applying data processing techniques before mining will typically improve the overall quality of the items mined, and/or reduce the required time for the actual mining.

In this study some general tasks of the data preprocessing have to be performed on the dataset, such as data integration, data cleaning, data reduction, data transformation.

The first task of the data preprocessing is the data integration. In this research merged data related to the students involved with four files. The first semesters and second semesters to one file (first year), third semesters of fourth semesters to one file (second year), fifth semesters of sixth semesters to one file (third year), seventh semesters of eight semesters to one file (four years).

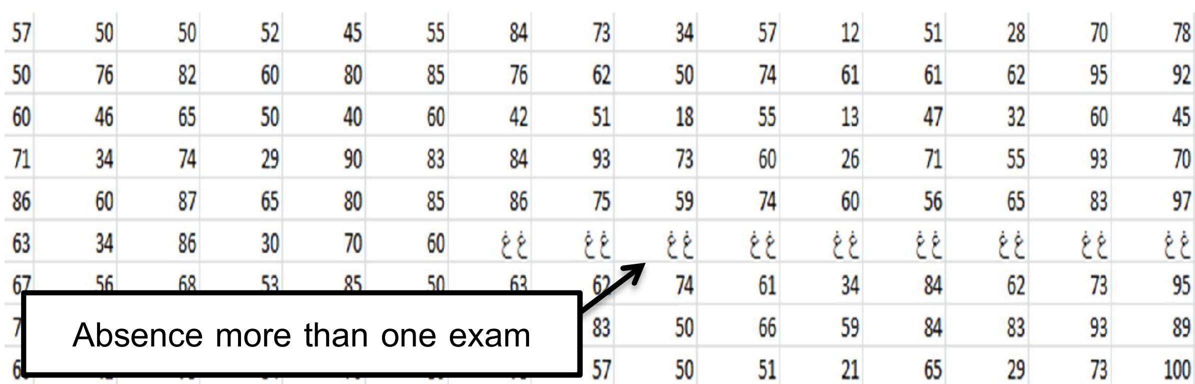
The second task of the data preprocessing is handling the missing Data. The problem of missing data generally arises due to the absence of data in a statement for any variable during the experiments, or when no information is provided or unavailable for the variables. In this study, there are two cases of the missing data will be handling. First case; when the student is absent from one exam as shown in Figure (3-3), which was handled by using the attribute mean for all samples belonging to the same class as the given tuple. The second case; when the student is absented of the more than one exam as shown in Figure (3-4), which was handled by ignoring the tuple (remove tuple).



| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 55 | 58 | 52 | 89 | 85 | 50 | 67 | 64 | 85 | 86 | 62 | 87 | 77 | 77 | 85 |
| 55 | 50 | 53 | 54 | 100 | 25 | 67 | 33 | 72 | 75 | 51 | 38 | 29 | 60 | 60 |
| 50 | 64 | 66 | 50 | 85 | 40 | 84 | 64 | 90 | 84 | 70 | 77 | 58 | 95 | ؟؟ |
| 50 | 37 | 66 | 65 | 70 | 90 | 87 | 59 | 67 | 81 | 52 | 33 | 50 | 88 | 75 |
| 50 | 55 | 59 | 50 | 85 | 35 | 86 | 66 | 72 | 83 | 62 | 66 | 61 | 80 | 95 |
| 79 | 52 | 69 | 79 | 70 | 95 | 94 | 71 | 90 | 97 | 79 | 76 | 50 | 100 | 90 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 70 | 92 | 78 | 85 | 90 | 74 | 77 | 87 | 77 | 95 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 95 | 77 | 71 | 67 | 92 | 70 | 50 | 60 | 90 | 65 |



Figure 3-3: missing data of one course



| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 57 | 50 | 50 | 52 | 45 | 55 | 84 | 73 | 34 | 57 | 12 | 51 | 28 | 70 | 78 |
| 50 | 76 | 82 | 60 | 80 | 85 | 76 | 62 | 50 | 74 | 61 | 61 | 62 | 95 | 92 |
| 60 | 46 | 65 | 50 | 40 | 60 | 42 | 51 | 18 | 55 | 13 | 47 | 32 | 60 | 45 |
| 71 | 34 | 74 | 29 | 90 | 83 | 84 | 93 | 73 | 60 | 26 | 71 | 55 | 93 | 70 |
| 86 | 60 | 87 | 65 | 80 | 85 | 86 | 75 | 59 | 74 | 60 | 56 | 65 | 83 | 97 |
| 63 | 34 | 86 | 30 | 70 | 60 | ؟؟ | ؟؟ | ؟؟ | ؟؟ | ؟؟ | ؟؟ | ؟؟ | ؟؟ | ؟؟ |
| 67 | 56 | 68 | 53 | 85 | 50 | 63 | 62 | 74 | 61 | 34 | 84 | 62 | 73 | 95 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 83 | 50 | 66 | 59 | 84 | 83 | 93 | 89 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 57 | 50 | 51 | 21 | 65 | 29 | 73 | 100 |



Figure 3-4: missing data of more than one

The third task of the data preprocessing in this study is the attribute Subset Selection. The feature selection is one of the important and frequently used techniques in data preprocessing for data mining. Feature selection is a process of identifying and selecting a useful subset from original features. In this study will be select two attributes as per names of courses and course degree, add to these attributes as GPA of students.

The last task in the data preprocessing is the data transformation. the data are transformed or consolidated into forms appropriate for mining. Data transformations involve many techniques, in this study use normalization technique. Normalizing the courses marks since the performance of students is compared by associating the grades with various percentages. The attribute data are scaled so as to fall within a small specified range, as represented in the table: 3-1.

Table 3-1: Normalization data by specified range

| The range of the point degree | Symbol of the point degree | Grade |
|---|---|---|
| 3.19 – 4.0 | A | Excellent |
| 2.79 – 3.18 | B | Very good |
| 2.39 – 2.98 | C | Good |
| 1.99 – 2.38 | D | Pass |
| 0 – 1.98 | F | Fail |

Figure 3-5 explains the marks normalized to the point degrees with various values. Since some implementations of K-means only allow numerical values for attributes.

Some techniques, such as association rule mining, can only be performed on categorical data. This requires performing discretization on numeric or continuous attributes. In this case, point degrees normalized to a symbol such as A, B, C, D, F as shown in Figure 3-6.

| accounting | operating1 | App1 | Islam2 | Eng2 | Arb2 | Commun | Pro1 | statis | calculus | operating2 | App2 | GPA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3.24 | 2.4 | 2.12 | 3.08 | 2.6 | 2.44 | 1.6 | 2.04 | 2.08 | 2.48 | 2.6 | 2 | pass |
| 3.28 | 1.52 | 2.84 | 2.2 | 3.48 | 3 | 2.6 | 2.08 | 2.6 | 2.12 | 2.8 | 2.32 | Good |
| 2.92 | 3.2 | 3.68 | 2.92 | 2.12 | 2.88 | 2.4 | 2 | 2.44 | 2.12 | 2.72 | 3.28 | Good |
| 2.92 | 2.48 | 2.2 | 2.44 | 2.2 | 2.92 | 2.76 | 2 | 2.12 | 2.64 | 2.6 | 2.08 | pass |
| 3.84 | 1.44 | 3.84 | 3.24 | 3.16 | 3.6 | 3.08 | 2.08 | 3.4 | 2.56 | 3 | 3.72 | Very good |
| 2.2 | 3.04 | 3.68 | 2.2 | 2.56 | 2.8 | 2.56 | 2.44 | 2.4 | 2.4 | 2.64 | 3 | Good |
| 1.72 | 2.56 | 3.48 | 2.4 | 2.24 | 2.96 | 2.36 | 2 | 2.12 | 2.2 | 2.84 | 2.76 | pass |
| 2.24 | 3.2 | 3.92 | 1.44 | 2.52 | 1.04 | 1.8 | 0.92 | 1.52 | 0.72 | 2.36 | 3.12 | Fail |
| 1.4 | 3.76 | 3.4 | 2.36 | 2.72 | 0.84 | 1.8 | 0.84 | 0.84 | 0.52 | 1.48 | 1.6 | Fail |
| 1.24 | 2.52 | 3.36 | 2.36 | 2.12 | 2.08 | 2.04 | 1.28 | 1.6 | 2.16 | 2.08 | 2.6 | Fail |
| 2.44 | 3 | 2.6 | 2.84 | 2.28 | 2.96 | 2.56 | 2.8 | 3.04 | 2.2 | 3.2 | 3.4 | Good |
| 2.28 | 2.4 | 3 | 1.84 | 2.04 | 2.32 | 2.4 | 1.16 | 1.56 | 2.12 | 2.84 | 1.8 | pass |
| 3.28 | 2.2 | 3.44 | 2.56 | 2.8 | 3.08 | 2.6 | 2 | 2.04 | 2 | 2.76 | 2.2 | Good |
| 3.28 | 2.52 | 3.52 | 2.32 | 2.88 | 2.92 | 3.24 | 2.72 | 2.36 | 2.48 | 3.16 | 3.32 | Very good |
| 3.68 | 3.6 | 3.8 | 2.36 | 2.96 | 3.2 | 2.36 | 2.52 | 3.48 | 2.72 | 2.76 | 3.72 | Very good |

Figure 3-5: normalize marks to the point degree

| B.mathB.r | Compter | accountin | operating | App1 | Islam2 | Eng2 | Arb2 | Commun | Pro1 | statis | calculus | operating | App2 | GPA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| C | D | A | C | D | B | C | C | F | D | D | C | C | D | pass |
| D | D | A | F | B | D | A | B | C | D | C | D | B | D | Good |
| D | B | B | A | A | B | D | B | C | D | C | D | C | A | Good |
| D | D | B | C | D | C | D | B | C | D | D | C | C | D | pass |
| B | B | A | F | A | A | B | A | B | D | A | C | B | A | Verygood |
| B | B | D | B | A | D | C | B | C | C | C | C | C | B | Good |
| D | C | F | C | A | C | D | B | D | D | D | D | B | C | pass |
| D | D | D | A | A | F | C | F | F | F | F | F | D | B | Fail |
| F | F | F | A | A | D | C | F | F | F | F | F | F | F | Fail |
| D | F | F | C | A | D | D | D | D | F | F | D | D | C | Fail |
| D | C | C | B | C | B | D | B | C | B | B | D | A | A | Good |
| D | D | D | C | B | F | D | D | C | F | F | D | B | F | pass |
| D | C | A | D | A | C | B | B | C | D | D | D | C | D | Good |
| D | C | A | C | A | D | B | B | A | C | D | C | B | A | Verygood |
| B | B | A | A | A | D | B | A | D | C | A | C | C | A | Verygood |

Figure 3-6: normalize the point degree to the symbol

### 3.3 Clustering based on Collaborative filtering

According to related work for using recommendation system in educational environment, most researches have proven that the Collaborative filtering technique of recommendation system was the more efficient technique used in filtering data that support students to improve their academic performance (Vialardi *et al.*, 2009) and (Schafer *et al.*, 2007) and (Hameed, 2012).

Clustering is the processes of grouping the data into classes or clusters so that the objects within a cluster have high similarity in comparison to one another, while are very dissimilar to objects in other clusters. Dissimilarities are assessed based on the attributes values describing the objects. Thus clustering is one of the most common methods used for collaborative filtering algorithms.

K-mean clustering technique is used in this work because of its simplicity and being suitable to be used with numerical unsupervised data like student courses' grades. The main idea of the K-mean clustering is to define k centers, one for each cluster. The next step is to take each point belonging to a given dataset and associate it with the nearest centers. At this point, need to re-compute *k* new centers of the clusters resulting from the previous step. After having these *k* new centers, a new binding has to be done between the same data set points and the nearest new centers; a loop has been generated. As a result of this loop, may notice that the k centers change their location step by step until no more changes are done. Similarity distance measures are commonly used for computing the objects similarity of the clusters described by interval scaled variables (Al-Badarenah and Alsakran, 2016). Interval scale variables are continuous measurements of roughly linear scale. The similarity between the objects described by interval-scaled variables is typically computed based on the distance between each pair of the objects. In this study Euclidean distance is used to calculate similarly items based on clustering collaborative filtering recommender which that mentioned in chapter two. The below figure 3-7 explains steps of the clustering process.

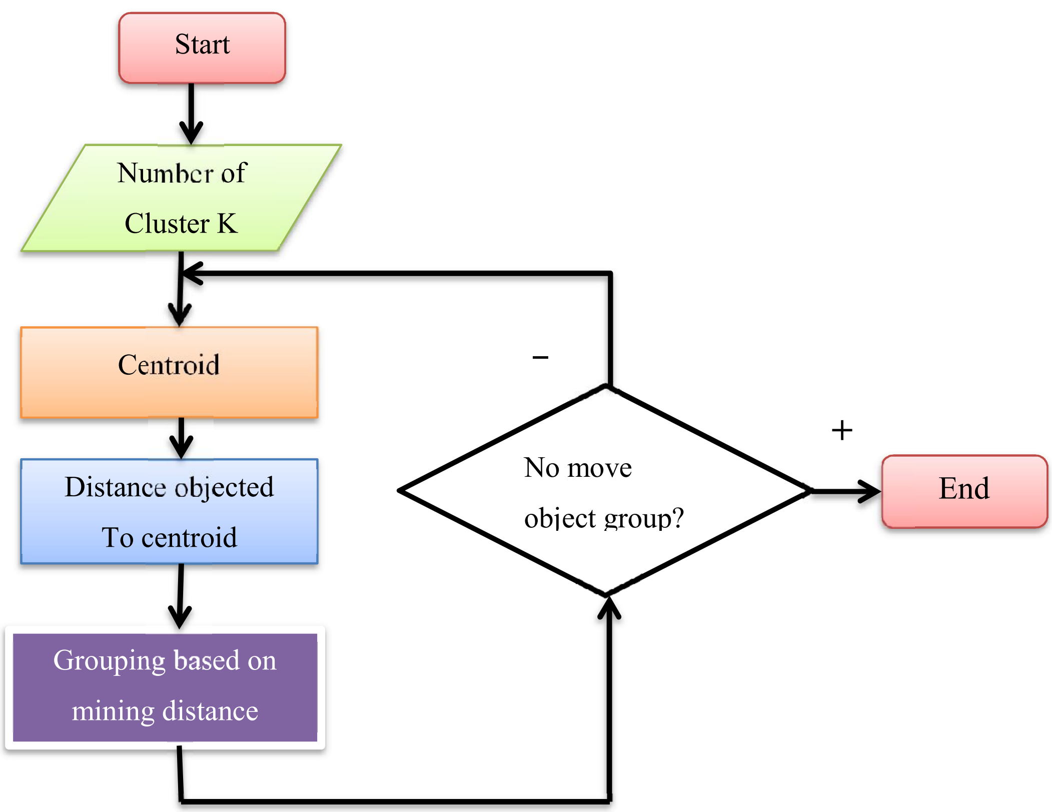



Figure 3-7: steps of the clustering process

The below figure 3-8 shown the Euclidean distance based on k means cluster which deployed on weka tool.

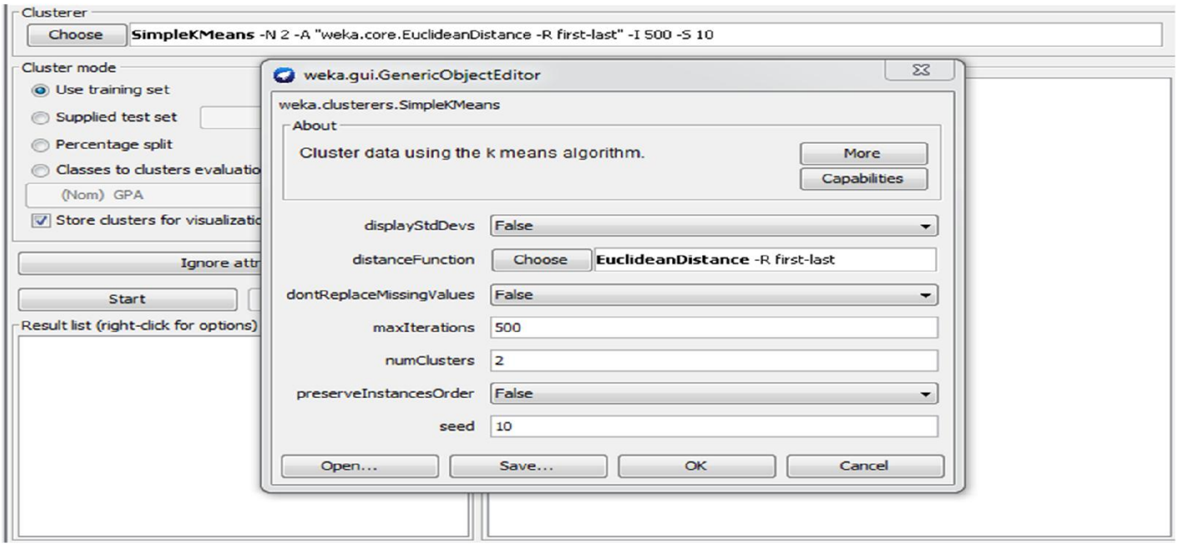
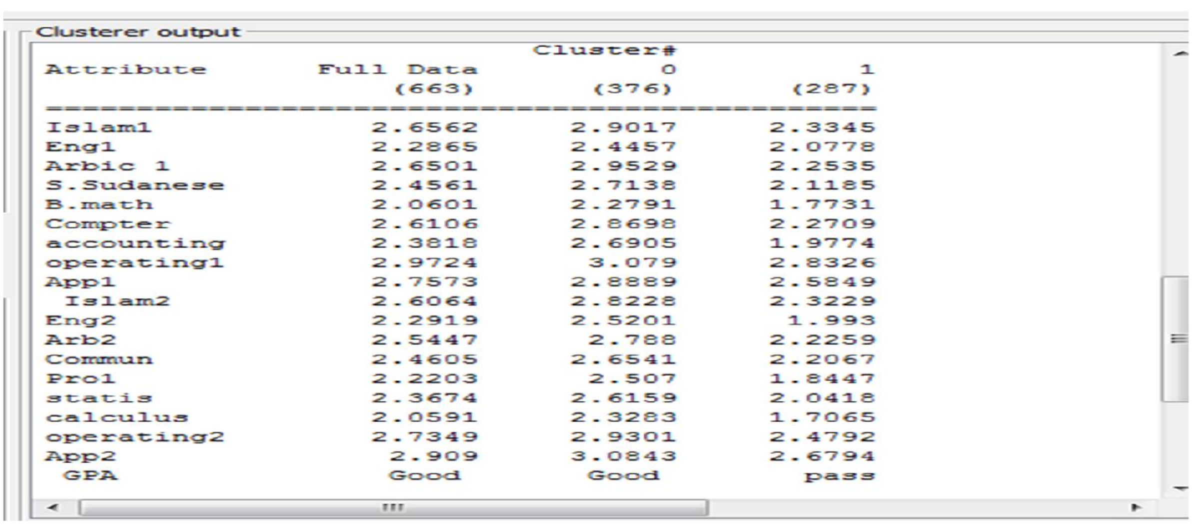



Figure 3-8: Euclidean distance

The datasets are deployed in Weka tool and then k means clustering algorithms are applied to the datasets with classes to the cluster. The datasets equivalent to four files that mentioned in section 3.2 in this chapter. Each file will be deploying separately on other. The figure 3-9 as explained the result of dataset implemented on one file (first year) using k means cluster algorithms in Wake tool. In the same way appalling k means on all files.


Clusterer output

| Attribute | Full Data | Cluster# 0 | 1 |
|---|---|---|---|
| | (663) | (376) | (287) |
| Islam1 | 2.6562 | 2.9017 | 2.3345 |
| Eng1 | 2.2865 | 2.4457 | 2.0778 |
| Arbic 1 | 2.6501 | 2.9529 | 2.2535 |
| S.Sudanese | 2.4561 | 2.7138 | 2.1185 |
| B.math | 2.0601 | 2.2791 | 1.7731 |
| Compter | 2.6106 | 2.8698 | 2.2709 |
| accounting | 2.3818 | 2.6905 | 1.9774 |
| operating1 | 2.9724 | 3.079 | 2.8326 |
| App1 | 2.7573 | 2.8889 | 2.5849 |
| Islam2 | 2.6064 | 2.8228 | 2.3229 |
| Eng2 | 2.2919 | 2.5201 | 1.993 |
| Arb2 | 2.5447 | 2.788 | 2.2259 |
| Commun | 2.4605 | 2.6541 | 2.2067 |
| Pro1 | 2.2203 | 2.507 | 1.8447 |
| statis | 2.3674 | 2.6159 | 2.0418 |
| calculus | 2.0591 | 2.3283 | 1.7065 |
| operating2 | 2.7349 | 2.9301 | 2.4792 |
| App2 | 2.909 | 3.0843 | 2.6794 |
| GPA | Good | Good | pass |

Figure 3-9: Deployed dataset on k mean

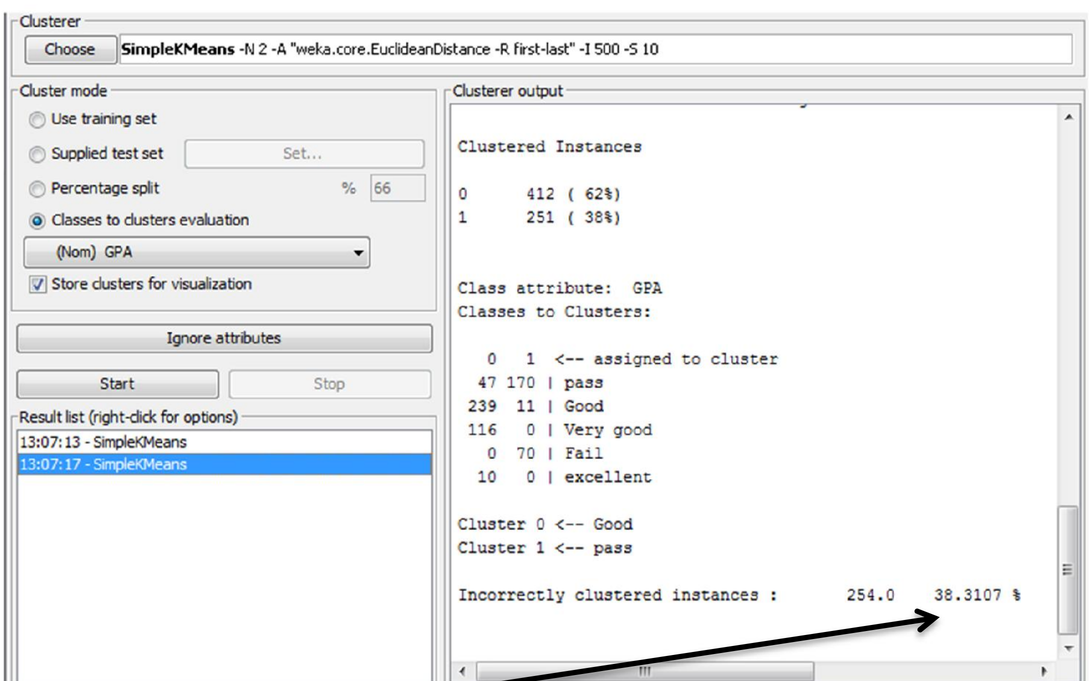
### 3.4 Evaluation clusters results

For evaluation, clusters results are deployed inside section 3-3. In this chapter the criterion function E is used to calculate error as explained in equation (3).

$$E=\sum_{i=1}^{K}\sum_{P\,\epsilon\,Ci}\left|P\text{-}M_{i.}\right| \qquad \text{equation (3)}$$

In which, *E* is a total square error of all the objects in the data cluster, *p* is given data object, *mi* is the mean value of cluster *Ci* (*p* and *m* are both multidimensional). The function of this criterion is to make the generated cluster be as compacted and independent as possible (Godara and Yadav, 2013).

Figure 3-10 explained the criterion function E which deployed on the dataset to calculate the error. The same way applied to calculated the error in all year study.




Incorrectly cluster when the number of k=2

Figure 3-10: results of incorrectly clustered

The accuracy of clustering algorithms calculated the correct instances for each year study. Table 3-3 explains the several experiments for the first year to determine the best cluster.

The total of correctly instances is high when the number of clusters equals three comparing the other numbers of clusters. As the similar way executed on second, three and four years studies. The results of the best cluster in each year as shows in below tables respectively, table 3- 4 when the number of clusters equals three, table 3-5 when the number of clusters equals two and table 3-6 when the number of clusters equals two.

Table 3-3: Evaluation cluster results in the first year

| Number of k | Total number of Instances | Correctly clustered instances | Accuracy |
|---|---|---|---|
| **3** | **663** | **390** | **58.82%** |
| 4 | 663 | 356 | 53.70% |
| 5 | 663 | 387 | 58.37% |

Table 3-4: Evaluation cluster results in the second year

| Number of k | Total number of Instances | Correctly clustered instances | Accuracy |
|---|---|---|---|
| 2 | 424 | 249 | 58.72% |
| **3** | **424** | **303** | **71.46%** |
| 4 | 424 | 265 | 62.5% |

Table 3-5: Evaluation cluster results in the third year

| Number of k | Total number of Instances | Correctly clustered instances | Accuracy |
|---|---|---|---|
| **2** | **249** | **175** | **70.28%** |
| 3 | 249 | 152 | 61.04% |
| 4 | 249 | 119 | 47.79% |

Table 3-6: Evaluation cluster results in four year

| Number of k | Total number of Instances | Correctly clustered instances | Accuracy |
|---|---|---|---|
| 2 | 146 | 86 | 58.90% |
| **3** | **146** | **87** | **59.59%** |
| 4 | 146 | 83 | 56.85% |

The below figures (3.11 to 3.14) explain the high accuracy of clusters numbers. Figure 3-11shows the content of three clusters with a size of 663 records. Clearly, the grade of pass is a high ratio as shown in the cluster (1) comparing the cluster (0) and cluster (2).
Figure 3.12 shows the content of three clusters with a size of 424 records. Obviously, the grade of pass is a high ratio as shown in the cluster (0) than cluster (1) grade of good and cluster (2) the grade of fail.
Figure 3.13 shows the content of two clusters with a size of 250 records. Clearly, the grade of pass is a high ratio as shown in the cluster (1) than cluster (0) grade of good.
Figure 3.11 shows the content of two clusters with a size of 146 records. Obviously, the grade of good is a high ratio as shown in the cluster (0) than cluster (1) grade of pass.
Each subcluster must be saving individually such as the three clusters in figure 3.11 kept corresponding to three files.

| Attribute | Full Data (663) | 0 (126) | 1 (287) | 2 (250) |
|---|---|---|---|---|
| Islam1 | 2.6562 | 3.1022 | 2.3345 | 2.8006 |
| Eng1 | 2.2865 | 2.5946 | 2.0778 | 2.3707 |
| Arbic 1 | 2.6501 | 3.2105 | 2.2535 | 2.823 |
| S.Sudanese | 2.4561 | 2.9432 | 2.1185 | 2.5982 |
| B.math | 2.0601 | 2.4502 | 1.7731 | 2.193 |
| Compter | 2.6106 | 3.047 | 2.2709 | 2.7805 |
| accounting | 2.3818 | 2.9098 | 1.9774 | 2.58 |
| operating1 | 2.9724 | 3.24 | 2.8326 | 2.9979 |
| App1 | 2.7573 | 3.0283 | 2.5849 | 2.8187 |
| Islam2 | 2.6064 | 2.9768 | 2.3229 | 2.7451 |
| Eng2 | 2.2919 | 2.633 | 1.993 | 2.4632 |
| Arb2 | 2.5447 | 3.0149 | 2.2259 | 2.6736 |
| Commun | 2.4605 | 2.8511 | 2.2067 | 2.5549 |
| Pro1 | 2.2203 | 2.7105 | 1.8447 | 2.4045 |
| statis | 2.3674 | 2.8594 | 2.0418 | 2.4931 |
| calculus | 2.0591 | 2.5235 | 1.7065 | 2.2299 |
| operating2 | 2.7349 | 3.1425 | 2.4792 | 2.823 |
| App2 | 2.909 | 3.3054 | 2.6794 | 2.9728 |
| GPA | Good | Very good | pass | Good |

Figure 3.11: divided data into three clusters in the first year

Cluster centroids:

| Attribute | Full Data (424) | Cluster# 0 (161) | 1 (186) | 2 (77) |
|---|---|---|---|---|
| Eng3 | 2.32 | 2.4 | 2.32 | 2.2 |
| Economy | 2.72 | 3.04 | 2.68 | 2.28 |
| Files | 2.36 | 2.64 | 2.32 | 2 |
| Pro.Methods 1 | 2.24 | 2.56 | 2.12 | 1.6 |
| Electricity | 2.4 | 2.76 | 2.32 | 2 |
| Apps.statist | 2.52 | 2.92 | 2.38 | 2 |
| Math2 | 2.12 | 2.48 | 2 | 1.32 |
| Pro. Methods 2 | 2.36 | 2.88 | 2.2 | 2 |
| OP.Research | 2.6 | 3 | 2.52 | 2.04 |
| MIS | 2.32 | 2.56 | 2.2 | 2 |
| DB1 | 2.2 | 2.52 | 2.16 | 1.72 |
| Algebra | 2.04 | 2.44 | 2 | 1.44 |
| Discrete.math | 2.12 | 2.4 | 2 | 1.64 |
| APP.net | 2 | 2.44 | 1.8 | 1.44 |
| GPA | pass | Good | pass | Fail |

Figure 3.12: divided data into three clusters in the second year.

Cluster centroids:

| Attribute | Full Data (250) | Cluster# 0 (99) | 1 (151) |
|---|---|---|---|
| OOP1 | 2.6432 | 2.9871 | 2.4178 |
| T.internet 1 | 2.1717 | 2.48 | 1.9696 |
| HCI | 2.3153 | 2.4614 | 2.2196 |
| Network | 2.4508 | 2.8222 | 2.2072 |
| Structure | 2.1902 | 2.4347 | 2.0299 |
| manage | 2.5552 | 2.8045 | 2.3918 |
| Analysis 1 | 2.3329 | 2.6699 | 2.1119 |
| AI | 2.5284 | 2.7341 | 2.3934 |
| Analysis 2 | 2.4618 | 2.7313 | 2.2852 |
| research | 2.4692 | 2.7572 | 2.2805 |
| OS | 2.2267 | 2.4149 | 2.1032 |
| Media | 2.3855 | 2.6408 | 2.2182 |
| OOP2 | 2.2361 | 2.7329 | 1.9104 |
| T.internet 2 | 2.0416 | 2.4335 | 1.7846 |
| Engineering 1 | 2.4289 | 2.7317 | 2.2304 |
| GPA | pass | Good | pass |

Figure 3.13: divided data into two clusters in the three year

Cluster centroids:

| Attribute | Full Data (146) | Cluster# 0 (96) | Cluster# 1 (50) |
|---|---|---|---|
| M.Network 1 | 2.2956 | 2.4792 | 1.9432 |
| Engineering 2 | 2.3548 | 2.5158 | 2.0456 |
| simulation | 2.7236 | 2.8912 | 2.4016 |
| E commerce | 2.7085 | 2.7962 | 2.54 |
| crpto 1 | 2.3964 | 2.4671 | 2.2608 |
| M.Internet | 2.4225 | 2.5629 | 2.1528 |
| M.Network 2 | 2.2753 | 2.4708 | 1.9 |
| crypto 2 | 2.257 | 2.3783 | 2.024 |
| visual | 2.2879 | 2.4921 | 1.896 |
| A.DB | 2.6808 | 2.7538 | 2.5408 |
| Media 2 | 2.2762 | 2.4146 | 2.0104 |
| ethical | 2.7825 | 2.9767 | 2.4096 |
| GPA | Good | Good | pass |

Figure 3.14: divided data into two clusters in the four year

### 3.5 Recommendations Using Association Rules Mining

Association rules are one of the most popular ways of the representing discovered knowledge and describe a close correlation between frequent items in a database. An $X \Rightarrow Y$ type association rule expresses a close correlation between items (attribute-value) in a database. There are many association rule algorithms such as Apriority algorithm. Apriori is the first and foremost algorithm in association rule. The algorithm implemented is especially useful in collaborative recommender systems, in order to produce recommendations that are increasingly useful and precise.

Most association rule mining algorithms require the user to set at least two thresholds, one of the minimum support and the other of minimum confidence. The support $S$ of a rule is defined as the probability that an entry has of satisfying both $X$ and $Y$. Confidence is defined as the probability an entry has of satisfying $Y$ when it satisfies $X$.

In this study, will be applied hybrid Recommendation techniques like apriori algorithm on results of correct clusters there are generated in the previous section (3-4) in this section. In order to the applying apriori algorithm on the clusters results to be covert the numerical values to the nominal values. Many experiments are deployed on the apriori algorithm that considered the results of clusters were generated.

In the first experiment: on the data of the first year study which is contents three clusters. The first cluster observed was generated strong rule with (Support 11% and Confidence 100%) as shown in figure (3-15). The second cluster as shown in figure (3-16) also generated strong rules it is observed with (Support 11% and Confidence 90%). The third cluster was generated strong rules as appear in figure (3-17) with (Support 1% and Confidence 100%).

In the second experiment: on the data of the second year study which is contents three clusters. The first cluster observed was generated strong rule with (Support 0.9% and Confidence 100%) as shown in figure (3-18). The second cluster as shown in figure (3-19) also generated strong rules it is observed with (Support 0.9% and Confidence 100%). The third cluster was generated strong rules as appear in figure (3-20) with (Support 3% and Confidence 100%).

The third experimented: applied to the data of four years led to generated strong rule corresponding to two clusters. The first cluster as appears in figure (3-21) with (Support 0.9% and Confidence 100%). The Second cluster with (Support 0.9% and Confidence 100%) as shown in figure (3-22).

The fourth experimented: applied on fourth-year led to extracted strong rules equivalents two clusters. The first cluster was presented in figure 3-23 with (Support 0.9% and Confidence 100%). The second cluster was presented in figure 3-24 with (Support 0.9% and Confidence 100%).

The last experimented: applied on a dataset without using clustering that is appear weak rules as shown in figure (3-25)

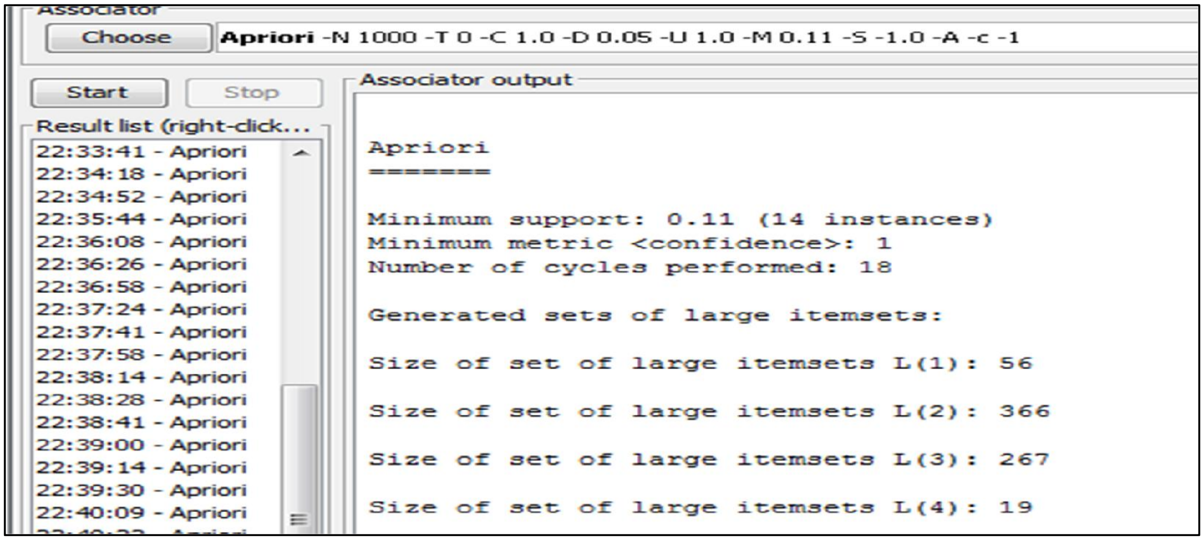



Figure 3-15: generated strong rules on cluster number (0) in the first year

Islam1=A S.Sudanese=A calculus=C App2=A==> GPA=Very good

Islam1=A Commun=C calculus=C App2=A ==> GPA=Very good

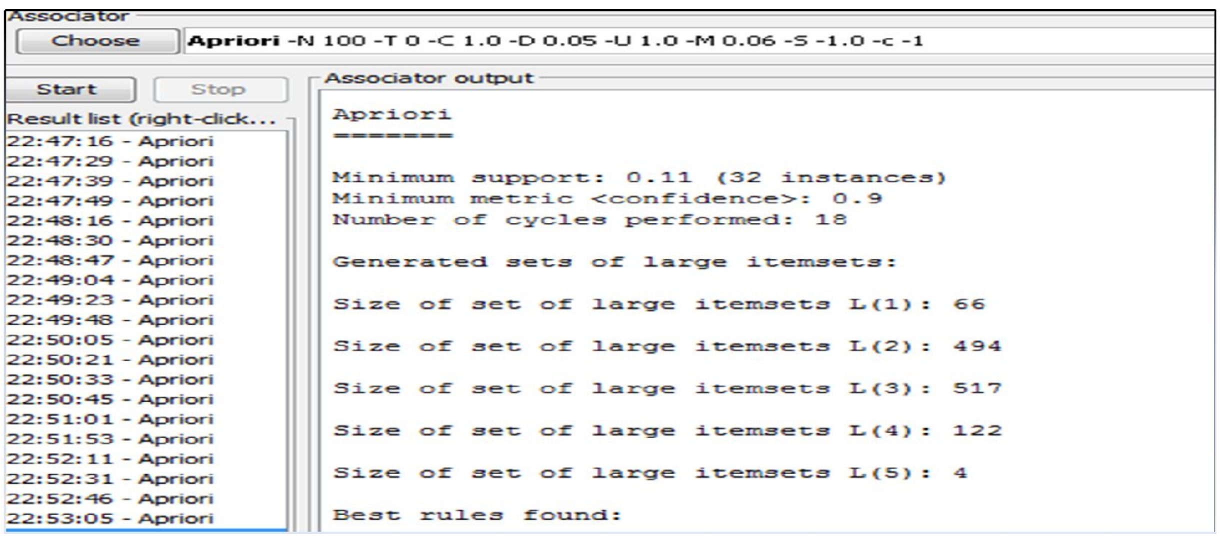



Figure 3-16: generated strong rules on cluster number (1) in the first year

B.math=F Computer=F Pro1=F statis=F calculus=F==> GPA=Fail

B.math=F Computer=F accounting=F Pro1=F calculus=F==> GPA=Fail

B.math=F Computer=F accounting=F statis=F calculus=F ==> GPA=Fail

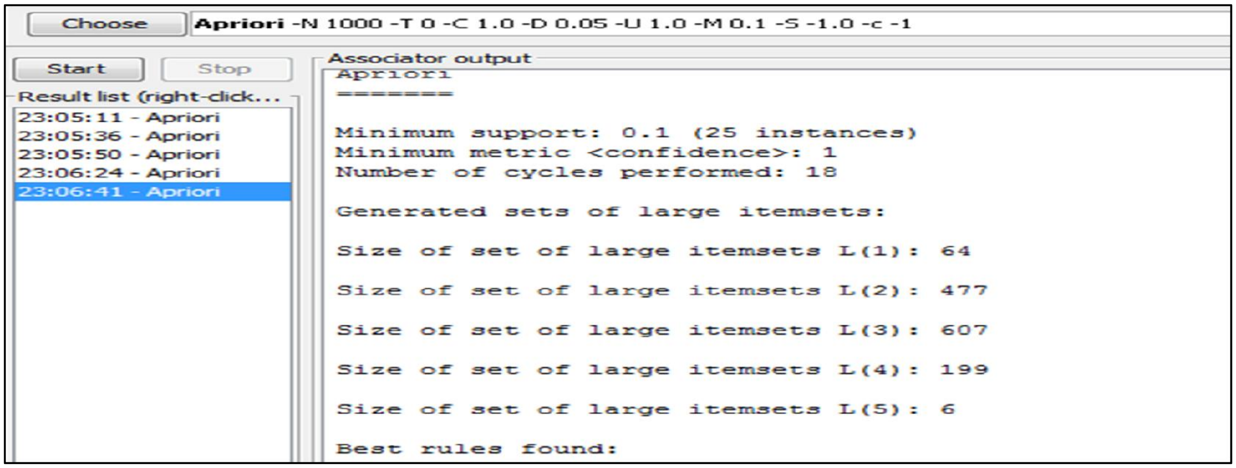



Figure 3-17: generated strong rules on cluster number (2) in the first year

B.math=D operating1=A App1=A calculus=D App2=A==> GPA= Good

B.math=D App1=A Pro1=D calculus=D App2=A ==> GPA=Good

B.math=D Eng2=D Commun=C calculus=D App2=A ==> GPA=Good

B.math=D Commun=C Pro1=D calculus=D App2=A==> GPA=Good

B.math=D Arb2=C Commun=C Pro1=D calculus=D ==> GPA=Good

B.math=D Commun=C Pro1=D calculus=D operating2=C==> GPA=Good

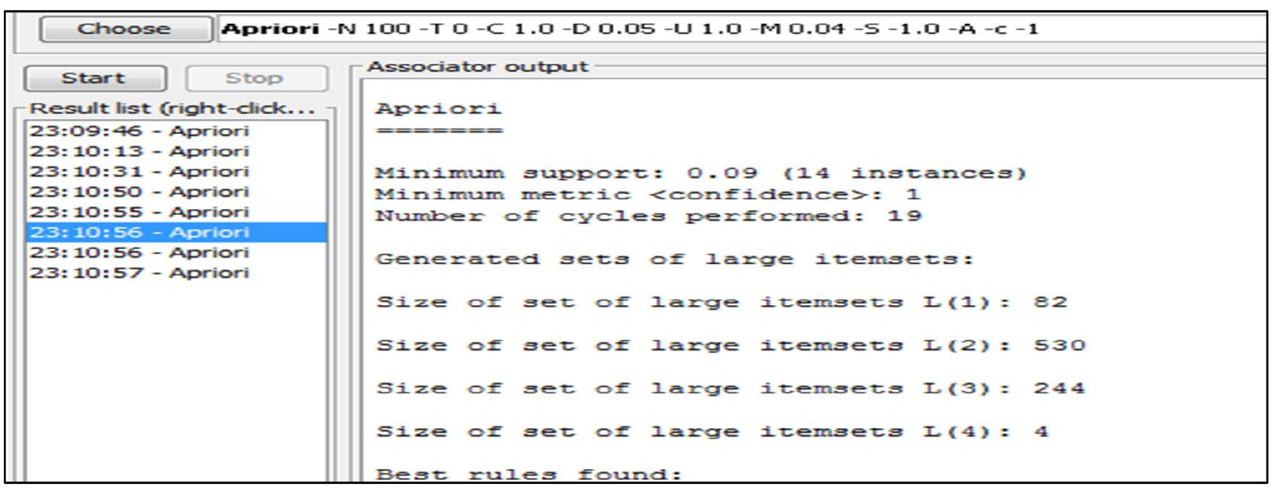



Figure 3-18: generated strong rules on cluster number (0) in the second year

Economy=A Algebra=D Discrete.math=D APP.net=D ==> GPA=Good

Pro.Methods 1=D MIS=C DB1=D Algebra=D==> GPA=Good

Pro.Methods 1=D MIS=C DB1=D APP.net=D ==> GPA=Good

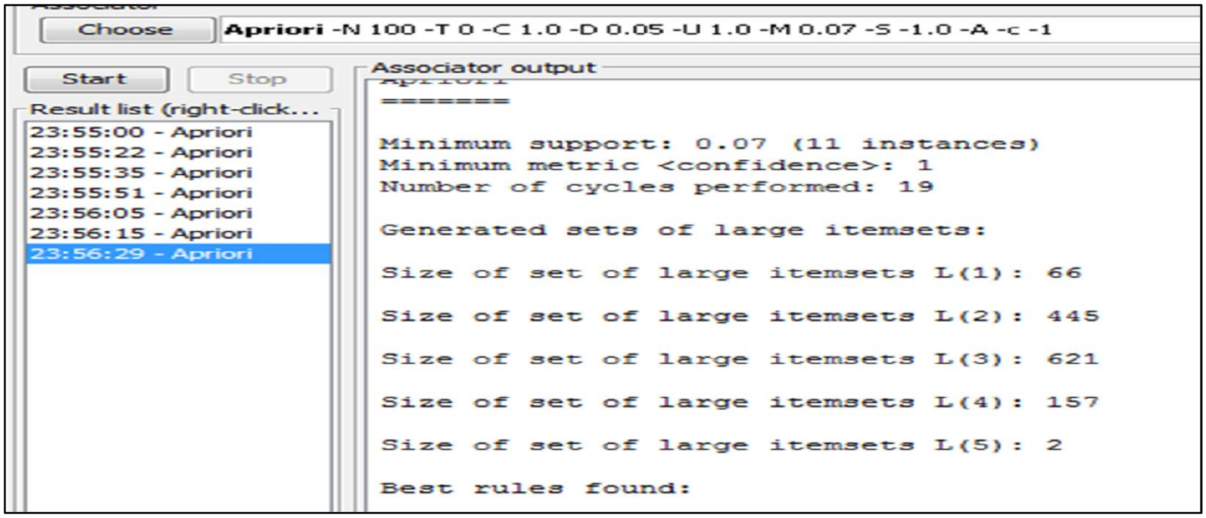



Figure 3-19: generated strong rules on cluster number (1) in the second year

Pro. Methods 2=D MIS=D Algebra=F APP.net=F==> GPA=pass

Pro.Methods 1=D Pro. Methods 2=D MIS=D APP.net=F==> GPA=pass

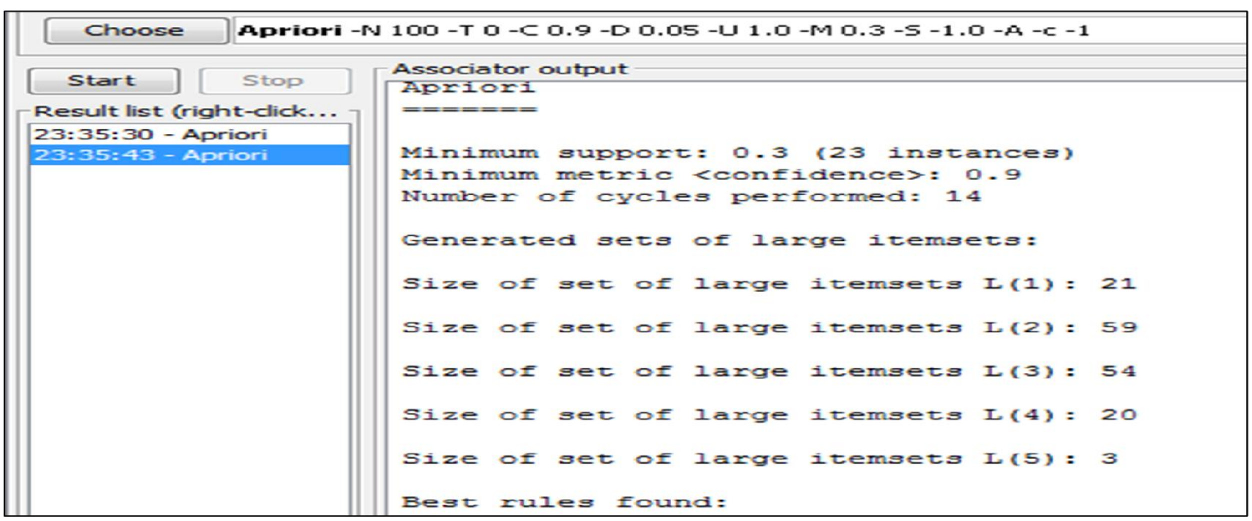



Figure 3-20: generated strong rules on cluster number (2) in the second year

Pro.Methods 1=F Math2=F DB1=F Algcbra=F APP.net=F==> GPA=Fail

Math2=F DB1=F Algebra=F Discrete.math=F APP.net=F==> GPA=Fail

Pro.Methods 1=F Math2=F DB1=F Algebra=F Discrete.math=F==> GPA=Fail

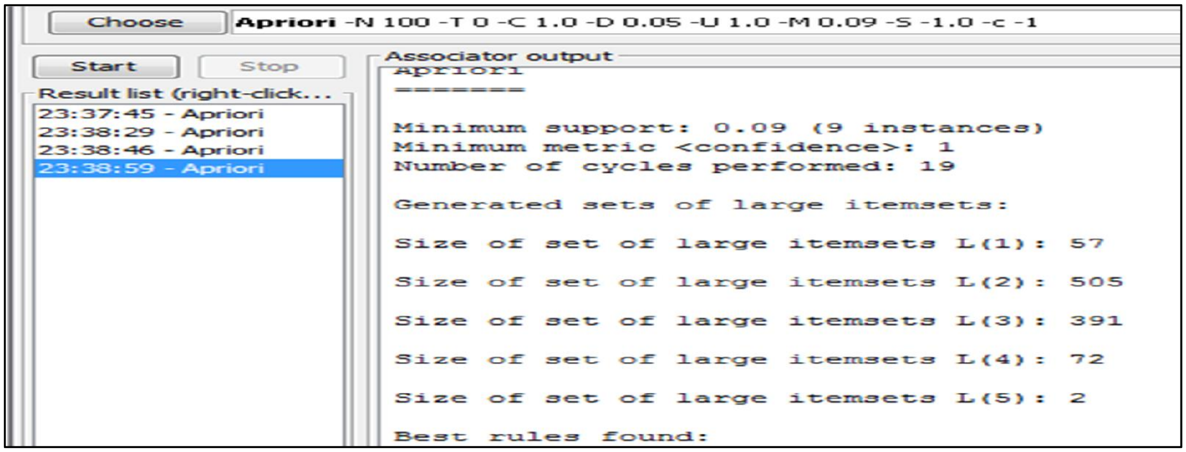



Figure 3-21: generated strong rules on cluster number (0) in the third year

T.internet 1=C HCI=D manage=C AI=D T.internet 2=D ==> GPA=Good

T.internet 1=C manage=C AI=D Media=C T.internet 2=D ==> GPA=Good

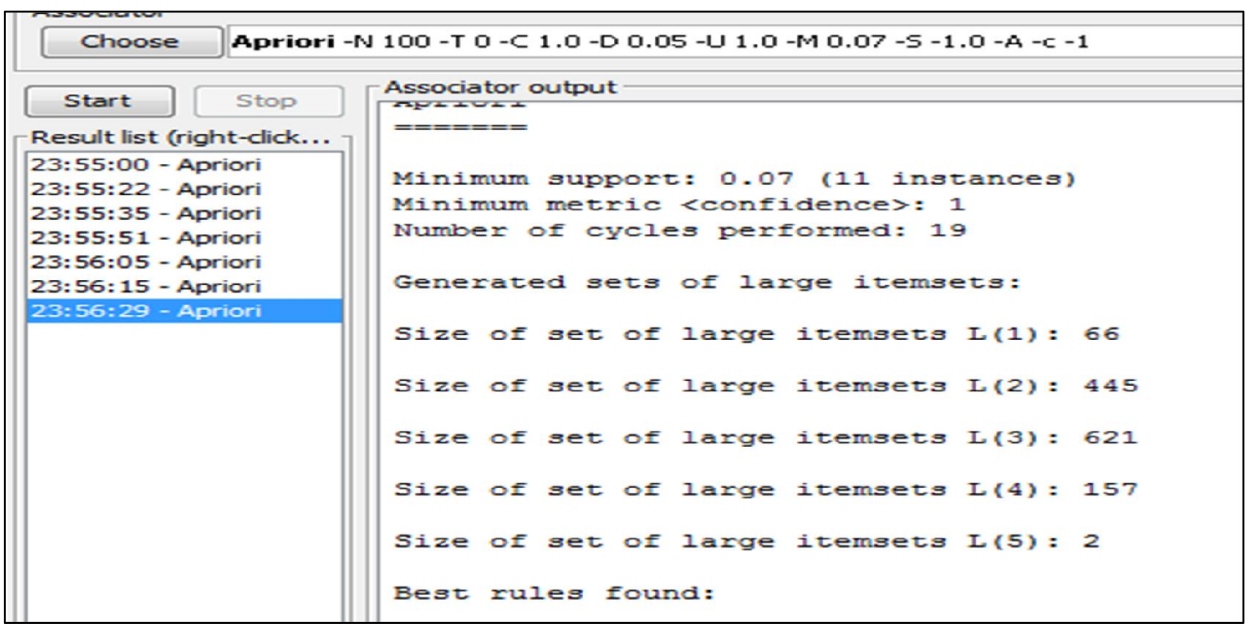



Figure 3-22: generated strong rules on cluster number (1) in the third year

HCI=D OS=D Media=D OOP2=D Engineering 1=D ==> GPA=pass

HCI=D OS=D Media=D T.internet 2=F Engineering 1=D ==> GPA=pass

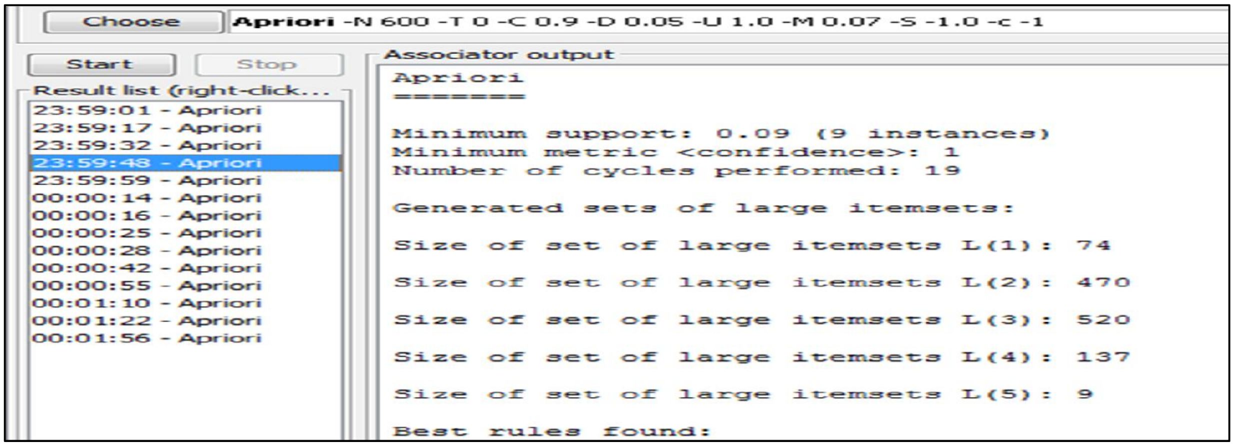



Figure 3-23: generated strong rules on cluster number (0) in four year

M.Network 1=C Engineering 2=D E commerce=C M.Network 2=D crypto 2=C==> GPA=Good

M.Network 1=C Engineering 2=D M.Network 2=D crypto 2=C visual=C==> GPA=Good

M.Network 1=C Engineering 2=D M.Network 2=D visual=C A.DB=C==> GPA=Good

M.Network 1=C E commerce=C M.Network 2=D crypto 2=C Media 2=D==> GPA=Good

M.Network 1=C E commerce=C M.Network 2=D visual=C Media 2=D==> GPA=Good

M.Network 1=C M.Network 2=D crypto 2=C visual=C Media 2=D ==> GPA=Good

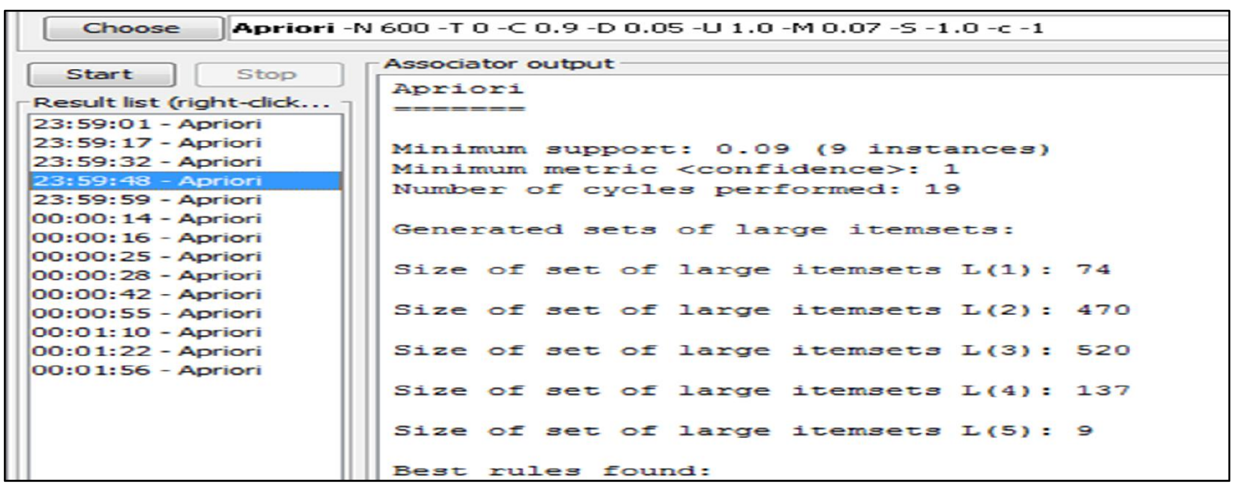



Figure 3-24: generated strong rules on cluster number (1) in four yea

M.Network 1=D Engineering 2=D E commerce=B crpto 1=D M.Internet=D==> GPA=pass

M.Network 1=D Engineering 2=D crpto 1=D M.Internet=D ethical=C==> GPA=pass

M.Network 1=D Engineering 2=D crpto 1=D crypto 2=D ethical=C==> GPA=pass

M.Network 1=D simulation=D E commerce=B crpto 1=D crypto 2=D==> GPA=pass

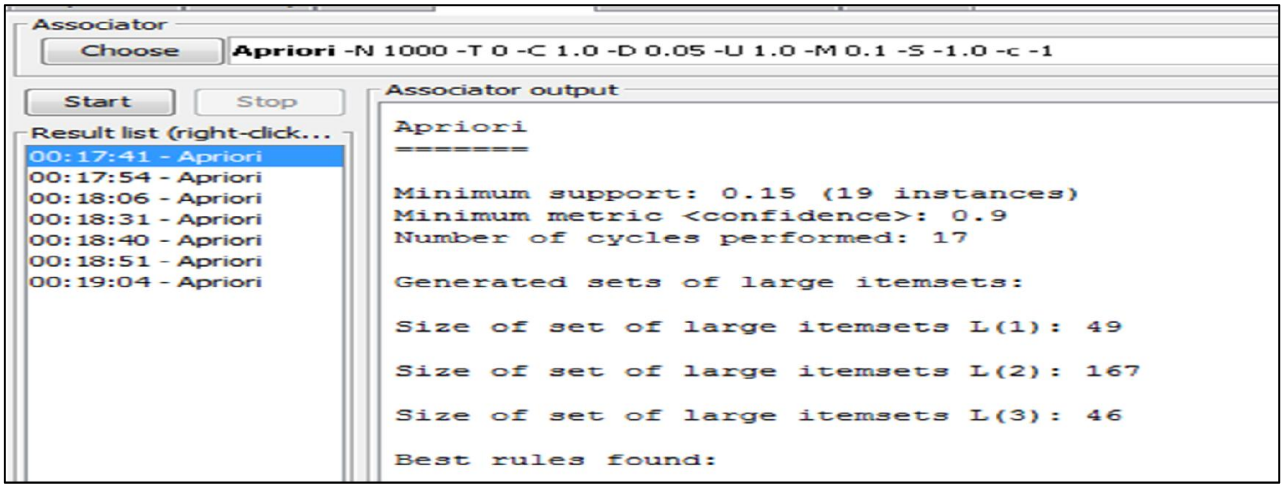



Figure 3-25: generated strong rules on a dataset without clustering

Arbic 1=D Eng3=D ==> Cumulative GPA=pass

S.Sudanese=D Commun=D M.Network 2=D==> Cumulative GPA=pass

B.math=D Pro. Methods 2=D APP.net=F ==> Cumulative GPA=pass

calculus=D Eng3=D Analysis 2=D==> Cumulative GPA=pass

Eng3=D APP.net=F Analysis 2=D ==> Cumulative GPA=pass

Eng3=D APP.net=F T.internet 2=F==> Cumulative GPA=pass

Eng3=D Analysis 2=D T.internet 2=F==> Cumulative GPA=pass

MIS=D Algebra=F M.Internet=D ==> Cumulative GP

B.math=D APP.net=F T.internet 2=F==> Cumulative GPA=pass

Islam2=D calculus=D Eng3=D ==> Cumulative GPA=pass

S.Sudanese=D Analysis 2=D==> Cumulative GPA=pass

OP.Research=D ethical=C==> Cumulative GPA=pass

S.Sudanese=D T.internet 2=F M.Network 2=D ==> Cumulative GPA=pass

B.math=D operating1=A APP.net=F ==> Cumulative GPA=pass

B.math=D operating2=C APP.net=F ==> Cumulative GPA=pass

B.math=D APP.net=F Engineering 2=D ==> Cumulative GPA=pass

Islam2=D Electricity=D Media 2=D ==> Cumulative GPA=pass

calculus=D Eng3=D OS=D 16 ==> Cumulative GPA=pass

### 3.6 Summary

The purpose of this Chapter is to present and discuss the approach and methods of the research. Hence, it covers the methodological aspects that have guided the present work. It starts with an introduction, which gives an overview of the methodology of work then data description and preprocessing approach. Several learning algorithms used this research are described, which are mainly used to determine the performance of the proposed methodology. Finally, the methodology used for evaluation of the research process and results are presented and discuss in the next chapter.

## CHAPTER FOUR
## DISCUSSION OF RESULTS

### 4.1 Introductions

This chapter discussed the results of the experiment have been implemented in chapter three. The results were carried on three phases first; explain the results of k-means clusters have been applied in chapter three. Second, intersection the strong rules were generated in chapter three. Third, discuss the recommendation results.

### 4.2. Experiments the k-means clusters based on similarity of collaborative filtering

At this point, the results have been implemented in chapters three, as shown in figures (4-1 to 4-2). Figure, 4-1 observed most students in the first year have been weak results as the cluster number one. In figure 4-2, most students in the second year have been weak results as the cluster number two. In figure 4-3, student in the third year has been weak results as the cluster number one. In figure 4-4, observed the most student in four years has been weak results as the cluster number one.

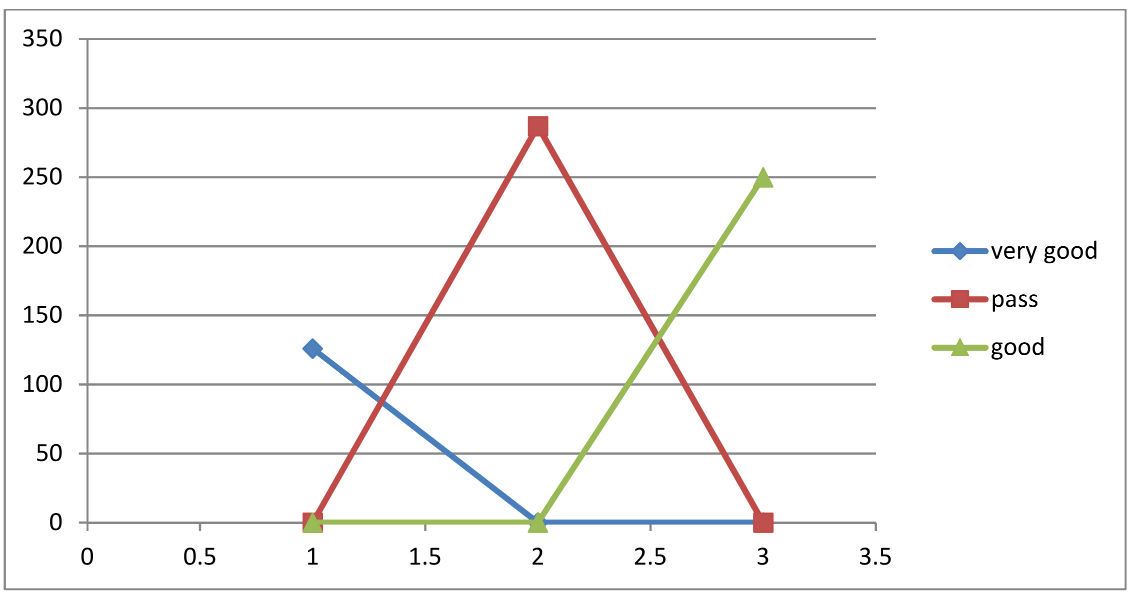



Figure 4-1: divided data into three clusters in the first year


Figure 4-2: divided data into three clusters in the second year

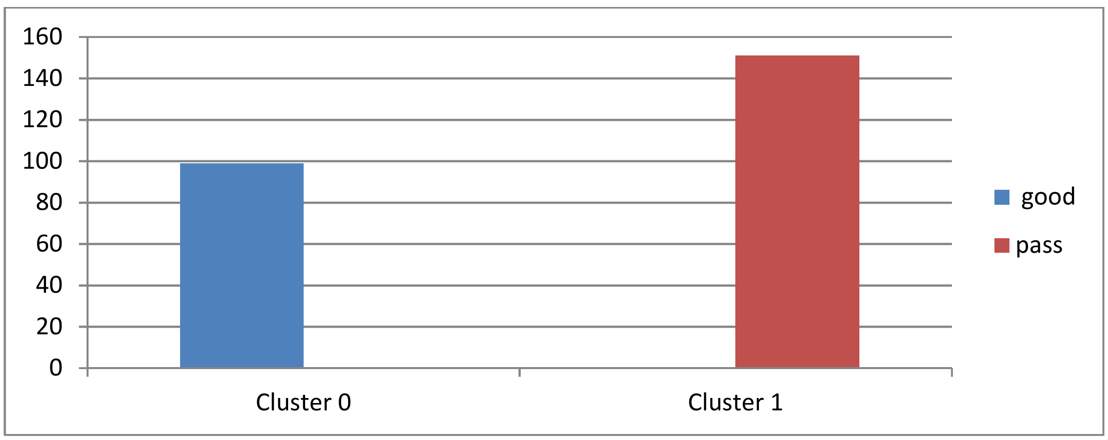



Figure 4-2: divided data into two clusters in the third year

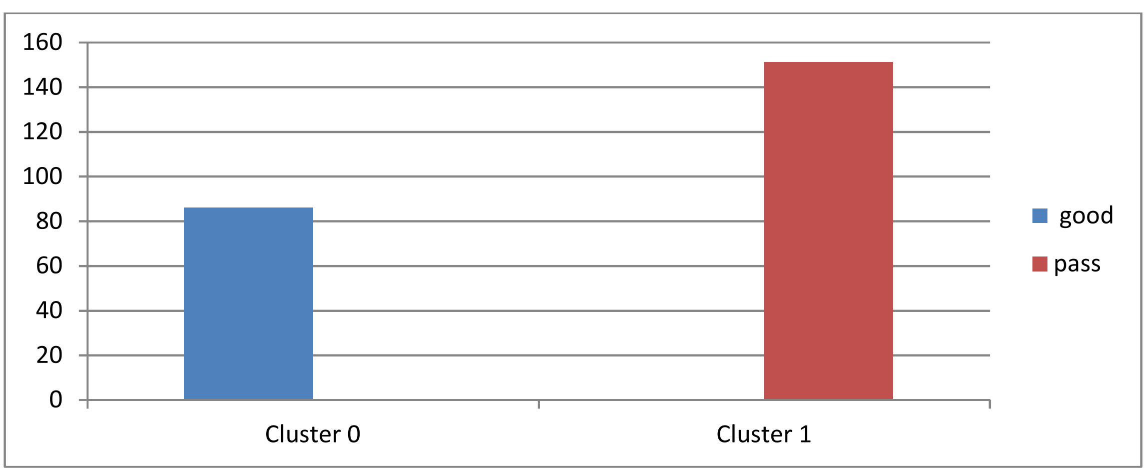



Figure 4-2: divided data two clusters in four year

### 4.3. results and knowledge

In this point intersection is used on the strong rules were generated in chapter three. The resulsts of intersection used to recommender students for improving their academic performance for each year study.

#### 4.3.1 Intersection rules in the first year

Rules of cluster number (0) in first year

Islam1=A S.Sudanese=A calculus=C App2=A==> GPA=Very good

Islam1=A Commun=C calculus=C App2=A ==> GPA=Very good

Arbic 1=A B.math=D calculus=C App2=A==>GPA=Very good

Arbic 1=A Compter=B calculus=C App2=A==> GPA=Very good

Arbic 1=A App1=A calculus=C App2=A==> GPA=Very good

S.Sudanese=A Compter=A calculus=C App2=A==> GPA=Very good

S.Sudanese=A Commun=C calculus=C App2=A==> GPA=Very good

S.Sudanese=A calculus=C operating2=A App2=A==> GPA=Very good

B.math=C operating1=A calculus=C App2=A==> GPA=Very good

Accounting =B Commun=C calculus=C App2=A==> GPA=Very good

operating1=A App1=A Eng2=C calculus=C==> GPA=Very good

operating1=A statis=C calculus=C App2=A==> GPA=Very good

operating1=A calculus=C operating2=A App2=A==>GPA=Very good

App1=A Eng2=C calculus=C App2=A==> GPA=Very good

App1=A calculus=C operating2=A App2=A==> GPA=Very good

⇩

calculus =C and App2=A

Rules of cluster number (1) in first year

B.math=D operating1=A App1=A calculus=D App2=A==> GPA= Good

B.math=D App1=A Pro1=D calculus=D App2=A ==> GPA=Good

B.math=D Eng2=D Commun=C calculus=D App2=A ==> GPA=Good

B.math=D Commun=C Pro1=D calculus=D App2=A==> GPA=Good

⇩

calculus=D and App2=A

Rules of cluster number (2) in first year

B.math=F Computer=F Pro1=F statis=F calculus=F==> GPA=Fail

B.math=F Computer=F accounting=F Pro1=F calculus=F==> GPA=Fail

B.math=F Computer=F accounting=F statis=F calculus=F==> GPA=Fail

⇩

B.math=F and Computer=F and calculus=F

- ✓ The above output clearly indicates that the performance of students having computer application II has significantly improved in Calculus. The grades of very good students from computer application II =A and Calculus=C. While in the case of students with non-computer application II =A and Calculus= f the grades are reverse.
- ✓ The grades of Good the students have been degraded in Calculus from C to D. It clearly indicates that the students are fronting difficulty in Calculus.
- ✓ The grades of fail the students have been degraded in Calculus from D to F with basic mathematics I and introduction of computer science the situation is a failure. It clearly indicates that the students are fronting falling in grades.

**4.3.2 Intersection rules in the second year**

Rules of cluster number (0) in second year

Pro.Methods 1=D MIS=C DB1=D Algebra=D==> GPA=Good

Pro.Methods 1=D MIS=C DB1=D APP.net=D==> GPA=Good

⇩

Pro.Methods 1=D and MIS=C and C DB1=D

Rules of cluster number (1) in second year

Pro.Methods 2=D MIS=D Algebra=F APP.net=F==> GPA=pass

Pro.Methods 1=D Pro. Methods 2=D MIS=D APP.net=F==> GPA=pass

⇩

Pro.Methods 2=D and MIS=D and APP.net=F

Rules of cluster number (2) in second year

Pro.Methods 1=F Math2=F DB1=F Algebra=F APP.net=F==> GPA=Fail

Math2=F DB1=F Algebra=F Discrete.math=F APP.net=F==> GPA=Fail

Pro.Methods 1=F Math2=F DB1=F Algebra=F Discrete.math=F==> GPA=Fail

⇩

Math2=F and Algebra=F and DB1=F

- ✓ The above output obviously indicates that the performance of students having the grades of good. Whether in this case the students have improved from programming methods I =D and management of information systems =C and database I=D.
- ✓ The grades of pass the students have been degraded in the management of information systems to D and fail in principles of using the internet. It clearly indicates that the students are meeting difficulty in these courses.
- ✓ The grades of fail the students have been a failure in basic mathematics II basic mathematics II= F with basic mathematics I=F and Algebra Geometry=F the situation is a failure. It clearly indicates that the students are fronting falling in grades.

### 4.3.3 Intersection rules in the third year

Rules of cluster number (0) in third year

T.internet 1=C HCI=D manage=C AI=D T.internet=D==> GPA=Good

T.internet 1=C manage=C AI=D Media=C T.internet=D==> GPA=Good

⇩

T.internet 1=C and manage=C and T.internet=D

Rules of cluster number (1) in third year

HCI=D OS=D Media=D OOP2=D Engineering 1=D==> GPA=pass

HCI=D OS=D Media=D T.internet 2=F Engineering 1=D==> GPA=pass

⇩

HCI=D and OS=D and Media=D and Engineering 1=D

- ✓ The above outcomes clearly indicate that the performance of students having the grades of good. Whether in this case the students have improved internet technologies I =C and fundamentals of management =C. While may increase the grades when improved the internet technologies II.
- ✓ The grades of pass the students have been degraded in the human and computer interaction, operating system, multimedia systems, software engineering I systems to D. It clearly indicates that the students are meeting difficulty in these courses.

**4.3.4 Intersection rules in four year**

Rules of cluster number (0) in four year

M.Network 1=C Engineering 2=D E commerce=C M.Network 2=D crypto 2=C==> GPA=Good

M.Network 1=C Engineering 2=D M.Network 2=D crypto 2=C visual=C==> GPA=Good

M.Network 1=C Engineering 2=D M.Network 2=D visual=C A.DB=C==> GPA=Good

M.Network 1=C E commerce=C M.Network 2=D crypto 2=C Media 2=D==> GPA=Good

M.Network 1=C E commerce=C M.Network 2=D visual=C Media 2=D==> GPA=Good

M.Network 1=C M.Network 2=D crypto 2=C visual=C Media 2=D ==> GPA=Good

⇩

M.Network 1=C and M.Network 2=D

Rules of cluster number (1) in four year

M.Network 1=D Engineering 2=D E commerce=B crpto 1=D M.Internet=D==> GPA=pass

M.Network 1=D Engineering 2=D crpto 1=D M.Internet=D ethical=C==> GPA=pass

M.Network 1=D Engineering 2=D crpto 1=D crypto 2=D ethical=C==> GPA=pass

M.Network 1=D simulation=D E commerce=B crpto 1=D crypto 2=D==> GPA=pass

⇩

M.Network 1=D and crypto 1=D

- ✓ The above output clearly indicates that the performance of students having the grades of good. Whether in this case the students have improved networks management I =C. While maybe increase the grades when improved the networks management II. It clearly indicates that the students are meeting difficulty in networks management courses.

- ✓ The grades of pass the students have been degraded in the cryptography and information, networks management I to D. It clearly indicates that the students are meeting difficulty in these courses.

### 4.3.5 Intersection rules without using hybrid recommendation techniques

B.math=D Pro. Methods 2=D APP.net=F==> Cumulative GPA=pass

B.math=D APP.net=F T.internet 2=F==> Cumulative GPA=pass

B.math=D operating1=A APP.net=F==> Cumulative GPA=pass

B.math=D operating2=C APP.net=F==> Cumulative GPA=pass

B.math=D APP.net=F Engineering 2=D==> Cumulative GPA=pass

B.math=D operating2=C APP.net=F==> Cumulative GPA=pass

Arbic 1=D MIS=D APP.net=F==> Cumulative GPA=pass

Commun=D Electricity=D APP.net=F==> Cumulative GPA=pass

Commun=D APP.net=F M.Network 2=D==> Cumulative GPA=pass

Eng3=D APP.net=F Analysis 2=D==> Cumulative GPA=pass

Eng3=D APP.net=F T.internet 2=F==> Cumulative GPA=pass

⇩

APP.net=F

- ✓ The above output clearly indicates that the performance of students having the Cumulative grades of pass. Whether in this case the students have weakness principles of using internet =F. In order to improve the performance of students from principles of using internet, some introductory courses need to be conducted at initial level or some courses can be taken.

Table 4-5: summary of intersection rule

| Recommendation based on hybrid recommendation techniques | | | |
|---|---|---|---|
| **Results of intersection rules** | | | **Recommendation** |
| **Course name** | **symbol** | **GPA** | |
| **First year** | | | |
| computer application II | A | V.good | effectiveness positive |
| Calculus | C | | |
| computer application II | A | Good | effectiveness positive |
| Calculus | D | | |
| basic mathematics I | F | Fail | effectiveness negative |
| introduction of computer science | F | | |
| Calculus | F | | |
| **Second year** | | | |
| programming methods I | D | Good | effectiveness positive |
| management of information systems | C | | |
| database I | D | | |
| programming methods II | D | Pass | effectiveness negative |
| management of information systems | D | | |
| principles of using internet | F | | |
| basic mathematics II | F | Fail | effectiveness negative |
| Algebra Geometry | F | | |
| Database I | F | | |
| **Third year** | | | |
| internet technologies I | C | Good | effectiveness positive |
| fundamentals of management | C | | |
| internet technologies II | D | | |
| human and computer interaction | D | Pass | effectiveness negative |
| operating system | D | | |
| multimedia systems | D | | |
| software engineering I | D | | |
| **Four year** | | | |
| networks management I | C | Good | effectiveness positive |
| networks management II | D | | |
| cryptography and information | D | Pass | effectiveness negative |
| networks management I | D | | |
| **Recommendation without using hybrid recommendation techniques** | | | |
| principles of using internet | F | Pass | effectiveness negative |

### 4.4 discussion of recommending a student to improve their performance

According to the results of strong rules was obtained in this study to be recommended for new students join in year study. The students must consider these recommendations to improve their academic performance as the flowing points.

- **In the first year:** the students when they enrolled in the first year study. To be conceding the courses of Calculus, the introduction of computer science, basic mathematics I lead to decrease the grade or failure.
- **In the second year:** the students before they join in the second year of study. Must be considering the courses of management of information systems, principles of using internet I lead to decrease the grade.
- **In third year:** the students before they join in the third year study. Must be considering the courses of human and computer interaction, operating system, multimedia systems, software engineering I lead to reduction the grade.
- **In the fourth year:** the students before they join in the third year study. Must be considering the courses in cryptography and information, networks management I lead to reducing the grade.

### 4.5 Summary

This Chapter present and discuss the results was implemented in chapter three. Hence, it covers the results and knowledge that have guided to students. It starts with a results of clusters, which gives an overview of the ratio of student's performance then. Several intersection rules are interpreted, which are mainly used to determine the decrease grades in a specific year and present the of rules intersection. Finally, present the recommendation results which are help students to improve their performance.

# CHAPTER FIVE

# CONCLUSION AND RECOMMENDATION

## 5.1 Conclusion

The study applied on real data as faculty of computer studies and statistics, department of information technologies, university of KORDOFAN. Some preprocessing phases are applied on data such as handle missing data, data transformation and attribute selected. Then applied hybrid recommendation techniques which are clustering based on collaborative filtering algorithm and Association rules algorithms to achieve the objective of study, thus applied Association rules algorithms on each cluster, of which leads to generated strong rules in each year study. The intersection is applied on strong rules which help for recommending students to care which courses are positive and negative effectiveness on GPA. The study also applied Association rules algorithms without using hybrid recommendation techniques. The experiments showed that the applied hybrid recommendation techniques are better.

### 5.2 Recommendation and future work

After the completion of this study, recommend the following points to enhance the educational section and to make it more beneficial and more powerful:

- Extend the work on all departments of the college of computer studies and statistics.
- Activate the optional courses to be selected online in a faculty of computer studies and statistics using collaborative filtering recommender system.
- Using data merging tools of the data warehouse to integrated academic data of student for all year’s studies.

# References


Aher, S. B. and L.M.R.J., L. (2012) 'Best Combination of Machine Learning Algorithms for Course Recommendation System in', *International Journal of Computer Applications (0975*, 41(6), pp. 2–11. doi: 10.5120/5542-7598.

Al-Badarenah, A. and Alsakran, J. (2016) 'An Automated Recommender System for Course Selection', *International Journal of Advanced Computer Science and Applications*, 7(3), pp. 166–175. doi: 10.14569/IJACSA.2016.070323.

Al-barrak, M. A. and Al-razgan, M. (2016) 'Predicting Students Final GPA Using Decision Trees : A Case Study', 6(7). doi: 10.7763/IJIET.2016.V6.745.

Badr, A., Din, E. and Elaraby, I. S. (2014) 'Data Mining : A prediction for Student â€™ s Performance Using Classification Method', 2(2), pp. 43–47. doi: 10.13189/wjcat.2014.020203.

Bambrah, C. *et al.* (2014) 'Mining Association Rules in Student Assessment Data', *International Journal of Advanced Research in Computer and Communication Engineering*, 3(3), pp. 2278–1021. Available at: www.ijarcce.com.

Bydžovská, H. (2013) 'Course Enrollment Recommender System', (x), pp. 312–317.

El-Halees, A. (2008) 'Mining Students Data To Analyze Learning Behavior : a Case Study Educational Systems', *Work*, (February). doi: 10.1504/IJTEL.2012.051816.

Felfernig, A. *et al.* (2014) 'Basic approaches in recommendation systems', in *Recommendation Systems in Software Engineering*, pp. 15–37. doi: 10.1007/978-3-642-45135-5_2.

Gadhavi, M. and Patel, C. (2017) 'STUDENT FINAL GRADE PREDICTION', 8(3), pp. 274–279.

Gamulin, J. ., Gamulin, O. . and Kermek, D. . (2014) 'Comparing classification models in the final exam performance prediction', in *2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2014 - Proceedings*, pp. 663–668. doi: 10.1109/MIPRO.2014.6859650.

Godara, S. and Yadav, R. (2013) 'Performance Analysis of Clustering Algorithms for Non-circular Sources', *Proceedings of IEEE International Conference on Acoustics, Speech and Siggnal Processing (ICASSP)*, 4(1), pp. 119–123.

Hameed, M. (2012) 'Collaborative Filtering Based Recommendation System: A survey.', *... Journal on Computer ...*, 4(5), pp. 859–876. Available at: http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=09753397&AN=82397440&h=gpJ9vFOnA8y/cUhcaiLchym8wVlha4MBxJi+gCJHlJZ4fXO50YdHmHdTVdde+o/lbnsSEHvxcAawX0EWFsyscw==&crl=c.

Jain, R. (no date) *Introduction to Data Mining Techniques*, *Iasri.Res.in*. Available at: http://www.iasri.res.in/ebook/expertsystem_o/c-10 Data Mining techniques.pdf.

Jiawei Han (2006) *Data Mining: Concepts and Techniques*. Diane Cerra.

Kadambande, A. *et al.* (2017) *Predict Student Performance by Utilizing Data Mining Technique and Support Vector Machine*. Available at: http://web.engr.illinois.edu/~hanj/bk2/slidesindex.htm.

Kumar, S. V. K. and Padmapriya, S. (2014) 'An Efficient Recommender System for Predicting Study Track to Students Using Data Mining Techniques', *International Journal of Advanced Research in Computer and Communication Engineering*, 3(9), pp. 7996–7998.

Mobasher, G., Shawish, A. and Ibrahim, O. (2017) 'Educational Data Mining Rule based Recommender Systems', 1(Csedu), pp. 292–299. doi: 10.5220/0006290902920299.

Mueen, A. (2016) 'Modeling and Predicting Students ' Academic Performance Using Data Mining Techniques', (November), pp. 36–42. doi: 10.5815/ijmecs.2016.11.05.

Osmanbegović, E. and Suljić, M. (2012) 'Data mining approach for predicting student performance’, *Journal of Economics and Business*, X(1), pp. 3–12.

Osofisan, A. O., Adeyemo, O. O. and Oluwasusi, S. T. (2014) ‘Empirical Study of Decision Tree and Artificial Neural Network Algorithm for Mining Educational Database’, 7(2), pp. 187–196.

Pazzani, M. and Billsus, D. (2007) ‘Content-based recommendation systems’, *The adaptive web*, (November), pp. 325–341.

Priya, K. S. (2013) ‘Improving the Student ’ s Performance Using Educational Data Mining’, 1685, pp. 1680–1685.

R.B, B. (2013) ‘“Importance of Data Mining in Higher Education System”’, *IOSR Journal of Humanities and Social Science*, 6(6), pp. 18–21. doi: 10.9790/0837-0661821.

Raut, A. B. (2017) ‘Students Performance Prediction Using Decision Tree Technique’, 13(7), pp. 1735–1741.

Schafer, J. *et al.* (2007) ‘Collaborative Filtering Recommender Systems’, *The Adaptive Web, LNCS 4321eb*, (April), pp. 291–324. doi: 10.19026/rjaset.5.4644.

Sorour, S. E. *et al.* (2014) ‘Predicting Students ’ grades Based on free style Comments Data by Artificial Neural Network’.

Sumitha, R. and Vinothkumar, E. S. (2016) ‘Prediction of Students Outcome Using Data Mining Techniques’, (6).

Thangavel, S. K. and Learn, S. (2017) ‘Student Placement Analyzer : A Recommendation System Using Machine Learning’, pp. 4–8.

Vialardi, C. *et al.* (2009) ‘Recommendation in higher education using data mining techniques’, *EDM’09 - Educational Data Mining 2009: 2nd International Conference on Educational Data Mining*, pp. 190–199. Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-84857488884&partnerID=tZOtx3y1.

Yadav, S. K. (2012a) ‘Data Mining : A Prediction for Performance Improvement of Engineering Students using Classification’, 2(2), pp. 51–56.

Yadav, S. K. (2012b) ‘Mining Education Data to Predict Student ’ s Retention : A comparative Study’, 10(2), pp. 113–117.