REPUBLIC OF TURKEY
FIRAT UNIVERSITY
THE GRADUATE INSTITUTION OF NATURAL AND APPLIED SCIENCES

EVALUATION OF SCHOOL ADMINISTRATORS BY USING DATA MINING TECHNIQUES

Usman Umar
(141137106)

Master Thesis
Department: Software Engineering
Supervisor: Asst. Prof. Dr. Murat KARABATAK

August-2016

REPUBLIC OF TURKEY
FIRAT UNIVERSITY
THE GRADUATE INSTITUTION OF NATURAL AND APPLIED SCIENCES

EVALUATION OF SCHOOL ADMINISTRATORS BY USING DATA MINING TECHNIQUES

MASTER THESIS

Usman Umar

(141137106)

Department: Software Engineering

Date Submitted to Institute : **26 July 2016**
Thesis Defense Date : **09 August 2016**

THESIS SUPERVISOR: **Assist. Prof. Dr. Murat KARABATAK (F.U)**
OTHER JURIES : **Prof. Dr. Asaf VAROL (F.U)**
Assoc. Prof. Dr. Muhammed Fatih TALU (I.U)

August-2016

DECLARATION

I hereby declare that the titled “Evaluation of School Administrators Using Data Mining Technique” thesis is my own research and prepared by myself. Except for the quoted lines, and single words that are quoted. It is being submitted for the Degree of Master of Science (in Software Engineering) at the Firat University.

DEDICATION

This thesis is dedicated to my father, who taught me that the best kind of knowledge to have is that which is learned for its own sake. It is also dedicated to my mother, who raised and taught me that even the biggest task can be accomplished if it is done patiently and one step at a time. I would like to thank the rest of my family members for their understanding, moral supports, encouragements, prayers, patience and all kind of support.

ACKNOWLEDGEMENTS

I will like to express my sincere appreciation and gratitude to my thesis supervisor Asst. Prof. Dr. Murat Karabatak for his advice, patience, guidance, kind support, directions and thorough supervision during my course and research work. I wish to express my gratitude to all individuals that have participated in this study in one way or the other.

I would like to also thank my housemates Abubakar Karabade, Abubakar Muhammed Dandala, Abubakar Muhammed Kolo, Kaloma Usman Majikumna and Usman Muhammed Taa. These people have always been there during my stressful time. I wish to express my gratitude to the Turkish people in Bolu and Elazig province for their support and hospitality.

Usman Umar

TABLE OF CONTENTS

Page No

ABSTRACT

EVALUATION OF SCHOOL ADMINISTRATORS BY USING DATA MINING TECHNIQUES

School leadership is a very important issue in developed countries where school administrators are trained according to the standard that school leadership should have. This thesis applies Data Mining (DM) techniques to evaluate school administrators' perceptions on Believes about the Principal-ship, Generalized Perceived Self-Efficacy, and Melbourne Decision Making. With the help of literature scale, data is collected in order to evaluate these three school administrators' attitudes. Later, the data is used and analysis is performed by applying data mining technique. The school administrators who work in Nigeria and Turkey are selected as the sample group of this thesis. For the evaluation in this thesis, as basic evaluation method of data mining techniques.

Firstly, in this studies attitude on occupation questionnaire, generalized self-efficacy questionnaire and Melbourne decision making questionnaire were applied to the school administrators who serve in Nigeria and Turkey. Data Mining is a field comprising many different techniques that can be used to acquire hidden and meaningful information from a large amount of data. In this thesis, the analysis and evaluation are performed by applying data mining methods to the obtained data set. Firstly, classification method is applied to estimate the demographics based on the answers given by school administrators. According to their responses, the administrators are can be estimated based on their country with the scale of **93.38%** accuracy rate. Secondly, clustering method is applied to the collected data and **88.97%** clustering performance results are obtained. Lastly, the analysis of the data is done by applying association rule method. After applying the association rules, relations between the questionnaire responses are detected. Consequently, many different rules are found with confidence value of **100%** and a maximum of **3.76** lift value. This suggests that the obtained rules are interesting and highly reliable. In conclusion, a questionnaire is applied and various attitudes of school administrators are analyzed using data mining techniques and remarkable results are obtained.

Keywords: Association Rule, Classification, Clustering, Data Mining, School Administrators.

ÖZET

VERİ MADENCİLİĞİ TEKNİKLERİ KULLANILARAK OKUL YÖNETİCİLERİNİN DEĞERLENDİRİLMESİ

Okul yöneticiliği, gelişmiş ülkelerde çok önemli bir konu olup, okul yöneticiliğinin sahip olması gereken standartlar belirlenerek okul yöneticileri bu standartlara göre yetiştirilmektedir. Bu tez çalışmasında, okul yöneticilerinin mesleğe ilişkin tutumları, öz yeterlikleri ve karar verme etkinliklerinin veri madenciliği (VM) teknikleri kullanılarak değerlendirilmesi amaçlanmaktadır. Okul yöneticilerini bu üç farklı tutumunu değerlendirmek için literatürde geliştirilmiş ölçekler yardımı ile veriler toplanmıştır. Daha sonra bu veriler, veri madenciliği teknikleri ile analiz edilmiştir. Çalışmanın örneklem grubunu, Nijerya ve Türkiye'de görev yapan okul yöneticileri oluşturmaktadır. Verilerin analizi aşamasında ise veri madenciliği teknikleri kullanılmıştır.

Tez çalışmasında öncelikle Türkiye ve Nijerya'da görev yapan okul yöneticilerine, *Mesleğe İlişkin Tutum Anketi*, *Genelleştirilmiş Öz Yeterlik Anketi* ve *Melbourne Karar verme Anketi* uygulanarak yöneticilerden gerekli bilgiler anket yolu ile toplanmıştır. Veri madenciliği, büyük ölçekli verilerden gizli kalmış ve anlamlı bilgileri elde etmek için bir dizi teknikler içeren bir çalışma alanıdır. Bu nedenle, elde edilen anket verilerinin analizi ve değerlendirilmesi aşamasında veri madenciliği yöntemleri kullanılmıştır. İlk olarak sınıflandırma algoritmaları uygulanmış ve okul yöneticilerinin verdiği cevaplara göre demografik özelliklerini tahmini yapılmıştır. Anketlere verilen cevaplara göre **%93.38** başarım oranı ile okul yöneticilerinin iki ülkeden hangisinde yönetici olduğu tahmin edilebilmektedir. İkinci yöntem olarak veri grubuna kümeleme yöntemleri uygulanmış ve elde edilen sonuçlara göre de **%88.97** başarım elde edilmiştir. Son olarak, birliktelik kuralı uygulanarak verilerin analizi yapılmıştır. Birliktelik kuralı uygulanarak, anketlere verilen cevaplar arasındaki ilişkiler tespit edilmiştir. Bunun sonucunda güven değeri **%100** ve maksimum **3,76** lift değerine sahip çeşitli kurallar elde edilmiştir. Bu da elde edilen kuralların ilginç ve güvenirliliğinin yüksek olduğunu göstermektedir. Sonuç olarak okul yöneticiliği ve yöneticilerinin çeşitli tutumları veri madenciliği teknikleri ile analiz edilmiş ve kayda değer sonuçlar elde edildiği görülmüştür.

Anahtar Kelimeler: Birliktelik Kural, Kümeleme, Okul Yöneticileri, Sınıflandırma, Veri Madenciliği.

LIST OF TABLE

Page No

LIST OF FIGURES

Page No

LIST OF SYMBOLS AND ABBREVIATIONS

ARFF : Attribute-Relation File Format
ARM : Association Rule Method
CART : Classification and Regression Tree
CSV : Comma Separated Values
DB : Database
DBMS : Database Management System
DM : Data Mining
DT : Decision Table
E-CRM : Electronic Customer Relationship Management
EM : Expectation Maximisation
FN : False Negative
FP : False Positive
IBK : Instance-Based learning
ICI : Incorrectly Classified Instances
ID3 : Iterative Dichotomiser 3
KDD : Knowledge Discovery in Databases
KNN : K-Nearest Neighbor
MBR : Memory-Based Reasoning
MCC : Multi-Class classifier
MDBC : Make Density Based Clusterer
N / A : Not Assigned
NB : Naïve Bayes
NCE : National Certificate of Education
NG : Nigeria
SL : Simple Logistic
SVM : Support Vector Machine
TN : True Negative
TP : True Positive
TR : Turkey
VFI : Voting-Feature-Intervals
VM : Veri Madenciliği
WEKA : Waikato Environment for Knowledge Analysis

1. INTRODUCTION

The Internet technology has made it possible to collect and evaluate huge amount of data from diverse community across the globe. Information is being dealt with each and every single day in people's lives, as in [1, 2] the term is described as "Information Age". Many technological tools are being used every day to access the information which generates huge stacks of data on the Internet. The data can be accessed anywhere using the Internet; however, reaching the most valuable and relevant information from the accessed data is a very important and challenging task. For example, a company may like to use the transaction behavior of certain customers in order to send a suggestion based on the categories of things the customers used to buy. But failing to save the customers' transactions records may cause problems when a company tries to analyze those customers. Database technology has been provided in order to help in saving a huge amount data electronically.

Database (DB), as defined in [3], is any structured collection of selected data or records. Nowadays, the state of dealing with huge amount of files that contain much information, connections between students' or people's records in many different schools, organizations, institutes, government offices, hospitals etc., the need for a specific field, environment or platform in which all these kind of records to be stored is highly demanded. Therefore, software technology that organizes saving and accessing the information is widely used. The formation of the database and managing or controlling it is an approach to Database Management System (DBMS). In DBMS, the input and storage of data differ from application programs that access the data. But, while using the classical file as [4] described, any small alteration in file structure or register pattern may result in the change of application programs and compilation again.

Also in [5], DBMS is defined as a computer software that allows its user to work with databases on a computer system, as the user of DBMS can form, remove, and alter a table within the database. Any created table in a database may contain information about a large number of people. The removal or deletion of any table may result in the loss of the entire information saved within that specific table. But the alteration of a table means either addition, deletion, rearrangement, or change of records or data in the table. The main issue arises if any information is to be drawn from the database with a large amount of information

gathered. This is where the field *data mining* comes into the thesis study, and to obtain what is needed simply and safely some procedures are followed.

The initiative of school as a place or an atmosphere where learning, research or honorary is held started back in time even before the Ancient Greece. In the school system, knowledge is taught systematically where scholars tend to teach to a certain group of pupils [6]. Knowledge was being obtained in religious centers because there were only religion-based knowledge. The scholars or teachers were the religion preachers and the students were both parents and children in the same environment. The location of schools was the same as the location where people go to worship.

Since then, students have been taught based on their level of knowledge. Mostly older ones are placed at a high level due to their experiences in the past. The students' group predicts their level of knowledge. The classes differ based on what religion the family or parents are from and schools had no, specific classroom teacher, board for writing, chairs, and mostly no even an exercise books. Everything was based on memorization because there wasn't a standard administration that organizes how the knowledge is to be passed and received.

Administration, in another word management, is defined by Cambridge dictionary as the way to arrange and control the operation of tasks or plans needed in an organization [7], also, can be described as activities taken by a group of people committed to run the school or organization to its better functions [8]. Administrations differ from one organization to another or from one school to another depending on what guidelines or regulations the organization, institute or the district is built on.

School administrators are individuals who obtained a certificate of bachelor degree or postgraduate degree. They manage the routine activities and often provide instructional leadership in schools. Their role include supervising staffs, making decisions that are effective within the school premises, controlling the budget or expenses of the school, and making sure that everything goes as expected.

The aim of this thesis is to analyze the data obtained from a questionnaire that is administered to some school administrators in Nigeria and Turkey. The analysis is done according to the steps and by applying the methods of data mining. The accuracies are found based on the removal process of some meaningless questions from the questionnaire after applying association rule. The relationships between the questions in the questionnaire about believes in principal-ship, self-efficacy, and Melbourne decision-making are predicted according to their perspectives. The prediction process is done using Apriori algorithm in association rule. The relationships could be either based on parts, based on inter-parts or both. The software of the algorithms that helps to analyze the obtained data is applied, evaluated and the results are discussed and presented in Section 4.

2. DATA MINING

As explained by Zaiane (1999), with the advent of machines (computers) and means for huge capacity digital storage databases, all sorts of data are being collected and stored. This depends on the capabilities of these computers in order to help sort through this massive information. However, these massive collections of data stored on disparate structures are rapidly becoming overwhelming. To access and use the information stored easily, a technique called Data Mining (DM) should be applied. Data mining can be defined as finding and revealing any information valuable to any institution, organization or establishment. Data mining is also popularly known as Knowledge Discovery in Databases (KDD), which means consequential extraction of implicit, unknown before and likely useful information from data in databases [1]. Data mining is used in fields like business, banking, insurance and electronic commerce etc. Simply, data mining is a way of attaining the most valuable data from a large amount of data or the practice of examining large pre-existing databases in order to generate new useful information [9, 10]. By considering the steps in data mining models, Leskovec et al. (2014) concluded that data mining resembles classic statistical methods. The classic statistical method is a well-organized and mostly applicable to summarized data; while data mining deals with multi-millions or even multi-billions of data stacks and many more variables. By using data mining, all the data generated in an institution that exists or could arise in the future can be accessed, categorized, analyzed and evaluated. Therefore, the formation of the mentioned achievements with certain data mining methods could be regarded as confidential information extraction process.

2.1. Method of Converting Data into Information

Nowadays, under information technology changes are being noticed always, not only in computer technology but also in data communication technologies [5]. The most drawing attention is the changes in getting cheaper products. Customers are getting familiar and being talented in using computer technology. Due to technology development nowadays, almost every piece of information is being saved in the digital environment. Owing to this, the following are three most overcoming questions about the saved data in digital environments:

- What are the importance of the saved data to the organizations?
- What are the advantages of the saved data to the organizations?
- How to obtain information from the saved data?

➢ To do some analysis of the gathered data, many different types of statistical and mathematical methods can be applied.

➢ To do such, new concept database and new analytical methods should be formed.

➢ To direct data, a data warehouse should be used, and to analyze and get valuable data, data mining is used.

2.2. Data Mining Application Fields

There is the widespread usage of data mining areas these days. For example, in marketing, banking and related areas such as insurance and e-commerce areas are widely used [4]. Below are some areas where data mining is used and purposes of its usage:

2.2.1. Marketing

- To identify customer's buying patterns behavior.
- To reveal customers demographic characteristics links.
- To analyze the shopping cart.
- To manage customer relationship.
- To evaluate customer.
- To guess (forecast) sales.

2.2.2. Banking

- Bringing out the hidden relationships between different financial indicators.
- Determination of credit card fraud and forgery.
- Determination of customer groups based on credit card payments.
- Evaluation of loan (credit) demand.

2.2.3. Insurance

- To forecasts on which customers will potentially purchase new policies.
- To detect fraudulent behavior.
- To detect risky customers' behavior patterns

2.2.4. Electronic Commerce

- Analysis of attack.
- Management of E-Customer Relationship Management (E-CRM) applications.
- Analysis of visitations to web pages.

2.3. Data Mining Process

Data mining has to be evaluated the due process way. The processes of data mining are: data cleaning, data integration, data reduction, data transformation, implementation of data mining algorithm, and interpretation of results and evaluation. Below is a diagram of how the process takes place.

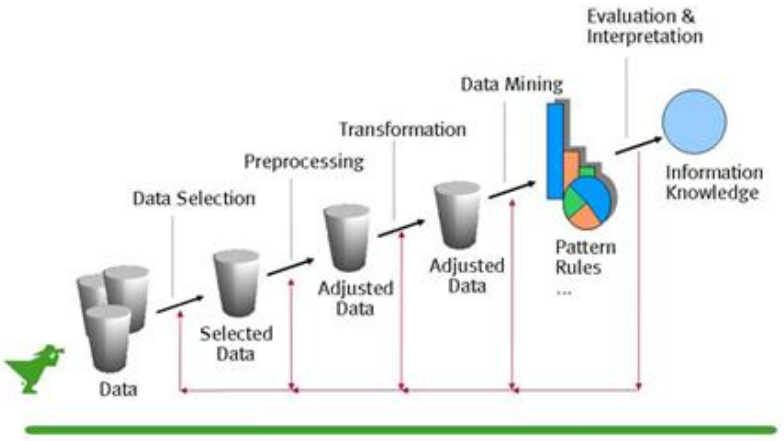

Figure 2.1. Data Mining Process

2.3.1. Data Cleaning

Data cleaning, also known as data cleansing, is defined by Zaiane (1999) as a process in which noised data or irrelevant data are removed from the collected data in the databases. This is due to the fact that handling data and databases possessed in a unique logical way is important. Fayyad et al. [11] noted that incomplete and missing data issues must be addressed and handled by either mapping or single naming conventionally whenever necessary. In some applications, the data to be analyzed can be seen not having the desired characteristics. For example, the incomplete data and unsuitable data that form data inconsistency can be encountered. The inconsistent data that takes place in databases are

defined as noise. In this case, the data needs to be cleaned up by assigning new data instead of incomplete data. The following methods are applied in order to clean up data:

- Disposal of records with missing values from the data set.
- A generally fixed value instead of missing value. The same value can be used instead of all missing value.
- The average of all variable data, specifically by calculating the average of class instance variable used in place of missing value.
- Assigning suitable guess to the data instead of the missing value.

2.3.2. Data Integration

Data integration is a process in which information is extracted and combined from different large databases [12]. The conversion of collected data from different databases into one data type is also called data integration [4]. To apply the data mining technique, if data warehouse was designed before, then, data integration must have been done. However, if there isn't a data warehouse designed before, then, a direct approach to data mining is applied to those data.

2.3.3. Data Reduction

Data reduction is also known as one of the problems in data mining. This problem is often seen as a way of pre-processing for information retrieval [13]. Sometimes in data mining applications, analysis takes a long time. If it is believed that after analyzation the result will not be changed, then, the amount of data or variable can be reduced. There are many ways to do the data reduction such as data merging or data cubing, size reduction, data compression, sampling, and generalization.

2.3.4. Data Transformation

In certain cases, adding data that are not transformed for the analysis may not be appropriate. While in some cases, the variables' mean and variance are significantly different from each other. As a consequence, the largest mean and variance of the variables pressure

on the others becomes more, and it significantly reduces their roles. Furthermore, very large and very small values of the variables prevent the analysis to be done in a robust way. Therefore, applying a transformation method to normalize or standardize the variables would be the appropriate way [4]. To normalize the variables min-max normalization, and to standardize the variables z-score standardization method can be used.

2.3.4.1. Min-Max Normalization

The min-max normalization method is applied to transform the data to a numeric value between 0 and 1. This method is used to determine the largest and the smallest numeric value in the obtained data. Likewise, the conversion of the others is appropriately based on the principle of obtained data. So, the below formula expresses how the conversion relation is done:

$$X^* = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{2.1}$$

Here X^* is the converted values, X is observation value, X_{min} is smallest observation value, X_{max} is largest observation value.

Example 2.1: In Table 2.1 the given X variable values are to be transformed using min-max normalization method. Firstly, the following values are assigned:

$X_{\min} = 30$, $X_{\max} = 62$.

Next, the calculation to the first member of X is done as below:

$$X^* = \frac{X - X_{min}}{X_{max} - X_{min}} = \frac{30-30}{62-30} = 0$$

This process is repeated to all the members and the values in Table 2.1 are obtained using min-max normalize transformation.

Table 2.1. Min-max normalize transformation

X	***X****
30	0.0000
36	0.1875
45	0.4688
50	0.6250
62	1.0000

2.3.4.2. Z-Score Standardization

Another conversion format commonly used in the statistical analysis is referred to as a z-score. This method is based on the average and standard error being transformed to new values. So, the conversion is expressed as:

$$X^* = \frac{X - \bar{X}}{\sigma_x} \quad (2.2)$$

Here X^* is the converted values, X is observation value, $\bar{X}$ is the data arithmetic mean, σ_x is standard deviation of observation value.

Example 2.2: We apply z-score to the previous example to transform the data. First, arithmetic mean needs to be found before the calculations proceed. Arithmetic mean is calculated as:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = 44.6$$

Secondly, the standard error of X series must be found before calculating the z-score standardization. The standard error is calculated as:

$$\sigma_X = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}} = 12.44$$

Lastly, the transformation of the first row is found as:

$$X^* = \frac{X - \bar{X}}{\sigma_x} = \frac{30 - 44.6}{12.44} = -1.1736$$

The same method is applied to find the rest of the observation values.

2.3.5. Implementation of Data Mining Algorithm

Before applying data mining method, the observed aforementioned process is first appropriately done for better results. After preparing and making the data ready then, data mining algorithms are applied to the relevant topic.

2.3.6. Interpretation of Results and Evaluation

After applying data mining algorithm on data, organized results are then presented to the relevant places. The results are often supported by graphics. For example, if a hierarchical

clustering model is applied, then the results are presented with special graphic named dendrogram.

2.4. Methods of Data Mining

In data mining, Fayyad et al. [11] explained that prediction and description are the high-level goals. Because of that, numerous methods and algorithms have been developed to help achieve these goals in data mining. Many of these methods were tried and tested techniques in machine learning, pattern recognition, and statistically based. Data mining methods can be basically grouped under three main headings. These are Classification, Clustering, and Association rule.

2.4.1. Classification

Classification in data mining, as illustrated [11], is to learn a function that classifies data items into one or more predefined classes. Classification in data mining is a method used frequently to uncover hidden patterns in databases [4] and a specific process is followed for the classification of data. Firstly, using a portion of the existing database for training purposes helps in forming the classification rule. Later, these rules, as [4, 14] described, help determine how to decide if a new situation occurs.

2.4.1.1. Classification with Decision Tree

A decision tree can be described as any efficient material or tool used to solve questions or problems in regression and classification [15]. The classification task is to predict the label or class for a given unlabeled point. It classifies the data (i.e., it constructs a model) according to the training-set and the values contained in a classifying attribute, and uses it to classify new data. Classification creates groupings in a dataset depends on the observation attributes' similarities. The goal in classification is that, previously unseen records should be designated to a class as factual as possible. As mentioned in [4], in statistics with application under machine learning, many decision algorithms are developed. Decision tree classification is mostly used because it's easy to understand and has affordable

implementation [16]. The foundation must be good in order to apply decision trees classification algorithms. Assuming two inputs as *X* and *Y* are defined. If the value of *X* is greater than 1, and in class1 the value of *Y* equals *B, Y* equals *A* and *Y* equals *C* are also in the same class as class2. Else, if *X* is greater than or equal to 1, it is then contained in class1.

Classification Process

Classification process uses a specific part of the database as a training set to form the classification rules by applying classification algorithms. Later, with the help of the rules decisions are made if new situations appeared.

Branching Criteria in Decision Tree

The most important of decision trees is to define the flow from the root to the branches or the ended leaf. The following algorithms can be categorized as:

- Entropy-based algorithm (ID3, C4.5)
- Classification and Regression Trees (CART)
- Memory based classification algorithms (KNN)

Quinlan (1986) developed many algorithms under decision trees, these include Iterative Dichotomiser 3 (ID3) and C4.5.

ID-3 Algorithm

In decision tree learning, Quinlan (1986) invented the ID3 (Iterative Dichotomiser 3) algorithm in order to create a decision tree from the dataset [17, 18]. The ID3 algorithm is used typically in machine learning and is used in natural language processing domains as well. The technique of a decision tree involves in the formation of a tree in order to model the classification process. After the tree is created, then it is enforced to each tuple inside the database and outcomes in a classification way for that tuple.

Entropy

Entropy can be defined as a measure of the disorder of a system; either the system is at uncertainty or at ignorance state [19]. In general, greater disorder means greater entropy. Let a set of *S* samples be given, and if to say that *S* is partitioned into two different intervals or layoffs as S_1 and S_2 by using a *T* boundary then, the information obtained after partitioning will be:

$$I(S,T) = \frac{|S_1|}{|S|} H(S_1) + \frac{|S_2|}{|S|} H(S_2) \quad (2.3)$$

The Entropy is then calculated according to the class distribution of the samples in the set. If *m* classes are given, the entropy of S_1 is illustrated as:

$$H(S_1) = -\sum_{i=1}^{m} p_i \log_2(p_i) \quad (2.4)$$

Here p_i is the class *i* probability in S_1

Entropy in Decision Tree

As Yalin (2013) explained, if the set of records obtained from the database for training is considered, the values obtained from the class attribute training set is$\{C_1, C_2, \ldots, C_k\}$, it is divided up into *k* classes. Regarding this class, an average amount of information may be needed. Here is how to calculate the probability distribution of *PT* class contained in set values of class *T*:

$$PT = \left(\frac{|C1|}{|T|}, \frac{|C2|}{|T|}, \ \ldots \ \frac{|Ck|}{|T|}\right) \quad (2.5)$$

$$H(T) = -\sum_{i=1}^{n} pi \ \log_2(pi) \quad (2.6)$$

Selection of Qualifications and Earnings Criteria for Branching

If by considering the separation of *T* into $T_1, T_2, \ldots, T_n$ which has attribute related to *X* attribute value. The class of *T* elements can be determined by the below formula.

$$H(X, T) = \sum_{i=1}^{n} \frac{|Ti|}{|T|} H(T_i) \quad (2.7)$$

The information obtained by partitioning of database *T* for test *X* is called gain measure. The below calculation is done as:

$$\text{Gain}(X, T) = \text{H}(T) - \text{H}(X,T) \quad (2.8)$$

C4.5 Algorithm

As an extension to decrease or eliminate the disadvantages of the ID3, C4.5 algorithm is a well-known algorithm that is used to create decision trees [20]. The C4.5 is an extension of the ID3 algorithm and it is used in order to overcome the disadvantages of the ID3. Any decision tree formed by the C4.5 algorithm can be used in classification studies, this is the reason C4.5 is also referred to as a statistical classifier [4]. To improve the ID3 algorithm number of changes are made to the C4.5 algorithm.

Handling Numerical Values

On the basis of *Y* attribute values, if considered the training samples are first sorted. There is only a finite number of the values, so let these be denoted as $\{v_1, v_2, \ldots, v_m\}$ in an arranged order. Any threshold value located between v_i and v_{i+1} will have effects the same as dividing the cases into the ones whose the value of the *Y* attribute is located between $\{v_1, v_2, \ldots, v_i\}$ and those whose value is in $\{v_{i+1}, v_{i+2}, \ldots, v_m\}$. Thus, only *m-1* possible splits on *Y* are found, and to get an excellent split they must be analyzed and evaluated systematically. The midpoint of each interval is usually been choose, as [4] described the midpoint is obtained as:

$$t_i = \frac{vi+vi+1}{2} \quad (2.9)$$

To calculate the representative threshold equation 2.9 is used. The smaller value of the threshold is chosen by the C4.5 algorithm as v_i for every $\{v_i, v_{i+1}\}$ interval, rather than choosing the midpoint by itself.

Unknown Attribute Values

The C4.5 algorithm accepts a principle which illustrated that a training sample obtaining unknown values are probabilistically distributed based on the known values' frequency relative. Below is a form of calculating the new gain criteria value:

$$F = \frac{\text{number of samples in a database with known value for a attribute given}}{\text{total number of samples in the data set}} \quad (2.10)$$

$$New\ Gain(\mathrm{X}) = F(H(T) - H(X, T))$$

2.4.1.2. Memory Based Classification

Memory-based classification is a member of nearest-neighbor, therefore it is useful in classification because [21]. Memory-Based Reasoning (MBR), is a particular version of the memory-based methods, it has been applied to the typical benchmark database tests. It has also been applied to the object identification or recognition and classifications of free text samples obtained from potentially large databases. The results after the classification method is applied seem to be better or comparable to the ones obtained by neural networks [21, 22]. Therefore, Duch et al. (1996) concluded that memory based methods classification are better suited to industrial or large scale applications than neural networks methods.

K-Nearest Neighbor Algorithm

K-Nearest-Neighbor (KNN) algorithm is considered to be one of the simpler algorithm as well as the most fundamental method in classification method. It should also be among the first choices for a classification analysis when there is little or there is no prior knowledge about the data distribution [4]. The basis of KNN classifier is on *Euclidean* distance between training samples that are specified and a test sample. Assuming x_i is an sample input containing p features as ($x_{i1}, x_{i2}, \ldots, x_{ip}$), and let n be the total number of the samples' input as ($i=1, 2, \ldots, n$) and p to be the total number of these features as ($j=1, 2, \ldots, p$). The Euclidean distance between sample x_i and x_l ($l=1, 2, \ldots, n$) can be illustrated as:

$$d(i,j)=\sqrt{\sum_{k=1}^{p}\left(x_{ik}-x_{jk}\right)^2} \quad (2.11)$$

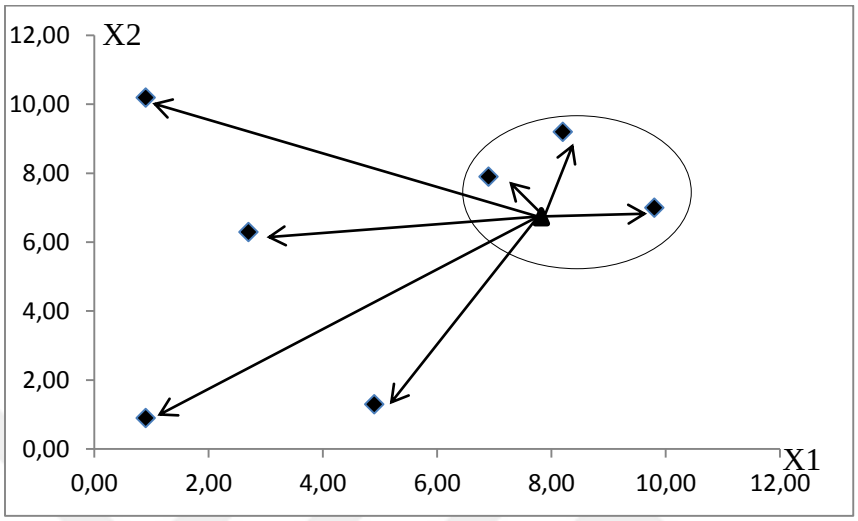

Figure 2.2. The observation values nearest to ▲ are k=3 neighbors

In K-nearest-neighbor algorithm, any cluster formed from observation value should be performed based on the following operations:

a) Determining K parameter.
b) Calculating the distance.
c) Determining the minimum distance.
d) Determination of the classes related to the selected lines.
e) Forming a new observation's class.

Weighted Voting

Weighted voting is described by Diplaris et al. [23] as one of the simplest methods in merging or combining homogeneous and heterogeneous models. The basis of weighted voting systems are on the idea of the voters are not all equal. Rather, it can be desired to identify the differences by providing some different amounts of weights regarding the outcome of a selection to the voters [4]. This is in contrast to the typical congressional process, in which it assumes that vote of each member can carry the same weight. The calculation of the weighted distance is based on the following relation:

$$d(i,j)' = \frac{1}{d(i,j)^2} \quad (2.12)$$

Here *d(i, j)* means the Euclid distance is between *i* and *j*. This distance is calculated for every class value and summing them to get the weighted voting value. A class value with highest weighted voting value is accepted as a class of new observation.

Statistical Classification Models

In [24], statistical classification is described as obtaining a set of discrete categories that can be assigned to a certain variable which is registered in an administrative file or any statistical survey. It can be used to present statistics and productions statistically. According to [25], in machine learning, the statistical classification model is used to identify categories of any set that belongs to new observations on the basis of training or educational set.

Conditional Probability

Conditional probability is a chance of an event to take place depending on the availability of a certain circumstances. As Triola [26] illustrated, a conditional probability of an event is a chance/probability attained together with some added information such that another event has occurred already. Considering two compatible (consonant) events *A* and *B*, there are common points between the event *A* and *B*. In this situation, it is said that $A \cap B \neq \emptyset$. For an event *B* to happen it totally depends on event *A*. The probability $P(B|A)$ is shown as follows:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \quad (2.13)$$

Based on the information given the compound probability is found as:

$$P(A \cap B) = P(A)P(B \mid A) \quad (2.14)$$

Then, the probability that two events will occur after one another like *A* and *B* equals to the product of the two events' probabilities. This can be written as:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

Which is explained as:

$$P(A \cap B) = P(B)P(A \mid B)$$

Rearranging the above equations produces the following:

$$P(A \mid B) = \frac{P(B)(A \mid B)}{P(A)} \quad (2.15)$$

Bayes Theorem

Bayes theorem describes the probability of an event to happen to watch the condition given that may relate to that event specifically [4, 27]. In calculating probability, Bayes theorem has a very important field. The classification process is possible using Bayes theorem. Let two events as T_1 and T_2 be consider. If to say that the inconsistency of these two events, i.e. $T_1 \cap T_2 = \emptyset$.

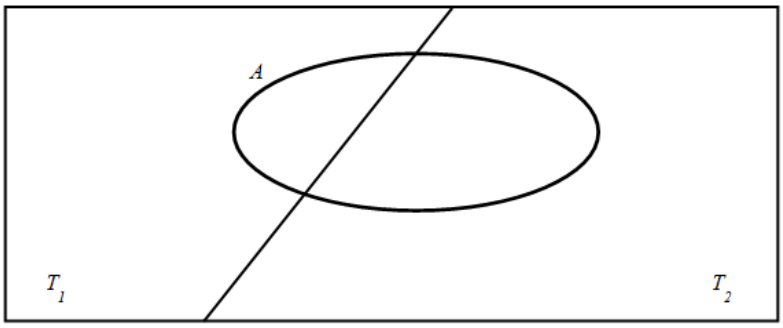

Figure 2.3. T_1, T_2, and A events

It can be said that event A is in event T_1 and T_2. For this, the conditional probability $P(T_1|A)$ is given as:

$$P(T_1 \mid \mathrm{A}) = \frac{P(A \mid T_1)}{P(A)} \quad (2.16)$$

As seen in Figure 2.3., event *A* happened in both event T_1 and T_2. The following relation for *A* can noted as:

$$A = (T_1 \cap A) \cup (T_2 \cap A)$$

Naturally, probability of an event *A* can be illustrated as:

$$P(A) = P(T_1 \cap A) + P(T_2 \cap A)$$

Using the above two equations, the following formula can be obtained:

$$P(T_1 \mid A) = \frac{P(A \mid T_1)}{P(T_1 \cap A) + P(T_2 \cap A)} \tag{2.17}$$

If the event is a set of clusters of $T_1, T_2, \ldots, T_n$ and the probability is not equal to zero, then the probability of event *A* happening in event T_j is shown as:

$$P(T_j \mid A) = \frac{P(A \mid T_j)}{\sum_{i=1}^{n} P(A \cap T_i)} \tag{2.18}$$

If $P(A \mid T_j) = P(A \mid T_j)P(T_j)$, then the following is obtained:

$$P(T_j \mid A) = \frac{P(A \mid T_j)P(T_j)}{\sum_{i=1}^{n} P(A \cap T_i)} \tag{2.19}$$

Bayes Classifier

Bayes classifier can reduce the probability of mis-classification to lesser value in statistical classification as explained in [4]. Suppose that two variables (*X*, *Y*) take values in $Rd * \{1, 2, \ldots, k\}$, where the class label of *X* is given as *Y*. Given that *r* value is assigned to the label *Y*, the conditional distribution of *X* will take the relation given as:

$$X \mid Y = r \sim Pr \text{ for } r = 1, 2, \ldots, k$$

where "~" means "is distributed as", and *Pr* represents the distribution of a probability. For example, if to say there is a stack of objects and suppose there are red round objects, the

hypothesis is “this object is an orange”. Here $P(H)$ is given as “priori probability”. That is what it is at first is known. But, because $P(X \mid H)$ probability is built under condition H then, it is evaluated as “posteriori probability” (i.e., if an X object is yellow and round then, the result is wil be “it is an orange”). Therefore, the Bayes relation is as follows:

$$P(H \mid X) = \frac{P(X \mid H)P(H)}{P(X)} \tag{2.20}$$

$$P(C_i \mid X) = \frac{P(X \mid C_i)P(C_i)}{P(X)} \tag{2.21}$$

To reduce the weight at the calculation process of $P(X \mid C_i)$ probability can be combined. To do that, the X_i values in the example are considered independent to be obtained:

$$P(X \mid C_i) = \prod_{k=1}^{n} P(x_k \mid C_i) \tag{2.22}$$

For classification of unknown X example, because the denominators in (2.20) are equal the comparison with the nominator must be done. Selecting the highest out of this and say the unknown example belongs to this class.

$$\underset{C_i}{\operatorname{argmax}} \{P(X \mid C_i)P(C_i)\} \tag{2.23}$$

The above used posteriori probability statement, *Maximum A Posteriori classification = MAP* is also known. In that case, as a result, because of (2.21), Bayes classifier can be used as below relation:

$$C_{MAP} = \underset{C_i}{\operatorname{argmax}} \prod_{k=1}^{n} P(x_k \mid C_i) \tag{2.24}$$

2.4.2. Clustering

As explained by Yalin in [4], clustering is the process of grouping dataset by considering the similarities between them. The clustering technique has been widely used in many different applications such as customer behavior analysis of purchased item, targeted

marketing, and a lot of others analysis [28]. So, it can be applied in many fields because of its features. For example, it is widely used in marketing research, pattern identification, image processing and in the analysis of spatial map data.

As such, according to [29], clustering method does not adopt class groups that are assigned before, except maybe while verifying to see if the clustering work well enough or not. When dealing with clustering the nearest neighboring algorithm and farthest-neighbour algorithm must be remembered, also known as a hierarchical method of clustering.

2.4.2.1. Cluster Analysis

Cluster analysis is defined in [30, 31], as the way of forming a group of abstract objects into classes of objects that are similar on the basis of gathered information in the data by illustrating/defining the relationship of objects. The basis of grouping objects in cluster analysis is on the information contained in the data by describing the objects or defining the objects' relationships. The aim of clustering analysis is that objects in a group should be identical or relevant to one another, and should be different from one another or unrelated to one another in the group at the same time. Figure 2.4 is an example of different cluster groups.

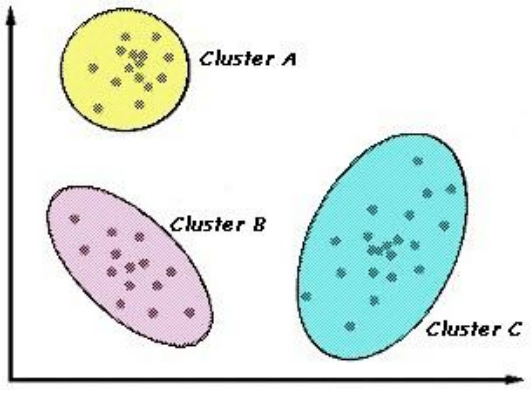

Figure 2.4. A view of cluster *A*, *B*, and *C*

2.4.2.2. Distance Measurements

As explained by Yalin (2013), each clustering problem is based on some kind of "distance" between points. That is why the distance between two points needs to be calculated. An observation value that is made of different kind of variables is called as *X*. For example, a matrix made of three variables and five observations are as in Figure 2.5

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \\ x_{51} & x_{52} & x_{53} \end{bmatrix}$$

Figure 2.5. An example of 3x5 *X* matrix

The 1st observation point location is (X_{11}, X_{12}, X_{13}) and the 2nd observation point location is (X_{21}, X_{22}, X_{23}). The distance between these two points is $d(1,2)$. The difference between every row to the other is defined as $d(i, j)$. D symmetric of distance symmetric can be shown in Figure 2.6.

$$D = \begin{bmatrix} 0 & & & & \\ d(2,1) & 0 & & & \\ d(3,1) & d(3,2) & 0 & & \\ d(4,1) & d(4,2) & d(4,3) & 0 & \\ d(5,1) & d(5,2) & d(5,3) & d(5,4) & 0 \end{bmatrix}$$

Figure 2.6. *D* symmetric of distance symmetric

There are many different types of distance measurements formula, three of them are given below:

a) *Euclidean distance*: is one of the most widely used distance measure in data mining field. Figure 2.7 below is an illustration of the method:

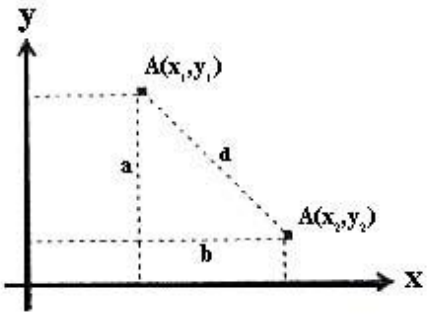

Figure 2.7. The d distance between two points

The distance between *A* and *B* is calculated using Euclid as:

$$d(A,B) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{2.25}$$

To generalize this relation, *i* and *j* points reach the following relation:

$$d(i,j) = \sqrt{\sum_{k=1}^{p} (x_{ik} - x_{jk})^2} \tag{2.26}$$

b) *Manhattan distance:* The distance if you had to travel along coordinates only, the total of absolute observation distance must be taken to do the operation:

$$d(i,j) = \sum_{k=1}^{p} (|x_{ik} - x_{jk}|) \quad i, j=1, 2, \ldots, n; k=1, 2, \ldots, p \tag{2.27}$$

c) *Minkowski distance:* is a metric in a normed vector space that can be considered as an integration of both the Euclidean and the Manhattan distance measures. The formula is:

$$d(i,j) = \left[\sum_{k=1}^{p} (|x_{ik} - x_{jk}|^m) \right]^{\frac{1}{m}} \quad i, j=1, 2, \ldots, n; k=1, 2, \ldots, p \tag{2.28}$$

Writing $m = 2$ in the formula to get Euclid distance

2.4.2.3. Hierarchical Clustering

Hierarchical clustering is described a way of producing a hierarchical series of nested clusters that range from the clusters at the bottom to an all-inclusive cluster at the top.

Unifying hierarchical method: Are methods that deal with separately handled clusters incrementally merging them together. The mostly considered algorithms are nearest neighbor algorithm and farthest neighbor algorithm.

K - Nearest Neighbor Algorithm

The k-nearest-neighbor classifier is an algorithm generally based on the Euclidean distance between the training samples that are specified and a sample test. Assuming x_i is an sample input containing p features as (x_{i1}, x_{i2} ,…, x_{ip}), and let n be the total number of the samples' input as ($i=1, 2, \ldots, n$) and p to be the total number of these features as ($j=1, 2, \ldots, p$). The Euclidean distance between sample x_i and x_l ($l=1, 2, \ldots, n$) can be illustrated as:

$$d(i,j)=\sqrt{\sum_{k=1}^{p}\left(x_{ij}-x_{jk}\right)^2}$$

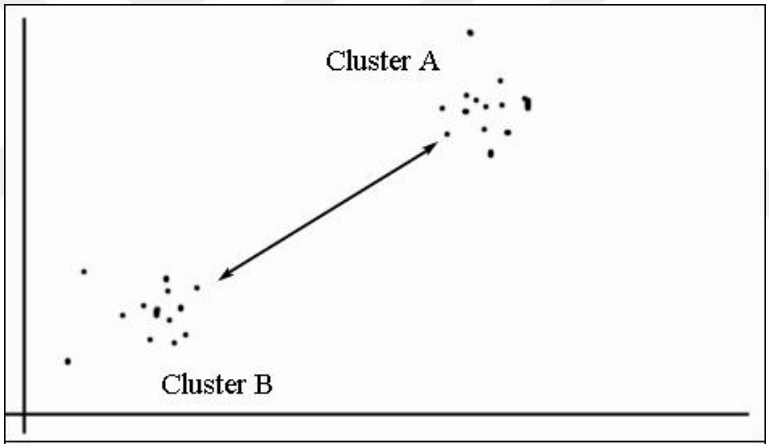

Figure 2.8. Nearest distance between two clusters

Farthest Neighbor Algorithm

This method is same as the nearest neighbor method, the only difference is, instead of calculating the shortest distance between two clusters rather the longest distance between them is being calculated as shown in Figure 2.9.

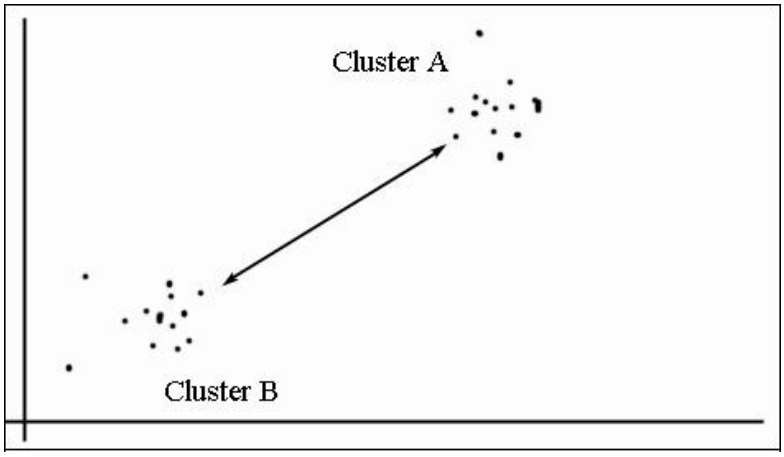

Figure 2.9. Farthest distance between two clusters

2.4.2.4. Non - Hierarchical Clustering

This method deals with non-overlapping groups that have no hierarchical relation among them. As Robbins (1994) illustrated, non-hierarchical methods demand not much computational effort when compared with hierarchical methods.

K – Mean Method

K-mean clustering method are frequently referred as a non-hierarchical clustering methods. Here, the aim is to minimize the mean error. If a space N is given, where space is in $\{C_1, C_2,\ldots,C_k\}$ format, divided into K clusters. So, $\sum n_k = N$ (k=1, 2, …, k), and C_k cluster's mean vector is calculated as M_k as shown below:

$$M_k = \frac{1}{n_k}\sum_{i=1}^{n_k} X_{ik}$$

The value of X_k for C_k cluster is the i^{th} example. The square error for C_k is the total Euclid distance between C_k example and its centroid. This error is called "intra-cluster change"

$$e_i^2 = \sum_{i=1}^{n_k}(x_{ik} - M_k)$$

Square error is total change inside the cluster, and is calculated as:

$$E_k^2 = \sum_{k=1}^{k} e_i^2$$

Algorithm: Before starting the *k*-mean algorithm, first, *k* cluster must be determined. After it is determined, then the observation value is assigned to all clusters and $C_1, C_2, \ldots, C_k$ is obtained. Then the following operations are done;

a) Every cluster's centroid is determined in $M_1, M_2, \ldots, M_k$ form.
b) $e_1, e_2, \ldots, e_k$ inter-cluster change is calculated to find total E^2_k
c) The distance between centroid value and observation value is found.
d) The above *b* and *c* steps are repeated if the condition is at state.

2.4.3. Association Rule

Association rule mining, as Chan [5], Jiang and Gruenwald [32] described, studies the frequency or repetition of items in the basis of finding all association rules by considering the minimum support and confidence threshold respectively. An Association Rule Method (ARM) is a data mining method used for analyzing the relationship of data contained in the database [4], determining situations that are related and can work together, where determining these relationships helps in attaining association rules. Association rules are found particularly in the marketing field. Applications that are named market cart analysis are based on this kind of data mining methods as [4] explained and with the help of market cart analysis, a certain probability can be applied that guesses if a customer purchases any product and detects a possible product added to the cart.

Association rule in data mining is a way of finding out the relationship(s) between one or more items that are in large number in the database [33], and Agrawal et al. [14] explained it as a way to find all the rules of association that fit both the minimum of support and confidence of transactions or constraints in the database.

2.4.3.1. Support, Confidence and Lift Measures

In Association Rule, *Support* is defined in [34, 35] as the percentage of records divided by the fraction of the records which contains the union of the two values that are contained in that database. In different transaction, whenever an item is encountered its count increases by one, while scanning is in progress in the database [36]. But a support count does not mean that its quantity is decreased anywhere. In market-cart analyzation, *support* and *confidence* are used to show the connection between the two or more products sold. A value called *support number* is used in these kinds of measure calculation. The *rule support measure* determines the percentage of connection that is repeated in all the transactions. Likely, *rule confidence measure* determines the possibility percentage of obtaining product *B* group if the customer gets product *A* group in the basket. The condition of customer that is taking product *A* group also takes product *B* group, i.e. the association rule is shown as: $A \rightarrow B$. Rule support measures can be calculated as:

$$support\,(A \rightarrow B) = \frac{number\ (A, B)}{N} \tag{2.29}$$

Here, *number*(*A*,*B*) shows the number of shopping number support in *A* and *B* product group together, and *N* is the total number of shopping done per store visit. Rule confidence measures determine the probability of purchasing product *A* and *B* at the same time, and it is calculated as:

$$confidence\ \ (A \rightarrow B) = \frac{number\,(A, B)}{number\,(A)} \tag{2.30}$$

In association rule, *lift* is defined by Ordonez, Bayardo and Agrawal in [37] and [38] as the fraction of the union of two records to the product of the two records. Below is the formula of *lift*:

$$lift\ \ (A \rightarrow B) = \frac{\sup port\,(A \cup B)}{\sup port\,(A)\ X\ \ \sup port(B)} \tag{2.31}$$

2.4.3.2. Apriori Algorithm

According to Agrawal and Srikant (1994), Apriori algorithm is an algorithm for association rule combinations that helps to achieve a feasible result from the databases and can be used also for frequent item set. As Agrawal et al. [14, 35, 39] explained, the term *association rule* in data mining can be illustrated as: Let the set of k binary attributes (items) be $I = \{i_1,i_2,...,i_k\}$, and let the set of transactions (database) be $D = \{t_1,t_2,...,t_k\}$. a unique *ID* is contained in every transaction made in D and it is the superset of the items in I. A rule is defined as the form $X \rightarrow Y$. here X, $Y \subseteq I$ and $X \cap Y = \emptyset$. Therefore the sets of X and Y items (for short itemsets) are called antecedent (i.e., left hand side or LHS) and consequent (right hand side or RHS) of the rule respectively. These key concepts are: frequent item sets (i.e., the sets of an item with minimum support, denoted by L_i for i^{th} item set), Apriori property (this property says that any subset of a frequent item set must also be frequent), join operation (i.e., a set of candidate of k-item sets must be created by integrating L_{k-1} item set with itself in order to find L_k).

Apriori algorithm can generate all potential large item-sets of $k+1$ from large k-item-sets by using joins on the large k-item-sets and it validates them in counter to the databases.

3. SCHOOL ADMINISTRATORS

School administrators are selected people with associate degree, undergraduate degree or postgraduate degree level. Their bases of selection is on severance and hardworking. In Nigeria, a person who is certified in National Certificate of Education (NCE) as the minimum or lowest teaching qualification could also be appointed as a school administrator [40]. The appointment of individuals with either of the certificates as a teacher does not require working experience but, a person has to work in the system enough to become a school administrator. Beale and Hall [41] explained the responsibility of school administrators in a condensed form as any activities held inside the school premises. To expand that, their roles consist of hiring or appointing teachers that are committed to teaching, professionals, and also (possibly) with high grades. They are also committed to appointing non-academic staffs with good manners and behavior, as well as managing their duties and activities. Moreover, it is among their duties to direct their staffs to work accordingly so that the school will become admirable both to the parents and to the students [42]. School administrators manage the routine activities and often provide instructional leadership in schools. They supervise their staffs, make sure students attend school on a daily basis, make decisions that affect the school positively, and control the budget or expenses of the school [43, 44]. The school management or the school administration often make the budget of the school and send it to the ministry for evaluation. After the ministry receives the budget of any school, it accepts and approves the requests of the school administration. Later then, the requests are sent together with a supervisor in order to make sure they are delivered. At the other part, the school principal assigns a personal to make sure anything delivered to the school is being recorded and saved in a safe place.

In every organization, there is always a person who is the last decision maker, the person who is superior and can give directives or orders to everyone else. This person can be called with different names such as head, director, chairman, chief, principal, leader, ruler, president, and so on. Every organization has its own term of naming the superior; for example, in high schools principal is the term used to describe a superior.

3.1. The School Principal

School principal is described in many words; however, it can shortly be said that school principal is a person responsible for the school's success or failure [45, 46]. Any school principal is expected to lead so that the school will reach its best targets in terms of knowledge and moral behavior acquirements. There is no school principal or any other leader as a person who can fulfill all the organization's tasks to greatness alone. As Spillane [45] illustrated, this kind of leadership must involve the collection of individuals from different districts of the organization in order to achieve such greatness. That is why the management must appoint the best possible individuals as staffs that will help in running different school districts. However, no matter how best or how many staffs serve, their duty is to give suggestions on how to improve the school activities for better quality outputs. Furthermore, it is the staffs' duty to execute the tasks based on the directives given by the school principal. The final decision of the suggested improvements on tasks rests on the shoulder of the school principal. If the long list of tasks listed by Telem and Buvitski [47] is considered, more than 13 different tasks are expected to be fulfilled by the school principal. However, Ekundayo [46] explained that these tasks can be reduced into five heads by marginalizing them, which makes the school principal be characterized as the planner, organizer, director, supervisor, and evaluator of the school system.

3.2. The School Administrators and other Staff Members

As mentioned above, the school principal needs staff that will help to get the organs of the school running to its best function. The staff include the school administrators and other staff members serving at the school. As illustrated in [48], the school as an organization is coordinating the tasks through the division of labor. According to Barnard [49], every individual of the staff has his own effectiveness (ability to reach the given target) and efficiencies (ability in reaching the state of satisfaction of individuals' motives through cooperation). Therefore, the staff members are required to perform the tasks stated to their districts according to the directives given by school principals, and give reports when required.

The effect of school and school environment to our lives may cause either success or failure. Therefore, any decision made or rule introduced on how the school activities should be exercised needs to be good, reliable, and strong for better and qualitative outcomes. Harris et al. [50] illustrated that the way to behave in the school premises matters. Adding to that, Boyd and others [51], mentioned that self-being does not work always in this kind of environments. Because people look to their superiors as their role models [50, 51]. This happens not only in schools but also in houses, neighborhoods, towns, cities and so on. For example, the attitudes of children are based on the behavior of the family members and the community members they are grown up with.

The aim of this thesis is to evaluate school administrators by using data mining techniques. That is why data mining, how data mining works, how to convert obtained data to useful information, what field data mining is being applied, how data mining is being processed, and the method used to mine data are explained. Also, school administrators, school principal, and other staff members, their roles and duties are explained. Along with the explanations, a questionnaire was applied to the school administrators in order to evaluate their perspectives of their leadership based on data mining techniques.

4. EVALUATION PROCESS OF SCHOOL ADMINISTRATORS

A questionnaire was used in order to gather information on the school administrators for the evaluation to get started. The questionnaire was applied to the school administrators in Nigeria and Turkey using a hardcopy. The evaluation process in this thesis consists of six phases. These phases are: about the questionnaire, methods, statistical findings, classification findings, clustering findings, and association rule findings. More about these phases are explained in the following subsections.

4.1. About Questionnaire

The questionnaire used in this thesis is a combination of questions collected from different papers. It's later designed to be questionnaire of four parts. The first part contains questions on participant's demographic information. In the second part, there are questions on believes about principalship. The third part consists of general perceived self-efficacy questions. The last part or the fourth part of the questionnaire contains Melbourne decision-making questions. For the first part, five questions (gender, age, rank, education level, and severance) were asked because the questionnaire was administered to school administrators. The obtained answers of the questions based on demographic information that help to know the person being analyzed. The second part of the questionnaire is composed of 14 questions used by Darling-Hammond et al. [52]. The first four questions are to measure the level of positive beliefs, the next four questions are to measure the level of negative beliefs, and the last six questions are to measure the level of commitment to the school management. For the third part of the questionnaire, the questions are based on Bandura's self-efficacy theory [53]. These questions were later developed by Jerusalem and Schwarzer [54]. The original questionnaire, having 10 questions, was translated into 31 different languages and the version of Aypay [55] is used in this thesis. The last or fourth part (i.e., Melbourne decision making) of the questionnaire is based on 31 questions where the first 22 are developed questions by Mann et al. [56], and the Turkish version used by Deniz [57] for validity and reliability study where the scale is divided into two sub-parts. The first part comprises six questions for determining self-esteem or self-respect in making a decision. The second part comprises 22 questions for measuring decision-making styles. The administrative strength

scale was developed by Kanungo and Menon [58] and was adapted to Turkish three-dimensional scale by Ersozlu [59]. For the first part of the questionnaire, the participants are asked in questions number 1, 3 and 4 to just select out from the options given, while the other two questions need to be filled by the participants. For the other three parts of the questionnaire, participants are asked to tick or select only one out of five options provided on the sheet. These options are Strongly Disagree, Somewhat Disagree, Neutral, Somewhat Agree and Strongly Agree respectively, which represent 5th Likert scale. For more about sample of the questionnaire applied (See Appendix A).

4.2. Methods

The analyses in this thesis are based on classification, clustering, and association rule. WEKA is the application used for all analysis methods. WEKA is a well-designed for easy navigation and it has many data mining algorithms for classification, clustering, and association rule methods. It also has many different parameters options for any algorithm chosen to do the analysis. If navigated, it will help to achieve many different results, clusters, or value of accuracies. The files supported for WEKA are files with Attribute-Relation File Format (`.arff`) or Comma Separated Values (`.csv`) file extensions. The questionnaire was applied containing five different options as mentioned before. However, during the analysis, some of the options are merged in order to make it three simpler, more understanding options, and also to reduce time processing for the algorithm. Firstly, the collected data from questionnaires are converted into the `.arff` file without tempering with any option, while the severance grouping method is organized in the file as 0-5, 6-10, 11-15, 16-20, 21-25, and 26-30. Later, the file is saved in a folder as `Data1.arff` waiting for classification, clustering, and association rule analyses. The same data file is used to form `Data2.arff` file but, the severance grouping is changed in this file to 0-10, 11-20, and 21-30. Another file is formed as `Data3.arff` with the same properties as the previous file with a change in the statements' options. The change made is by merging the two options nearest in meaning to reduce time consumption of the program. Options 1 and 2 (Strongly Disagree and Somewhat Disagree) are merged to make it option 1, option 3 (Neutral) is changed to be option 2, and option 4 and 5 (Somewhat Agree and Strongly Agree) are

merged to make it option 3. By using `Data3.arff` some `.arff` files are formed again but, instead of making changes, it is decided to separate the data file into three. These three data files are named as `Data4_1.arff`, `Data4_2.arff`, and `Data4_3.arff`. The formation of these files is based on the sections or parts (i.e., Believes about the principal-ship, Generalized perceived self-efficacy, and Melbourne decision making) in the questionnaire and the demographic information is added to each file. After the formation of data files above, it is decided again to form another data file by updating Data3.arff file, making the severance grouping to be 0-5, 6-10, and >10 then saved it as `Data4.aff`. As it was done to Data3.arff to formed next three sectional (i.e., Believes about the principal-ship, Generalized perceived self-efficacy, and Melbourne decision making) files, the same method is applied to form another three files and saved them as `Data5_1.arff`, `Data5_2.arff`, and `Data5_3.arff`.

In addition, the analysis of some questionnaires applied in Turkey was also done in this thesis. The Turkish questionnaire contains the same number of questions in the second and third part of the questionnaire applied in Nigeria. The questions positioning in the second part are the same as organized (listed) in the Nigerian questionnaire. Those in the third part are not organized as the Nigerian questionnaire. However, in the fourth part of the Turkish questionnaire, there are only 22 questions same as their Nigerian counterpart. Therefore, data files of the obtained feedback are formed in order to analyze the Turkish questionnaire. At first, a data file named `DataTR1.arff` is formed with the exact format of questions from the questionnaire. Later, the copy of the file is generated as `DataTR2.arff` but, with just 22 questions that are equal to the Nigerian questionnaire from the fourth part. Lastly, the next data file named `DataTR3.arff` is formed with the same as previous but rather with three options instead of five after merging the ones closer in meaning.

To conclude, the questionnaires applied in Nigeria and Turkey are merged together by considering the common questions they both contained. After combining these two different questionnaires, some data files are generated and the analysis was performed. The generated data files are named `NG-TR1.arff` and `NG-TR2.arff`. The first file contains five options but, the second one contains only three options as some options are merged. The following lists the files used and the explanation of the data they contained:

Data1.arff: *Nigerian questionnaire, age group is 18-25, 26-40, >40, severance group is 0-5, 6-10, 11-15, 16-20, 21-25, 26-30, all parts are contained, answer options are 1-5.*

Data2.arff: *Nigerian questionnaire, age group is 18-25, 26-40, >40, severance group is 0-10, 11-20, >20, all parts are contained, answer options are 1-5.*

Data3.arff: *Nigerian questionnaire, age group is 18-25, 26-40, >40, severance group is 0-10, 11-20, >20, all parts are contained, answer options are 1-3.*

Data4.arff: *Nigerian questionnaire, age group is 18-25, 26-40, >40, severance group is 0-5, 6-10, >10, all parts are contained, answer options are 1-3.*

Data4_1.arff: *Nigerian questionnaire, age group is 18-25, 26-40, >40, severance group is 0-10, 11-20, >20, only second part is contained, answer options are 1-3.*

Data4_2.arff: *Nigerian questionnaire, age group is 18-25, 26-40, >40, severance group is 0-10, 11-20, >20, only third art is contained, answer options are 1-3.*

Data4_3.arff: *Nigerian questionnaire, age group is 18-25, 26-40, >40, severance group is 0-10, 11-20, >20, only fourth part is contained, answer options are 1-3.*

Data5_1.arff: *Nigerian questionnaire, age group is 18-25, 26-40, >40, severance group is 0-5, 6-10, >10, second parts is contained, answer options are 1-3.*

Data5_2.arff: *Nigerian questionnaire, age group is 18-25, 26-40, >40, severance group is 0-5, 6-10, >10, third parts is contained, answer options are 1-3.*

Data5_3.arff: *Nigerian questionnaire, age group is 18-25, 26-40, >40, severance group is 0-5, 6-10, >10, fourth parts is contained, answer options are 1-3.*

DataTR1.arff: *Turkish questionnaire, age group is 18-25, 26-40, >40, severance group is 0-10, 11-20, >20, all parts are contained, Turkish based, answer options are 1-5.*

DataTR2.arff: *Turkish questionnaire, age group is 18-25, 26-40, >40, severance group is 0-10, 11-20, >20, all parts are contained, equivalent to Nigerian but complex, answer options are 1-5.*

DataTR3.arff: *Turkish questionnaire, age group is 18-25, 26-40, >40, severance group is 0-10, 11-20, >20, all parts are contained, equivalent to Nigerian not complex, answer options are 1-3.*

NG-TR1.arff: *Nigerian and Turkish combined questionnaires, age group is 18-25, 26-40, >40, severance group is 0-10, 11-20, >20, all parts are contained, answer options are 1-5.*

NG-TR2.arff: *Nigerian and Turkish combined questionnaires, age group is 18-25, 26-40, >40, severance group is 0-10, 11-20, >20, all parts are contained, answer options are 1-3.*

4.2.1. Classification Method

The classification method of analysis is being one of the methods used in this thesis to classify and find the best accuracy from the data collected. There are many algorithms in the classification panel which are grouped based on their classifiers' type. All performance results under classification method are based on the evaluation of the confusion matrix that comprises three cases: Sensitivity, Specificity, and Accuracy [27]. True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) values are used to calculate the cases [60]. TP, properly reported true positive test result in the data, FP, false positive is any error reported in the data indicating improper test results which are negative as positive, TN, properly reported true negative test result in the data, FN, false negative is any error reported in the data indicating improper test results which are positive as negative. Below are the formulas for calculating the *sensitivity*, *specificity*, and *accuracy*:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{4.1}$$

$$\text{Spesificity} = \frac{TN}{TN+FP} \tag{4.2}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+TN+FP} \tag{4.3}$$

The confusion matrix is a dimension of **NxN** table containing TP, TN, FP, and FN. Below is the illustration of **2x2** confusion matrix table:

Table 4.1. A 2x2 confusion matrix table

		Classified	
Actual		True Positive (TP)	False Negative (FN)
		False Positive (FP)	True Negative (TN)

There are different algorithms used for classification analysis such as Naïve Bayes, Simple Logistics, Multi Class Classifier, J48, and so on. The reason of applying more than one algorithm is to evaluate and discuss the results' differences.

4.2.2. Clustering Method

The other method used in this thesis is clustering method. This method helps in grouping objects according to the information obtained from the data which describes the objects or describes the relationships of the objects. The aim of clustering is that the similar objects or related objects to one another in a group will be grouped seperately and likewise different objects or unrelated objects to one another in a group will be grouped seperately.

4.2.3. Association Rule Method

Association rule of analysis under data mining is one of the methods used in this thesis as a method to analyze the data obtained from the applied questionnaire. The analyzation is done to obtain rules as results that tell more and give predictions about school administrators' perspectives. Association rule is used in order to point out the relationships between categories or questions in the dataset. There are six algorithms appeared in the "associations panel" and two of these algorithms are passive for the file used. Apriori algorithm is considered for the analysis because it was among the active ones, and some parameters are changed during analyzation.

4.3. Statistical Findings

As explained above, the questionnaire is formed with four sections. In the first section, some questions on demographic information are asked. Questions about school administrators' believes in principalship are asked in the second section. Questions about self-efficacy are asked in the third section. Questions about Melbourne decision-making are asked in the last section (fourth section). Upon the analyzed data, there are 79 participants in Turkey and 57 participants in Nigeria who had filled the questionnaire. The statistics of the questionnaire applied in Nigeria are as follows: Out of 100%, the analysis showed 66.67% are male and 33.33% are female. For *age* classification only 3.51% aged between 18 and 25, 68.42% aged between 26, and 40, 28.07% aged more than 40. For *ranking* classification, 52.63% of the participants are school principals, 31.58% are vice principals on administrative issues and 15.79% were vice principal on academic issues. The analysis

shows that 22.81% of the participants are graduated from National Certificate in Education, 63.16% are having bachelor degree certificate and 14.03% are administrators with postgraduate certificate. Lastly, 82.46% of the participants have served between 0 to 10 years, 15.79% have served between 11 to 20 years and 1.75% have served for more than 20 years. Table 4.2 shows the statistical results of the demographic information from both the Nigerian and Turkish questionnaire in details.

Table 4.2. Statistical findings about the participants

Gender	**Frequency**		**Percentage**	
	NG	**TR**	**NG**	**TR**
1) Male	38	68	66.67	86.08
2) Female	19	10	33.33	12.66
3) N / A	0	1	0	1.26
Age				
1) 18 - 25	2	0	3.51	0.00
2) 26 - 40	39	31	68.42	39.24
3) > 40	16	45	28.07	56.96
4) N / A	0	3	0	3.8
Rank				
1) School Principal	30	77	52.63	97.47
2) Vice Principal Administration	18	1	31.58	1.27
3) Vice Principal Academic	9	0	15.79	0.00
4) N / A	0	1	0	1.26
Education Level				
1) NCE / Undergraduate	13	8	22.81	10.13
2) Bachelor / Degree	36	55	63.16	69.62
3) Post-graduate	8	14	14.03	17.72
4) N / A	0	2	0	2.53
Severance				
1) 0 - 10	47	4	82.46	5.06
2) 11 - 20	9	38	15.79	48.10
3) > 20	1	29	1.75	36.71
4) N / A	0	8	0	10.13
Total	**57**	**79**	**100**	**100**

4.4. Classification Findings

Since this thesis aims at evaluating school administrators; therefore, the first step is to find the classification accuracies of their perspectives. To obtain the accuracies, different algorithms found in WEKA application are exercised. For the analysis process in WEKA, the selected algorithm can be executed using the standard parameters of the selected Test Options or by navigating and changing the inputs to obtain different results. For example,

the standard selected Test Option comes as Cross-validation, and the standard number of folds are 10. There are also default parameters if the name of the selected algorithm is clicked, such as debug, use-of-kernel-estimator etc. These parameters can also be changed by selecting either true or false. However, only the value of Cross-validation is changed to 20 for the analysis in this thesis while other parameters are left with their default inputs. As mentioned in Methods subsection, there are different files containing the data used to be uploaded to the WEKA for analysis purpose. The analysis of the above-mentioned files is performed with different algorithms. These algorithms are selected from different classifier function folders found in the application. Below is the list of the algorithms' names and folders they are being extracted from:

- Naïve Bayes (NB) from bayes folder
- Simple Logistic (SL) from function folder
- Instance-Based learning (IBK) from lazy folder
- Multi-Class-Classifier (MCC) from meta folder
- Voting-Feature-Intervals (VFI) from misc folder
- Decision-Table (DT) from rules folder
- J48 from trees folder

Even though most of the selected algorithms are not explained in this thesis but the algorithms are similar and have the same function with the ones explained before. The selection of different algorithms is due to the unavailability of simply known algorithms of data mining methods in the WEKA application. If the algorithms are taken into consideration, for NB the algorithm is explained in the second section of this thesis. For SL algorithm, the algorithm is based on regression model where the binary in this model is used to calculate a binary response probability according to one or more independent features. The IBK algorithm is a pattern recognition algorithm and implements the KNN algorithm method, where for a test instance instead of building a model it creates a prediction. For the MCC algorithm, it is an algorithm on the basis of Support Vector Machine (SVM) where the instances are classified into one or more classes like in the binary classification. SVM in machine learning are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. The VFI algorithm is said to be

simple algorithm where it also consider each future separately like NB classifier, and it is believed to produce higher accuracy in some cases than the NB classifier. In machine learning DT is an algorithm used for prediction in classification or regression models which resembles decision tree. The DT algorithm implement the C4.5 (extended version of ID3 algorithm) features to find classification result, and contains hierarchical tables for the purpose of visualization analysis. For J48 algorithm, it is also an algorithm that resembles the decision tree like DT, and implement C4.5 features as well. Unlike ID3, the J48 algorithm has additional features as it accounts for missing values and prune trees etc.

The listed classification algorithms are applied to the created files mentioned before. The classification results are given in a series based on the two countries the questionnaires are administered. These results are presented in tables for better understanding, and accuracies are provided based on the demographic information (selected as outputs). After the presentation of the classification results based on countries then, comparison of the two results is also given.

4.4.1. Classification Findings Based on Questionnaire Applied in Nigerian

The data obtained from the questionnaire applied in Nigeria is used to form 10 different WEKA data files for the analysis. These data files are uploaded to the application and classification algorithms are applied to them and the results are obtained. In the following tables, classification accuracies of the algorithms applied to the data files are given. Below is a table showing the highest accuracy of an output is illustrated in bold, showing which algorithm happened to find the highest accuracy value based on outputs.

Table 4.3. The percentage of the algorithms' accuracies of `Data1.arff` file

	NB	**SL**	**IBK**	**MCC**	**VFI**	**DT**	**J48**
Gender	50.88	66.67	49.12	49.12	**68.42**	49.12	47.37
Age	52.63	**64.91**	64.40	56.14	63.16	50.88	54.39
Rank	33.33	**49.12**	36.84	42.11	38.60	**49.12**	43.86
Educ. Level	43.86	56.14	43.86	59.65	54.39	43.86	**63.16**
Severance	29.85	33.33	43.86	**56.14**	49.12	**56.14**	38.57

As shown in the above table, the algorithm that found the highest value of accuracy for *gender* output is Voting Feature Intervals (VFI). The accuracy of the VFI algorithm appeared to be better than the accuracy of NB algorithm just like [61] illustrated. The accuracy value is 68.42% as the relative absolute error was 78.47%. Table 4.4 contains the accuracy values of `Data2.arff` obtained from different algorithms.

Table 4.4. The percentage of the algorithms' accuracies of `Data2.arff` file

	NB	**SL**	**IBK**	**MCC**	**VFI**	**DT**	**J48**
Gender	50.88	66.67	54.37	47.37	**68.42**	57.89	56.14
Age	50.88	59.65	**64.91**	57.89	61.40	54.39	**64.91**
Rank	35.09	49.12	33.33	40.35	30.60	**52.63**	43.86
Educ. Level	38.60	56.14	43.86	54.39	54.39	43.86	**66.67**
Severance	71.93	**82.46**	82.46	80.70	**82.46**	40.70	**82.46**

The accuracy results in Table 4.4 above showed improvements in *rank*, *education level*, and *severance* outputs due to the alteration made in severance range. It can be concluded that Table 4.4 has the best accuracy results from the analyzed data files. To justify that, in following tables, the results of the same algorithms are also presented and compared with Table 4.3 which proves that Table 4.4 presents the better results.

Table 4.5. The percentage of the algorithms' accuracies of `Data3.arff` file

	NB	**SL**	**IBK**	**MCC**	**VFI**	**DT**	**J48**
Gender	45.61	**66.67**	49.12	43.86	42.11	57.89	47.37
Age	64.91	**66.67**	63.16	**66.67**	61.40	56.14	59.65
Rank	36.84	**52.63**	43.86	42.11	35.09	**52.63**	38.60
Educ. Level	50.88	**57.89**	52.63	45.61	47.37	52.63	56.14
Severance	63.16	78.95	73.68	63.16	71.93	77.19	**82.46**

The alteration of severance range appeared to produce better results, especially in the last three outputs if Table 4.3 and Table 4.4 are examined. However, this method worked only on `Data2.aff` data file while the analysis was performed. Because some alterations were also made to merge some options and see whether or not there will be better results found. Thus, by examining the results in Table 4.5, there is no improvement in the Rank output.

The same situation arose after changing the severance range in `Data3.arff` data file to form the `Data4.arff` data file. The table below contains the results of the analysis and when compared with the above table there is an improvement only in the *age* output.

Table 4.6. The percentage of the algorithms' accuracies of `Data4.arff` file

	NB	SL	IBK	MCC	VFI	DT	J48
Gender	42.11	**66.67**	49.12	38.60	36.84	57.89	45.61
Age	66.67	61.40	64.91	**71.93**	59.65	68.42	49.12
Rank	33.33	49.12	40.35	49.12	33.33	**52.63**	38.60
Educ. Level	52.63	56.14	50.88	45.61	49.12	50.88	**57.89**
Severance	**45.61**	35.09	40.35	33.33	**45.61**	28.07	**45.61**

As seen in previous tables, different algorithms and different demographic information statements are used as outputs to find accuracies. The selected parameter under test option is cross-validation with 20 folds as mentioned before. After observing the results of the data files in previous tables, the best from `Data2.arff` file, for *gender* output the VFI algorithm found the highest accuracy as 68.42%. For *age* output, the SL algorithm found the highest accuracy as 64.91%. For *rank* output, the SL and DT algorithms both found the highest accuracy as 49.12 %. For *education level* output the J48 algorithm found the highest accuracy as 63.16%. For *severance* output, the MCC and DT algorithms both found the highest accuracy as 56.14%. This process is repeated to the other data files (`Data2`, `Data3`, and `Data4.arff`) and highest accuracies are boldly marked in the given tables.

The same method is applied to the rest of the data files that are separated based on of the questionnaire. These files are `Data4_1.arff`, `Data4_2.arff`, `Data4_3.arff`, `Data5_1.arff`, `Data5_2.arff`, and `Data5_3.arff` data files. Some accuracies of the algorithms to the files are also presented in following tables:

Table 4.7. The percentage of the algorithms' accuracies of `Data4_1.arff` file

	NB	SL	IBK	MCC	VFI	DT	J48
Gender	54.39	61.40	54.37	49.12	56.14	**64.91**	54.39
Age	**64.91**	**64.91**	52.63	63.16	56.14	63.16	**64.91**
Rank	43.86	43.86	47.37	**50.88**	42.11	49.12	49.12
Educ. Level	54.39	57.90	52.63	57.90	47.37	**63.16**	50.88
Severance	71.93	78.95	73.68	56.14	63.16	80.70	**82.46**

By observing the accuracy values in Table 4.7 above, it is noticed that the results started with incredible values especially for the Severance output. For the other outputs, the results are high as they are all above 50%. In contrast, the remaining two data files of `Data4` show the incremental values of accuracy.

Table 4.8. The percentage of the algorithms' accuracies of `Data4_2.arff` file

	NB	SL	IBK	MCC	VFI	DT	J48
Gender	50.88	63.16	59.65	54.39	47.37	**66.67**	61.40
Age	59.65	**66.67**	**66.67**	52.63	40.35	63.16	**66.67**
Rank	57.90	**59.65**	49.12	57.90	47.37	56.14	56.14
Educ. Level	54.39	63.16	50.88	52.63	33.33	**66.67**	63.16
Severance	73.68	**82.46**	78.95	63.16	66.67	80.70	**82.46**

If Table 4.8 above is examined, the increment of some accuracies can be noticed compared to the accuracies in Table 4.7. The increment in *gender* and *age* outputs are about 1.76% while in Rank output had the highest percentage value of 8.77% increment. For the *education level* output, the increment is 3.51% while the *severance* shows the same value.

Table 4.9. The percentage of the algorithms' accuracies of `Data4_3.arff` file

	NB	SL	IBK	MCC	VFI	DT	J48
Gender	45.61	**66.67**	57.90	45.61	49.12	63.16	54.39
Age	64.91	64.91	**77.19**	68.42	66.67	54.39	68.42
Rank	29.82	**47.37**	43.86	29.82	19.30	45.61	42.11
Educ. Level	47.37	**64.91**	50.88	43.86	36.84	50.88	57.90
Severance	64.91	80.70	80.70	61.40	66.67	80.70	**82.46**

After comparing Table 4.8 and Table 4.9, it is noticed that *gender* output accuracy values are same but found by different algorithms. The *severance* output is also the same but found both by the previous and another algorithm. For *age* output, the accuracy one of the algorithm from Table 4.8 shows increment in Table 4.9 while the other decreases. For the *rank* and *education level* outputs, the highest accuracy values were found by different algorithms but are decreased compared to the previous values.

This method was applied to the other sub-file of `Data5` file but the results found are no better than the results of sub-file of `Data4`. The highest accuracy value among the files is 73.68% which is found by Instance-Based learning (IBK) algorithm from `Data5_3.arff` data file. The differences in accuracy are often changed on the basis of algorithms and outputs. The confusion matrix of the results for the questionnaire applied in Nigerian is not given in this thesis. This is because the accuracy results of sensitivity and specificity were very low.

4.4.2. Classification Findings Based on Questionnaire Applied in Turkish

The data obtained from the questionnaire applied in Turkey is used to form four different WEKA data files for the analysis. These data files are uploaded to the application and classification algorithms are applied to them and the results are obtained. In the following tables, classification accuracies of the algorithms applied to the data files are given. The highest accuracy of an output is illustrated in bold, showing which algorithm happened to find the highest accuracy value based on outputs. Table 4.10 below shows the results from analysis on `DataTR1.arff` file.

Table 4.10. The percentage of the algorithms' accuracies of `DataTR1.arff` file

	NB	SL	IBK	MCC	VFI	DT	J48
Gender	78.21	**87.18**	79.49	82.05	84.62	85.90	**87.18**
Age	67.11	**72.37**	61.84	57.89	**72.37**	68.42	**72.37**
Rank	94.87	98.72	96.15	96.15	**100**	98.72	98.72
Educ. Level	59.74	62.34	49.35	64.94	64.94	58.44	**71.43**
Severance	61.97	61.97	46.48	64.79	47.89	64.79	**69.01**

The Table 4.10 shows the results of the Turkish questionnaire with much higher accuracy value when compared with the Nigerian questionnaire. If the highest accuracy value of *gender* output found by the Simple Logistic (SL) and J48 algorithm is considered, the accuracy value is 87.18%. For *age* output, the highest value of accuracy is found as 72.37% by Simple Logistic, Voted Feature Intervals (VFI), and J48 algorithms. The results in Table 4.10 also show the highest accuracy value as 100% from the *rank* output. But the accuracy value falls to 71.43% and 69.01% as the outputs are being changed to *education level* and *severance* respectively. Table 4.11 is a confusion matrix table for the best algorithm solution based on sensitivity and specificity accuracies.

Table 4.11. Confusion matrix of J48 algorithm based on *age* output

		Classified value		
		18-25	**26-40**	**>40**
Actual value	**18-25**	0	0	0
	26-40	0	22	9
	>40	0	12	33

The confusion matrix table given above is based on the values found by J48 algorithm for Age output. This is because, even though VFI found the highest accuracy, J48 algorithm has the best sensitivity and specificity values calculated as 70.97% and 73.33% respectively. Table 4.12 below is an extended version of the above used table containing values of accuracy obtained from different classification algorithms.

Table 4.12. The percentage of the algorithms' accuracies of `DataTR2.arff` file

	NB	**SL**	**IBK**	**MCC**	**VFI**	**DT**	**J48**
Gender	78.21	**87.18**	76.92	75.64	84.62	**87.18**	**87.18**
Age	65.79	**76.32**	63.16	65.79	73.68	68.42	75.0
Rank	93.59	98.72	97.44	97.44	**100**	98.72	98.72
Educ. Level	63.64	59.74	49.35	58.44	63.64	49.35	**68.83**
Severance	60.56	**69.01**	50.70	57.75	49.30	63.38	67.61

The results in this new version of Turkish data file show increments for *age* output compared to Table 4.10. The *gender*, *rank*, and *severance* outputs stay the same while the

education level output falls. The stability of the outputs may not always mean that the accuracy value is found with the same algorithm(s). Sometimes different algorithm manage to find the same values of the accuracy found by another algorithm previously. Additionally, other algorithms may also find the same highest values previously found by certain algorithms. For example, in Table 4.10 the highest accuracy value found for *gender* output is 87.18% and it is found by SL and J48 algorithms. But, in Table 4.12 the DT algorithm also managed to find the same value while the SL and J48 algorithms remained constant. Table 4.14 is a confusion matrix table for the best algorithm solution based on sensitivity and specificity accuracies.

Table 4.13. Confusion matrix of SL algorithm based on *age* output

		Classified value		
		18-25	**26-40**	**>40**
Actual value	**18-25**	0	0	0
	26-40	0	25	6
	>40	0	12	33

The confusion matrix table given above is based on the values found by SL algorithm for *age* output. This is because, even though VFI found the highest accuracy, SL algorithm has the best sensitivity and specificity values calculated as 80.65% and 73.33% respectively. Table 4.15 shows the results from analysis on `DataTR3.arff` file.

Table 4.14. The percentage of the algorithms' accuracies of `DataTR3.arff` file

	NB	**SL**	**IBK**	**MCC**	**VFI**	**DT**	**J48**
Gender	79.49	**87.18**	79.49	66.67	82.05	**87.18**	**87.18**
Age	64.47	**75.00**	61.84	73.68	76.32	59.21	80.26
Rank	93.59	98.72	97.44	98.72	**100**	98.72	98.72
Educ. Level	64.94	68.83	54.55	46.75	50.65	59.74	**64.94**
Severance	69.01	**69.01**	52.11	52.11	54.93	59.16	74.65

After examining the results in Table 4.14 above, some of the accuracy result values are the same as in Table 4.12. However, for *age* and *education level* outputs the accuracies dropped. For *gender*, *rank* and *severance* outputs, the same algorithms found the same

accuracy values as in Table 4.12 above. The same SL and J48 algorithms found the highest accuracy values for *age* and *education level* outputs. However, these results represent a less accuracy if compared with the results in Table 4.12. Table 4.15 is the confusion matrix table for the best algorithm solution based on sensitivity and specificity accuracies.

Table 4.15. Confusion matrix of SL algorithm based on *gender* output

		Classified value		
		18-25	**26-40**	**>40**
Actual value	**18-25**	0	0	0
	26-40	0	26	5
	>40	0	14	31

The confusion matrix table given above is based on the values found by SL algorithm for *age* output. This is because, even though VFI found the highest accuracy, SL algorithm has the best sensitivity and specificity values calculated as 83.87% and 68.89% respectively.

4.4.3. Findings of Combined Nigerian and Turkish Questionnaires

The data obtained from the questionnaire applied in Nigeria and Turkey are used to form 2 different WEKA data files for the analysis. These data files are uploaded to the application and classification algorithms are applied to them and results are obtained. In the following tables, classification accuracies of the algorithms applied to the data files are given. The highest accuracy of an output is illustrated in bold, showing which algorithm happened to find the highest accuracy value based on outputs. Table 4.16 below shows the results from analysis on NG-TR1.arff file.

Table 4.16. The percentage of the algorithms' accuracies of `NG-TR1.arff` file

	NB	SL	IBK	MCC	VFI	DT	J48
Gender	66.67	77.04	69.63	67.41	65.19	75.56	**78.16**
Age	61.65	**72.18**	66.17	58.65	71.43	**72.18**	71.43
Rank	65.93	**78.52**	77.04	62.96	69.63	74.81	74.07
Educ. Level	50.75	60.45	48.51	62.69	51.49	61.94	**67.16**
Severance	59.38	72.66	60.16	49.23	61.72	74.22	**75.78**
Nationality	91.18	87.50	84.56	80.88	85.29	**93.38**	88.24

The results in Table 4.16 indicate that the combination of data obtained from Nigerian and Turkish questionnaires can produce high accuracy value. In Table 4.16, the highest result found only by J48 is 67.16% accuracy for *education level* output. For *age* output, two algorithms (i.e., SL and DT) found 72.18% as the highest value accuracy. Also, only J48 found 75.78% as the highest accuracy value for *severance* output. For *gender* output, the accuracy value found by J48 is 78.16% making it almost close to the value found by SL for *rank* output as 78.52% accuracy. For the *nationality* output, the accuracies found are stunning values with all exceeding 80% and DT algorithm obtained the highest value as 93.38% accuracy. Due to the high value of accuracy found by DT algorithm, Table 4.17 is a confusion matrix that shows the number of correctly classified as well as the incorrectly classified instances.

Table 4.17. Confusion matrix of DT algorithm based on *nationality* output

		Classified value	
		NG	**TR**
Actual value	**NG**	53	4
	TR	5	74

The confusion matrix given in Table 4.17 indicates that the sensitivity, specificity, and accuracy values are incredible when calculated. As a result, sensitivity is 92.98%, specificity is 93.67% and the accuracy value is 93.38% as mentioned. Considering the results above, the analysis is performed to the other `NG-TR2.arff` data file too, and the following table shows the analysis results of the file.

Table 4.18. The percentage of the algorithms' accuracies of `NG-TR2.arff` file

	NB	SL	IBK	MCC	VFI	DT	J48
Gender	59.26	**78.52**	69.63	65.93	54.81	76.30	74.07
Age	61.65	72.93	60.90	68.42	68.42	69.17	**73.68**
Rank	67.41	**80.74**	74.07	73.33	67.41	**80.74**	76.30
Educ. Level	50.75	61.94	55.22	47.01	45.52	58.21	**64.93**
Severance	68.75	71.88	57.81	47.66	61.72	73.44	**75.00**
Nationality	**90.44**	**90.44**	**90.44**	83.09	87.50	89.71	87.50

After comparing the results of Table 4.18 and that of Table 4.16, both increment and decrement happened. For *gender*, *age,* and *rank* outputs there was increment in values of the accuracy compared to the ones in Table 4.16. However, for *education level*, *severance,* and *nationality* outputs there was decrement in the accuracy values. For the *nationality* output, the minimum and maximum accuracy values found in Table 4.16 are 80.88% and 93.38%, while in Table 4.18 the values are 83.09% and 90.44% of accuracy respectively. Table 4.19 is a confusion matrix that shows the number of correctly classified as well as the incorrectly classified instances.

Table 4.19. Confusion matrix of NB algorithm based on *nationality* output

		Classified value	
		NG	**TR**
Actual value	**NG**	55	2
	TR	11	68

The confusion matrix given in Table 4.19 above indicates that the sensitivity, specificity, and accuracy values are also incredible. As a result, sensitivity is 96.49%, specificity is 86.08% and the accuracy value is 90.44% as mentioned. As shown in the results of the obtained accuracies of different algorithms, there is no specific algorithm that is set to be the best in all the outputs provided. To select a specific algorithm for data analysis, the datasets needs to be unique. The results from classification analysis in this thesis often fluctuate from bright results to less good results or the other way around. This reflects the fact that the used questionnaire consists of different questions collected from different papers. However, it can be concluded that the school administrators can be classified by their nationality due to the

highest accuracy value found as **93.38%** after the classification algorithms are being executed. Conclusively with this result, the classification analysis is ended.

4.5. Clustering Findings

The clustering analysis used in this thesis is based on three different algorithms found in WEKA application. These three algorithms include: Simple K-Means, Expectation Maximization (EM), and Make Density Based Clusterer (MDBC) algorithms. As any other classification or association rule algorithms in the WEKA application, clustering algorithms have parameters too. However, this thesis will consider the default entries of these parameters. The EM algorithm is a statistical model that finds a maximum likelihood probabilities of parameters iteratively. In clustering analysis MDBC algorithm select an alternative algorithm in order to identify dense clusters of points, and allowing it to learn clusters of arbitrary shape and identify outliers in the data.

For this thesis, the clustering results of `NG-TR2.arff` data file is obtained by using the *nationality* as output after selecting the cluster mode to be "classes to clusters evaluation". Table 4.20 below presents a comparison between the three algorithms showing cluster outputs based on the *nationality* output.

Table 4.20. Performance comparison table of the three algorithms

	Simple K-Means	**EM**	**MDBC**
Nationality (%)	81.62	88.97	88.97

As seen in the comparison table above, the highest percentage value was found by MDBC algorithm. The following list contains the details of the above mentioned three algorithms:

1. Simple K-Means algorithm clustered the instances into two, namely; cluster 0 as NG and cluster 1 as TR. Both clusters contain 68 instances and 50% of the instances each. Table 4.21 provides the confusion matrix of the result.

Table 4.21. Confusion matrix of Simple K-Means algorithm

		Classified value	
		NG	**TR**
Actual value	**NG**	50	7
	TR	18	61

Based on the confusion matrix result in Table 4.21 above, the performance value of the cluster is found as 81.62%.

2. EM algorithm clustered the instances into two, namely cluster 0 as TR and cluster 1 as NG. These clusters contain 66 and 70 instances, and their percentages are 49% and 51% respectively. Table 4.22 shows the confusion matrix of the result.

Table 4.22. Confusion matrix of EM algorithm

		Classified value	
		TR	**NG**
Actual value	**TR**	65	14
	NG	1	56

Based on the confusion matrix result in Table 4.22 above, the performance value of the cluster is found as 88.97%.

3. MDBC algorithm clustered the instances into two, namely cluster 0 as NG and cluster 1 as TR. These clusters contain 70 and 66 instances, and their percentages are 51% and 49% respectively. Table 4.23 represents the confusion matrix of the result.

Table 4.23. Confusion matrix of Make Density Based Clusterer algorithm

		Classified value	
		NG	**TR**
Actual value	**NG**	56	1
	TR	14	65

Based on the confusion matrix result in Table 4.20 above the performance value of the cluster is found as 88.97%.

As seen in the previous tables, the highest performance result is found by EM and MDBC algorithm as 88.97%. In brief, the data file is clustered in the WEKA application with a high percentage of probability. With this value, the clustering method in this thesis is said to have achieved a promising result.

4.6. Association Rule Findings

Since the aim of this thesis is to evaluate and find the relations between one or more statements. Relations in a particular part of the questionnaire or between some of the inter-part statements of the questionnaire. Apriori algorithm is used, to predict the next step or action happening in the datasets. There are a lot of parameters for all the algorithm such as Apriori or Tertius algorithm under Associate in the WEKA application. Any WEKA user may be able to navigate and change some parameters to obtain better rules or results. Mostly the parameters are left standard but can be changed.

4.6.1. Rules Obtained from Nigerian Questionnaire

The following parameters are used for Apriori algorithm to generate one hundred thousand rules: car = false, class-index = -1, delta = 0.05, lower-bound-min-support = 0.1, metric-type = lift, min-metric = 0.9, number-rules = 100000, out-put-item-set = false, remove-all-missing-cols = false, significant-level = -1.0, upper-bound-min-support = 1.0 and verbose = false. Based on the above parameter entries the following results are obtained:

Rule 1: Confidence is 100% and lift value is 2.19

Severance = 0-10 and *If someone opposes me, I can find the ways and means to get what I want* = 3 and *I can solve most problems if I invest the necessary effort* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 ==> *I can always manage to solve difficult problems if I try hard enough* = 3 and *I am confident that I could deal efficiently with unexpected events* = 3 and *Thanks to my resourcefulness. I can handle unforeseen situations* = 3

Rule 2: Confidence is 100% and lift value is 2.19

Severance = 0-10 and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I am confident that I could deal efficiently with unexpected events* =3 and *Whenever I face a difficult decision I feel pessimistic about finding a good solution* = 3 ==> *If someone opposes me, I can find the ways and means to get what I want* = 3 and *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3

Rule 3: Confidence is 100% and lift value is 2.11

Severance = 0-10 and *If someone opposes me, I can find the ways and means to get what I want* = 3 and *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 ==> *I can always manage to solve difficult problems if I try hard enough* = 3 and *I am confident that I could deal efficiently with unexpected events* = 3

Rule 4: Confidence is 100% and lift value is 1.73

Age = 26-40 and *I forget or overlook important information about choice alternatives* = 3 and *The possibility that some small thing might go wrong causes me to swing abruptly in my preferences* = 3 ==> *If a decision can be made by me or another person I let the other person make it* = 3

Rule 5: Confidence is 100% and lift value is 1.63

Age = 26-40 and *I forget or overlook important information about choice alternatives* = 3 and *If a decision can be made by me or another person I let the other person make it* = 3 ==> *The possibility that some small thing might go wrong causes me to swing abruptly in my preferences* = 3

Rule 6: Confidence is 100% and lift value is 1.63

Age = 26-40 and *I can handle whatever comes my way* = 3 and *Whenever I face a difficult decision I feel pessimistic about finding a good solution* = 3 ==> *If I am in trouble, I can think of a good solution* = 3

Rule 7: Confidence is 100% and lift value is 1.58

Gender = Male and *When I am confronted with a problem. I can find several solutions* = 3 ==> *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3

Rule 8: Confidence is 100% and lift value is 1.46

Gender = Male and *I feel as if I am under tremendous time pressure when making decisions* = 3 and *I feel better about choosing if I can convince myself that the decision is not all that important* = 3 ==> *I like to consider all of the alternatives* = 3

Rule 9: Confidence is 100% and lift value is 1.43

Age = 26-40 and *I like to consider all of the alternatives* = 3 and *The possibility that some small thing might go wrong causes me to swing abruptly in my preferences* = 3 ==> *I feel better about choosing if I can convince myself that the decision is not all that important* = 3

Rule 10: Confidence is 100% and lift value is 1.43

Edu. Level = Bachelor and *I like to consider all of the alternatives* = 3 and *After making a decision I am inclined to undervalue the worth of the alternatives I did not choose* = 3 ==> *I feel better about choosing if I can convince myself that the decision is not all that important* = 3

Rule 11: Confidence is 100% and lift value is 1.43

Edu. Level = Bachelor and *If someone opposes me, I can find the ways and means to get what I want* = 3 and *I like to consider all of the alternatives* = 3 ==> *I feel better about choosing if I can convince myself that the decision is not all that important* = 3

Rule 12: Confidence is 100% and lift value is 1.39

Gender = Male and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I am certain that I can accomplish my goals* = 3 ==> *If someone opposes me, I can find the ways and means to get what I want* = 3

Rule 13: Confidence is 100% and lift value is 1.39

Gender = Male and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I can solve most problems if I invest the necessary effort* = 3 ==> *If someone opposes me, I can find the ways and means to get what I want* = 3

Rule 14: Confidence is 100% and lift value is 1.39

Gender = Male and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I feel better about choosing if I can convince myself that the decision is not all that important* = 3 ==> *If someone opposes me, I can find the ways and means to get what I want* = 3

Rule 15: Confidence is 100% and lift value is 1.39

Rank = School-principal and *I can always manage to solve difficult problems if I try hard enough* = 3 ==> *If someone opposes me, I can find the ways and means to get what I want* = 3

Rule 16: Confidence is 100% and lift value is 1.39

Edu. Level = Bachelor and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I am certain that I can accomplish my goals* = 3 ==> *If someone opposes me, I can find the ways and means to get what I want* = 3

This thesis aims to analyze and find out the relationships between different questions from different or within parts of the questionnaire. After the analyzation, the results and rules are found based on Apriori algorithm. The results show that most of the participants are males, aged between 26 and 40, have a certificate of a bachelor degree, and happened to be school principals. The results of rule 4 and rule 5 show that if a person within the age of 26 and 40 agreed to overlook or forget an important information, and also afraid that a small thing might go wrong, therefore, preferably that person will agree on allowing someone to make the decision instead. Considering the results from rule 10, it is noticed that if a person with gender male agreed that he is under pressure while making a decision, then he definitely considers all the alternatives or facts before making decisions. After evaluating the results from rule 9, 10 and 11 it is noticed that if a person's age is between 26 and 40, or education level is bachelor degree and agreed to consider all the alternatives before making decision, then that person certainly agrees to choose if convinced that the decision is not that important. Surprisingly, the results from rule 12, 13, 14, 15 and 16 show some interesting outcomes that if either the person's gender is male, rank is school principal or education level is bachelor degree and happens to always manage in solving difficult problems when tried hard enough, then the person can also find a way and means to get what the person

wants when being opposed by someone. As noted, all the rules have 100% confidence and with lift value of more than 1.30. Most of the results (i.e., rule 1, 3, 4, 5, 7, 8, 9, 10, 12, 13, 15, and 16), show that the relationships are part-based in the questionnaire while only four of the results (i.e. rule 2, 6, 11, and 14) defined the relationships as parts-mixed-based.

4.6.2. Rules Obtained from Turkish Questionnaire

The following parameters are used for Apriori algorithm to generate one hundred thousand rules; car = false, class-index = -1, delta = 0.05, lower-bound-min-support = 0.1, metric-type = lift, min-metric = 0.9, number-rules = 100000, out-put-item-set = false, remove-all-missing-cols = false, significant-level = -1.0, upper-bound-min-support = 1.0 and verbose = false. Based on the above parameter entries, the following results are obtained:

Rule 1: Confidence is 100% and lift value is 3.76
Rank = School-Principal and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 5 and *When making decisions I like to collect lots of information* = 5 ==> *I like to consider all of the alternatives* = 5 and *I consider how best to carry out the decision* = 5 and *I take a lot of care before choosing* = 5

Rule 2: Confidence is 100% and lift value is 3.59
Rank = School-Principal and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 5 and *When making decisions I like to collect lots of information* = 5 ==> *I like to consider all of the alternatives* = 5 and *I consider how best to carry out the decision* = 5

Rule 3: Confidence is 100% and lift value is 3.04
Gender = Male and *I waste a lot of time on trivial matters before getting to the final decision* = 1 ==> *Rank = School-Principal* and *Have too many responsibilities* = 5 and *I do not like to take responsibility for making decisions* = 1

Rule 4: Confidence is 100% and lift value is 2.32

Edu. Level = Degree and *Thanks to my resourcefulness. I can handle unforeseen situations* = 5 and *I like to consider all of the alternatives* = 5 ==> *Rank = School-Principal* and *I am certain that I can accomplish my goals* = 5

Rule 5: Confidence is 100% and lift value is 2.03

Age = >40 and *If someone opposes me, I can find the ways and means to get what I want* = 4 and *Thanks to my resourcefulness. I can handle unforeseen situations* = 4 and *I consider how best to carry out the decision* = 4 ==> *Gender = Male* and *When I am confronted with a problem. I can find several solutions* = 4

Rule 6: Confidence is 100% and lift value is 2.03

Age = >40 and *I am certain that I can accomplish my goals* = 4 and *I am confident that I could deal efficiently with unexpected events* = 4 and *When I am confronted with a problem. I can find several solutions* = 4 ==> *Rank = School-Principal* and *Thanks to my resourcefulness. I can handle unforeseen situations* = 4

Rule 7: Confidence is 100% and lift value is 1.98

Severance = >20 and *I like to consider all of the alternatives* = 5 ==> *Gender = Male* and *Age = >40*

Rule 8: Confidence is 100% and lift value is 1.8

Severance = >20 and *I try to find out the disadvantages of all alternatives* = 4 ==> *Age = >40* and *Rank = School-Principal*

Rule 9: Confidence is 100% and lift value is 1.76

Severance = >20 and *The possibility that some small thing might go wrong causes me to swing abruptly in my preferences* = 2 ==> *Age = >40*

Rule 10: Confidence is 100% and lift value is 1.76

Severance = >20 and *I try to find out the disadvantages of all alternative* = 4 ==> *Age = >40*

After the analyzation, the results and rules are found based on Apriori algorithm. The results show that most of the participants are males, aged above 40, have a certificate of a bachelor degree, and happened to be school principals. The first and second rules show that

if the individual is a male and a school principal, then being calm is not an issue to that person while facing difficulties in making a decision. Also, individuals consider all the possible alternatives and collect a lot of information before making a decision. This is due to the fact mentioned in rule 3 that the participants see the principalship as a job with many responsibilities, and require a lot of time before reaching a conclusion of a decision. Adding to that, the rule says that the participants do not want to be responsible for decision making. The results of the 4th rule tells that if an individual has a bachelor degree and can solve any problem because of his resources by considering all alternatives, then the individual can achieve own goals. Surprisingly, the rest of the results (i.e., from rule 5th to 10th) show some interesting outcomes that all the rules are connected to an individual with age higher than 40. As noted, all the rules have 100% confidence and with lift value higher than 1.76. All the rules showed that the relationships are part-mixed-based in the questionnaire, except for the 6th rule defined the relationships as parts-based.

4.6.3. Rules Obtained from Merged Questionnaires

The following parameters are used for Apriori algorithm to generate one hundred thousand rules: car = false, class-index = -1, delta = 0.05, lower-bound-min-support = 0.1, metric-type = lift, min-metric = 0.9, number-rules = 100000, out-put-item-set = false, remove-all-missing-cols = false, significant-level = -1.0, upper-bound-min-support = 1.0 and verbose = false. Based on the above parameter entries, the following results are obtained:

Rule 1: Confidence is 100% and lift value is 1.42

Rank = School-Principal and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I am certain that I can accomplish my goals* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *If I am in trouble. I can think of a good solution* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *When I am confronted with a problem. I can find several solutions* = 3

Rule 2: Confidence is 100% and lift value is 1.42

Rank = School-Principal and *I am certain that I can accomplish my goals* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *If I am in trouble. I can think of a good*

solution = 3 and *I take a lot of care before choosing* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *When I am confronted with a problem. I can find several solutions* = 3

Rule 3: Confidence is 100% and lift value is 1.42
Rank = School-Principal and *I am certain that I can accomplish my goals* = 3 and *I am confident that I could deal efficiently with unexpected events* = 3 and *If I am in trouble. I can think of a good solution* = 3 and *I take a lot of care before choosing* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *When I am confronted with a problem. I can find several solutions* = 3

Rule 4: Confidence is 100% and lift value is 1.42
Rank = School-Principal and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I am certain that I can accomplish my goals* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *If I am in trouble. I can think of a good solution* = 3 and *I take a lot of care before choosing* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *When I am confronted with a problem. I can find several solutions* = 3

Rule 5: Confidence is 100% and lift value is 1.42
Rank = School-Principal and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I am certain that I can accomplish my goals* = 3 and *I am confident that I could deal efficiently with unexpected events* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *If I am in trouble. I can think of a good solution* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *When I am confronted with a problem. I can find several solutions* = 3

Rule 6: Confidence is 100% and lift value is 1.42
Rank = School-Principal and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I am certain that I can accomplish my goals* = 3 and *I am confident that I could deal efficiently with unexpected events* = 3 and *If I am in trouble. I can think of a good solution* = 3 and *I take a lot of care before choosing* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *When I am confronted with a problem. I can find several solutions* = 3

Rule 7: Confidence is 100% and lift value is 1.42
Rank = School-Principal and *I am certain that I can accomplish my goals* = 3 and *I am confident that I could deal efficiently with unexpected events* = 3 and *I can remain calm when facing difficulties because I can rely*

on my coping abilities = 3 and *If I am in trouble. I can think of a good solution* = 3 and *I take a lot of care before choosing* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *When I am confronted with a problem. I can find several solutions* = 3

Rule 8: Confidence is 100% and lift value is 1.4
Rank = School-Principal and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *I try to be clear about my objectives before choosing* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *I take a lot of care before choosing* = 3

Rule 9: Confidence is 100% and lift value is 1.4
Rank = School-Principal and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I am confident that I could deal efficiently with unexpected events* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *I like to consider all of the alternatives* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *I take a lot of care before choosing* = 3

Rule 10: Confidence is 100% and lift value is 1.4
Rank = School-Principal and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I am confident that I could deal efficiently with unexpected events* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *I try to be clear about my objectives before choosing* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *I take a lot of care before choosing* = 3

Rule 11: Confidence is 100% and lift value is 1.4
Rank = School-Principal and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *I like to consider all of the alternatives* = 3 *And I consider how best to carry out the decision* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *I take a lot of care before choosing* = 3

Rule 12: Confidence is 100% and lift value is 1.4
Rank = School-Principal and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *When I am confronted with a problem. I can find several solutions* = 3 *I try to be clear about my objectives before choosing*

= 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *I take a lot of care before choosing* = 3

Rule 13: Confidence is 100% and lift value is 1.4
Rank = School-Principal and *I am certain that I can accomplish my goals* = 3 and *I am confident that I could deal efficiently with unexpected events* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *I like to consider all of the alternatives* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *I take a lot of care before choosing* = 3

Rule 14: Confidence is 100% and lift value is 1.4
Rank = School-Principal and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I am confident that I could deal efficiently with unexpected events* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *When I am confronted with a problem. I can find several solutions* = 3 *I like to consider all of the alternatives* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *I take a lot of care before choosing* = 3

Rule 15: Confidence is 100% and lift value is 1.4
Rank = School-Principal and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *When making decisions I like to collect lots of information* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *I take a lot of care before choosing* = 3

Rule 16: Confidence is 100% and lift value is 1.4
Rank = School-Principal and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *I consider how best to carry out the decision = 3 and I try to be clear about my objectives before choosing* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *I take a lot of care before choosing* = 3

Rule 17: Confidence is 100% and lift value is 1.4
Rank = School-Principal and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I am certain that I can accomplish my goals* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *I try to be clear about my objectives before choosing* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *I take a lot of care before choosing* = 3

Rule 18: Confidence is 100% and lift value is 1.4

Rank = School-Principal and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I can solve most problems if I invest the necessary effort = 3 and I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *I like to consider all of the alternatives* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *I take a lot of care before choosing* = 3

Rule 19: Confidence is 100% and lift value is 1.4

Rank = School-Principal and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I can solve most problems if I invest the necessary effort = 3 and I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *I try to be clear about my objectives before choosing* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *I take a lot of care before choosing* = 3

Rule 20: Confidence is 100% and lift value is 1.4

Rank = School-Principal and *I can always manage to solve difficult problems if I try hard enough* = 3 and *I am certain that I can accomplish my goals* = 3 and *I am confident that I could deal efficiently with unexpected events* = 3 and *I can remain calm when facing difficulties because I can rely on my coping abilities* = 3 and *I like to consider all of the alternatives* = 3 ==> *Thanks to my resourcefulness. I can handle unforeseen situations* = 3 and *I take a lot of care before choosing* = 3

After the analyzation, the results and rules are found based on Apriori algorithm. The results are found only based on Rank output if outputs are considered. Out of the most repeating rules, 20 are extracted as listed above. The results show that only rule 1 and 5 are part-based while the rest of the rules are parts-mixed-based relationship. The parts in which the relationship took place are only the 3rd and 4th part of the questionnaire. Results from rule 1 to 7 (i.e., rule 2, 3, 4, 6, and 7) showed that every rule contained only 1 statement from the 4th part and the other statements (i.e., from 8 to 20) are from the 3rd part. That only one statement contained is: "I take a lot of care before choosing". Because of the participants are all school principals, the results show that the individuals are responsible persons and they know what they are doing. Another finding in all the rules is the usage of "Thanks to my resourcefulness, I can handle unforeseen situations," and "I can remain calm when facing difficulties because I can rely on my coping abilities" statements. Apparently, the individuals

can use the available resources on hand to solve the upcoming situations. Except for the 3rd and 6th rules which do not contain the second statement of calmness. Owing to this, it is safe to say the result is promising to have 90% of the results needed from the outcomes. In conclusion to the commenting on the obtained rules, if a rule is said to be parts-mixed-based rule and contained "Thanks to my resourcefulness, I can handle unforeseen situations" statement from 3rd part then, it must also contain "I take a lot of care before choosing" statement from 4th part of the questionnaire.

5. CONCLUSION

This thesis explained the introduction to data mining where data mining is being defined as the extraction of valuable or meaningful information to organization or institution which will lead to the revelation of truth depending on the information contained in the database. Later on, it discussed most common used data mining algorithms in classification, clustering, and association rule methods. These algorithms are applied in many different fields such as store sales, schools, organizations, companies, hospitals and so on. In classification part algorithms under "classification with a decision tree, memory-based classification and statistical classification" such as ID3, k-nearest neighbor and Bayes theorem respectively are revised and discussed. In clustering part, algorithms like farthest neighboring, k-mean are studied. Last but not the least in association rule, algorithms such as support and confidence measure, and Apriori are also investigated. Most of the workplaces or business organizations use specific kinds of algorithms that help in handling information drawn from the database fast enough.

In the context of this thesis, a questionnaire is applied and the various attitudes of school administrators was analyzed using data mining techniques and remarkable results were obtained. With the provided explanations and good examples given in this thesis, many of these algorithms are applied to numerous data for analyses. After studying this thesis, some points are noted like the way data has been dealt with in terms of extracting a large amount of information about things or persons from the databases. Also, the unwanted data are cleaned by selecting the desired sets needed to help get close to specific about what is being checked over and over again.

However, there is a large amount of data stored on the Internet from different organizations and institutions. Therefore, neither the selected algorithms nor other algorithms perform best specifically in classification, clustering or association rule method. Rather, this depends on what kind of datasets are to be analyzed. This is due to the fact that most of the algorithms are for specific problem applications. One will have to develop own algorithm or use a specific algorithm based on the kind of information which is available in the database.

In the beginning, a questionnaire is applied to school administrators in Nigeria and Turkey. Later, to evaluate the school administrators we've used the data mining techniques and obtained results. Firstly, we applied the classification method to the data files formed from the applied questionnaire by considering seven different algorithms. The evaluation results of the questionnaire for both countries are obtained separately and a maximum of **93.38%** accuracy is obtained based on nationality output while the questionnaire data are integrated. Secondly, the clustering method is applied to the integrated file from the two countries and results are obtained with high value as **88.97%** based on the nationality output. This shows that the two country school administrators can be predicted using the clustering method. Lastly, association rule method is also applied in order to find some interesting rules, the results of the association rule analysis showed that most of the rules from Nigerian analysis are **part-based**. But for the Turkish files analysis most are **part-mixed-based**. For the integrated file, if a rule is said to be parts-mixed-based rule and contained *Thanks to my resourcefulness, I can handle unforeseen situations* statement from 3rd part then, it must also contain *I take a lot of care before choosing* statement from 4th part of the questionnaire.

REFERENCES

[1] **O. R. Zaiane,** "Principles of Knowledge Discovery in Database," 1999. [Online]. Available: https://webdocs.cs.ualberta.ca/~zaiane/courses/cmput690/slides/ch0s.pdf. [Accessed 23 October 2015].

[2] **E. Tekkanat and M. Topaloglu,** "The Assessment of High Schoolers' Internet Addiction," *Procedia - Social and Behavioral Sciences,* vol. 205, pp. 664-670, 2015.

[3] **R. J. Robbins,** "Databases Fundamentals," 1994. [Online]. Available: http://129.237.201.53/db-fund.pdf. [Accessed 23 October 2015].

[4] **O. Yalin,** "Veri Madenciligi," in *Papatya Yayincilik Egitim,* Istanbul, 2013.

[5] **A. Chan,** "Introduction to Computers for the Arts and Social Sciences," 17 May 2004. [Online]. Available: http://people.scs.carleton.ca/~achan/. [Accessed 5 November 2015].

[6] **W. Teams,** "Wikipedia," Wikipedia, 08 April 2016. [Online]. Available: https://en.wikipedia.org/wiki/School. [Accessed 11 April 2016].

[7] **C. press,** "Cambridge," Cambridge University Press, 2016. [Online]. Available: http://dictionary.cambridge.org/dictionary/english/administration. [Accessed 12 April 2016].

[8] **M. Alshurman,** "Democratic Education and Administration," *Procedia - Social and Behavioral Sciences,* vol. 176, pp. 861-869, 2015.

[9] **M. J. Zaki and W. Meira Jr.,** Data mining and analysis: fundamental concepts and algorithms, New York: Cambridge University Press, 2014.

[10] **J. Leskovec, A. Rajaraman and J. D. Ullman,** Mining of Massive Datasets, United Kingdom: Cambridge University Press, 2014.

[11] **U. Fayyad, G. Piatetsky-Shapiro and P. Smyth,** "From Data Mining to Knowledge Discovery in Databases," *Al Magazine,* vol. 17, no. 3, p. 37, 1996.

[12] **W. W. Cohen and J. Richman,** "Learning to Match and Cluster Large High-Dimensional Data Sets For Data Integration," in *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining,* Edmonton, Alberta, Canada, 2002.

[13] **Q. Hu, D. Yu and Z. Xie,** "Information-preserving hybrid data reduction based on fuzzy-rough techniques," *Pattern recognition letters,* vol. 27, no. 5, pp. 414-423, 2006.

[14] **R. Agrawal, T. Imielinski and A. Swami,** "Mining Association Rules between Sets of Items in Large Databases," *ACM SIGMOD Record,* vol. 22, no. 2, pp. 207-216, 1993.

[15] **M. Xu, P. Watanachaturaporn, P. K. Varshney and M. K. Arora,** "Decision tree regression for soft classification of remote sensing data," *Remote Sensing of Environment,* vol. 97, no. 3, pp. 322-336, 2005.

[16] **M. N. Anyanwu and S. G. Shiva,** "Comparative analysis of serial decision tree classification algorithms," *International Journal of Computer Science and Security,* vol. 3, no. 3, pp. 230-240, 2009.

[17] **J. Han, M. Kamber and J. Pei,** Data mining: concepts and techniques, Elsevier, 2011.

[18] **B. Kumar and S. P. Baradwaj,** "Mining Educational Data to Analyze Students Performance," *IJACSA International Journal of Advanced Computer Science and Applications,* vol. 2, no. 6, 2011.

[19] **M. Sheraz, S. Dedu and V. Preda,** "Entropy Measures for Assessing Volatile Markets," *Procedia Economics and Finance,* vol. 22, pp. 655-662, 2015.

[20] **K. Adhatrao, A. Gaykar, A. Dhawan, R. Jha and V. Hanrao,** "Predicting Students' Performance using ID3 and C4. 5 Classification Algorithms," *International Journal of Data Mining & Knowledge Management Process,* vol. 3, no. 5, pp. 39-52, 2013.

[21] **W. Duch, R. Adamczak and N. Jankowski,** "IMPROVED MEMORY-BASED CLASSIFICATION," in *Proc. EANN'96.*

[22] **B. Masand, G. Linoff and D. Waltz,** "Classifying news stories using memory based reasoning," in *Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval. ACM,* Massachusetts, 1992.

[23] **S. Diplaris, G. Tsoumakas, P. A. Mitkas and I. Vlahavas,** "Protein Classification with Multiple Algorithms," in *Advances in Informatics,* Springer Berlin Heidelberg, 2005.

[24] **E. Hoffmann and M. Chamie,** "Standard statistical classifications: basic principles," *Statistical Journal of the United Nations Economic Commission for Europe,* vol. 19, no. 4, pp. 223-241, 2002.

[25] **E. Alpaydin,** "Introduction to machine learning," *sl*, pp. 249-256, 2010.

[26] **M. F. Triola,** "Bayes' Theorem," in *Elementary Statistics*, 2010.

[27] **M. Karabatak,** "A new classifier for breast cancer detection based on Naïve Bayesian," *Measurement,* vol. 72, pp. 32-36, 2015.

[28] **S. R. D. M. Oliveira and O. R. Zaiane,** "Privacy preserving clustering by data transformation," *Journal of Information and Data Management,* vol. 1, no. 1, p. 37, 2010.

[29] **L. T. Wille,** New Directions in Statistical Physics: Econophysics, Bioinformatics and Pattern Recognition, Florida: Springer, 2004.

[30] **V. Kumar,** "An introduction to cluster analysis for data mining," Computer Science Department, University of Minnesota, USA, 2000.

[31] **P.-N. Tan, M. Steinbach and V. Kumar,** "Data mining cluster analysis: Basic concepts and algorithms," in *Introduction to Data Mining*, 2013, pp. 487-568.

[32] **N. Jiang and L. Gruenwald,** "Research issues in data stream association rule mining," *ACM Sigmod Record,* vol. 35, no. 1, pp. 14-19, 2006.

[33] **F. Tao, F. Murtagh and M. Farid,** "Weighted Association Rule Mining Using Weighted Support and Significance Framework," in *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM*, 2003.

[34] **C. C. Aggarwal, C. Procopiuc and P. S. Yu,** "Finding Localized Associations in Market Basket Data," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 14, no. 1, pp. 51-62, 2002.

[35] **C. C. Aggarwal and P. S. Yu,** "A new framework for itemset generation," in *Proceedings of the seventeenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems*, 1998.

[36] **Q. Zhao and S. S. Bhowmick,** "Association Rule Mining: A Survey," Singapore, 2003.

[37] **C. Ordonez,** "Association rule discovery with the train and test approach for heart disease prediction," *Information Technology in Biomedicine, IEEE Transactions on,* vol. 10, no. 2, pp. 334-343, 2006.

[38] **R. J. Bayardo Jr. and R. Agrawal,** "Mining the most interesting rules," in *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, 1999.

[39] **R. Agrawal and R. Srikant,** "Fast Algorithms for Mining Association Rules," in *Proc. 20th int. conf. very large data bases, VLDB*, 1994.

[40] **O. Akinbote,** "Problems of Teacher Education for Primary Schools in Nigeria: Beyond Curriculum Design and Implementation," *International Journal of African & African-American Studies*, vol. 6, no. 2, pp. 64-71, 2009.

[41] **A. V. Beale and K. R. Hall,** "Cyberbullying: What school administrators (and parents) can do.," *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, vol. 81, no. 1, pp. 8-12, 2007.

[42] **W. Ariratana, S. Sirisookslip and T. K. Ngang,** "Development of Leadership Soft Skills Among Educational Administrators," *Procedia-Social and Behavioral Sciences*, vol. 186, pp. 331-336, 2015.

[43] **V. E. Lee, A. S. Bryk and J. B. Smith,** "The Organization of Effective Secondary Schools," *Review of research in education*, vol. 19, pp. 171-267, 1993.

[44] **H. F. Ladd,** "Teachers' Perceptions of Their Working Conditions: How Predictive of Policy-Relevant Outcomes? Working Paper 33.," *National Center for Analysis of Longitudinal Data in Education Research*, 2009.

[45] **J. P. Spillane,** "Distributed leadership," *The educational forum*, vol. 69, no. 2, pp. 143-150, 2005.

[46] **H. T. Ekundayo,** "Administering secondary schools in Nigeria for quality output in the 21st century; The principals' challenge," *European Journal of Educational Studies*, vol. 2, no. 3, 2010.

[47] **M. Telem and T. Buvitski,** "The potential impact of information technology on the high school principal: a preliminary exploration," *Journal of Research on Computing in Education*, vol. 27, no. 3, pp. 281-296, 1995.

[48] **P. C. Godfrey and J. T. Mahoney,** "The Functions of the Executive at 75 An Invitation to Reconsider a Timeless Classic," *Journal of Management Inquiry*, vol. 23, no. 4, pp. 360-372, 2014.

[49] **C. Barnard,** "The Functions of the Executive," *Cambridge/Mass*, 1938.

[50] **D. N. Harris, S. A. Rutledge, W. K. Ingle and C. C. Thompson,** "Mix and match: What principals really look for when hiring teachers," *Education,* vol. 5, no. 2, pp. 228-246, 2010.

[51] **D. Boyd, P. Grossman, M. Ing, H. Lankford and J. Wyckoff,** "The influence of school administrators on teacher retention decisions," *American Educational Research Journal,* vol. 48, no. 2, pp. 303-333, 2011.

[52] **L. Darling-Hammond, M. Lapointe, D. Meyerson and M. T. Orr,** "Preparing school leaders for a changing world: Executive summary," 2007.

[53] **A. Bandura,** "Self-efficacy: toward a unifying theory of behavioral change," *Psychological review,* vol. 84, no. 2, pp. 191-215, 1977.

[54] **M. Jerusalem and R. Schwarzer,** "Self-efficacy as a resource factor in stress appraisal processes," in *Self-efficacy: Thought control of action*, New York, Taylor & Francis Group, 1992, pp. 195-213.

[55] **A. Aypay,** "The Adaptation Study of General Self-Efficacy (GSE) Scale to Turkish," *INONU UNIVERSITY JOURNAL OF THE FACULTY OF EDUCATION,* vol. 11, no. 2, pp. 113-131, 2010.

[56] **L. Mann, P. Burnett, M. Radford and S. Ford,** "The Melbourne Decision Making Questionnaire: An Instrument for Measuring Patterns for Coping with Decisional Conflict," *Journal of Behavioral Decision Making,* vol. 10, pp. 1-19, 1997.

[57] **E. M. Deniz,** "Investigation of the relation between decision making Self-esteem, decision making style and problem solving Skills of university students," *Eurasian Journal of Educational Research,* vol. 4, no. 15, pp. 23-35, 2004.

[58] **R. N. Kanungo and S. T. Menon,** "Managerial resourcefulness: The construct and its measurement," *Journal of Entrepreneurship,* vol. 13, no. 2, pp. 129-152, 2004.

[59] **A. Ersozlu,** "The effect of managerial resourcefulness of school administrators on organizational commitment, organizational citizenship behaviors and job satisfaction levels of teachers," *Unpublished Ph.D. Thesis, Elazig: Firat University Institute of Educational Sciences,* 2012.

[60] **W. Zhu, N. Zeng and N. Wang,** "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations," in *NESUG proceedings: health care and life sciences*, Baltimore, Maryland, 2010.

[61] **G. Demiröz and H. A. Güvenir,** "Classification by voting feature intervals," in *European Conference on Machine Learning*, Berlin Heidelberg, 1997.

[62] **R. J. Quinlan,** "Induction of decision trees," *Machine learning*, vol. 1, no. 1, pp. 81-106, 1986.

APPENDIX A: SAMPLE OF THE QUESTIONNAIRE APPLIED

Firat University

Questionnaire (Survey)

The questionnaire comprises of 4 parts. The first part is for personal information about the participant while the other three parts are the main questionnaire statements and each statement comprises of five (5) selective choices as follows: "*Strongly Disagree, Somewhat Disagree, Neutral, Somewhat Agree, Strongly Agree*" respectively. For each statement one has to select **only one** choice or the statement will be counted as "not answered"

We assure you that any information you gave to us will not be used for anything or anywhere else but for this research purpose only.

Thank you for your heartily and honestly cooperating with us.

PART 1

1. **Gender** ( ) Female ( ) Male
2. **Age** ……..
3. **Duty / Rank** ( ) School Principal ( ) Vice principal Admin. ( ) Vice principal Acad.
4. **Edu. Level** ( ) NCE ( ) Bachelor ( ) Post graduate
5. **Severance / Seniority** ..... (Please specify. i.e. "1, 2, 3, or more")

PART 2

BELIEVES ABOUT THE PRINCIPALSHIP QUESTIONNAIRE

1-Strongly Disagree 2- Somewhat Disagree 3- Neutral 4- Somewhat Agree 5- Strongly Agree

	1	2	3	4	5
1. Make a difference in the lives of students and staff					
2. Provide opportunities for professional growth					
3. Enable me to develop relationships with others inside and outside the school					
4. Enable me to influence school change					
5. Require very long work hours					
6. Have too many responsibilities					
7. Decrease my opportunity to work directly with children					
8. Create a lot of stress					
9. The stress and disappointments involved in serving as principal of this school aren't really worth it					
10. If I could get a higher paying job, I'd leave education as soon as possible					
11. I think about transferring to another school					
12. I will continue being a principal until something better comes along					
13. I plan to remain principal of this school as long as I am able					
14. I plan to remain a principal until I retire					

PART 3

GENERALIZED PERCIEVED SELF EFFICACY QUESTIONNAIRE

	1	2	3	4	5
1. I can always manage to solve difficult problems if I try hard enough.					
2. If someone opposes me, I can find the ways and means to get what I want.					
3. I am certain that I can accomplish my goals.					
4. I am confident that I could deal efficiently with unexpected events.					
5. Thanks to my resourcefulness, I can handle unforeseen situations.					
6. I can solve most problems if I invest the necessary effort.					
7. I can remain calm when facing difficulties because I can rely on my coping abilities.					
8. When I am confronted with a problem, I can find several solutions.					
9. If I am in trouble, I can think of a good solution.					
10. I can handle whatever comes my way.					

PART 4

MELBOURNE DECISION MAKING QUESTIONNAIRE

	1	2	3	4	5
1. I feel as if I'm under tremendous time pressure when making decisions.					
2. I feel better about choosing if I can convince myself that the decision is not all that important.					
3. I like to consider all of the alternatives.					
4. When I have a decision to make I try not to think about it.					
5. I prefer to leave decisions to others.					
6. Whenever I get upset by having to make decisions I choose on the spur of the moment.					
7. I try to find out the disadvantages of all alternatives.					
8. I am inclined to blame others when decisions turn out badly.					
9. I waste a lot of time on trivial matters before getting to the final decision.					
10. I feel uncomfortable about making decisions.					
11. I consider how best to carry out the decision.					
12. Even after I have made a decision I delay acting upon it.					
13. After making a decision I am inclined to undervalue the worth of the alternatives I did not choose.					
14. When making decisions I like to collect lots of information.					
15. I avoid making decisions.					
16. I only want to hear information about my preferred alternative.					
17. When I have to make a decision I wait for a long time before starting to think about it.					
18. I don't like to take responsibility for making decisions.					
19. I try to be clear about my objectives before choosing.					
20. I forget or overlook important information about choice alternatives.					
21. The possibility that some small thing might go wrong causes me to swing abruptly in my preferences.					
22. If a decision can be made by me or another person I let the other person make it.					
23. Whenever I face a difficult decision I feel pessimistic about finding a good solution.					
24. I take a lot of care before choosing.					
25. I choose on the basis of some small thing.					
26. I don't make decisions unless I really have to.					
27. I delay making decisions until it is too late.					
28. I prefer that people who are better informed decide for me.					
29. After a decision is made I spend a lot of time convincing myself it was correct.					
30. I put on making decisions.					
31. I can't think straight if I have to make decisions in a hurry.					

CURRICULUM VITAE (CV)

Personal Information

Name	Usman Umar
Date of Birth	25.11.1989
Place of Birth	Gashu'a, Yobe state, Nigeria
Nationality	Nigeria
Marital Status	Single
Address-(TR)	Çaydaçıra Mah. Anguzubaba Sok. DG8/A. No: 17 D: 11, Elazığ, Turkey
Address-(NG)	SSQ: 01 Uscoe gashu'a, Yobe state, Nigeria
Mobile No:	+90 553 433 0665 (Turkey), +234 803 887 0508 (Nigeria)
E-mail	usman.g.umar@gmail.com; usman_g_umar@hotmail.com

Education

1995-2001 (*6 years*) attended primary school in College of education staff primary school, Gashu'a, Yobe State, Nigeria.

2001-2004 (*3 years*) attended junior secondary school in Government Senior Science Secondary school Gashu'a, Yobe state, Nigeria.

2004-2007 (*3 years*) attended senior secondary school in Government Senior Science Secondary school Gashu'a, Yobe state, Nigeria.

2009-2010 Attended Turkish language courses for one year in Istanbul University, Istanbul, Turkey.

2010-2014 (*4 years*) Undergraduate degree (B. Art Computer Education and Instructional Technology) in Abant Izzet Baysal University, Faculty of Education, Department of Computer Education, Bolu, Turkey.

2014-2016 (*2 years*) Master degree (M. Sc. Software Engineering) Firat University, Faculty of Technology, Department of Software Engineering, Elazig, Turkey.

Working Experience

One year teaching practice as a teacher of *programming languages* at "Bolu Ticaret Meslek Lisesi"

Certificates of Attended Courses

- Web Design
- Graphics and Animations Design
- SolidWorks
- AutoCAD
- Turkish Language