Volume 3, Issue 2, February 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Data Mining: A Comparative Study on Various Techniques and Methods

Madhuri V. Joseph
Computing & Muscat College
Oman

Lipsa Sadath
Computing & Muscat College
Oman

Vanaja Rajan
Computing & Muscat College
Oman

Abstract- Data mining may be viewed as the extraction of patterns and models from observed data or a method used for analytical process designed to explore data. There are many different methods, which may be used to predict the appropriate class for the objects. The majority of data mining techniques can deal with different data types. The paper provides an explanation and comparative study on some of the most common data mining techniques and methods in use today in day to day life and business predictions. Though there are a number of other techniques and many variations of the methods described, one of the techniques from the mentioned group is almost always used in real world deployments of data mining systems.

Keywords: Data Mining, Statistics, Nearest Neighbour, Clustering, Neural Networks, Rule Induction.

I. INTRODUCTION

Data mining is an interdisciplinary subfield of computer science which involves computational process of large data sets' patterns discovery. The goal of this advanced analysis process is to extract information from a data set and transform it into an understandable structure for further use. The methods used are at the juncture of artificial intelligence, machine learning, statistics, database systems and business intelligence. Data Mining is about solving problems by analyzing data already present in databases [1].

Data mining is commonly misrepresented to mean any form of large-scale data or information processing like collection, extraction, warehousing and analysis but is also globalized to any kind of computer decision support system. In the 1960s, statisticians had a bad practice of analyzing data without an a-priori hypothesis termed "Data Fishing" or "Data Dredging". The term "Data Mining" appeared around 1990 in the database community which later on became more widespread in the business and press communities. Currently, Data Mining and Knowledge Discovery are used interchangeably. Data mining is often considered to be "A blend of statistics, AI and Database research.

II. DEFINITIONS

The actual data mining task is the automatic or semi-automatic analysis of large quantities of data to extract previously unknown interesting patterns such as groups of data records (cluster analysis), unusual records (anomaly detection) and dependencies (association rule mining). This usually involves using database techniques such as spatial indexes . According to David Bolton Data mining is the process of processing large volumes of data (usually stored in a database), searching for patterns and relationships within that data [2]. At the same time Gartner, the global leader in technology research and IT services define Data Mining as the process of discovering meaningful correlations, patterns and trends by sifting through large amounts of data stored in repositories [3]. Data mining employs pattern recognition technologies, as well as statistical and mathematical techniques.

III. TECHNIQUES

Data Mining is commonly defined as a stage used in Knowledge Discovery in Databases (KDD).There are numerous key data mining techniques that have been developing and used in data mining projects lately. These include statistics, association, classification, clustering, prediction, sequential patterns and decision tree. Basically Data Mining techniques have been broken up into two categories based on their evolution. Although these techniques are mentioned in the same breath, they have different analytical approaches. The final component of any data mining algorithms is data management strategy: The ways in which data are stored, indexed and accessed [4]. The focus of this paper is on a comparative study to analyze the similarities and differences between these distinct approaches namely Classical and Next Generation techniques. The basic methods covered here on Classical approaches are Statistics, Clustering and Nearest Neighborhood and that of Next Generation are Trees, Networks and Rules.

While choosing a technique, there is no specific rule that would let you select particular technique over another one. At times those choices are made reasonably arbitrarily based on the accessibility of data mining analysts who are most skilled in one technique over another. Selecting classical techniques over new generation techniques is often dependent on the obtainability of good tools and good analysts. Any realistic knowledge discovery process is not linear, but

rather iterative and interactive [5]. However the techniques chosen here for the study have been available and tried for more than two decades.

Statistics: Significance in Data mining

The two disciplines statistics and data mining have common aims and objectives. So statisticians and data miners use similar techniques to solve similar problems. Both statistics and data mining deal with the process of getting knowledge from data [6]. We use these attained knowledge for proper decision making and for further actions. The sequence can be given as follows.

DATA →INFORMATION→FACTS→KNOWLEDGE→DECISION→ACTION

Therefore, it is important to have some idea about statistical methods and techniques and how they can be applied. Statistics is the traditional field that deals with collection, analysis, interpretation and drawing conclusions from data, often collected to answer a specific question whereas data mining assumes that data have already been collected and mainly concerned with how to discover its secrets [7]. Normally these data sets are comparatively small, simple, mostly static and randomly sampled. But in the recent years, data sets have grown in size and complexity. Modern databases often contain billions or trillions of records. For a classical statistician, a database with a few thousand observations itself may be large and they work only with a sample and then makes statements about the larger data from which the sample was chosen. The advent of computers has made these works easy and the computer scientists respond with newer and more sophisticated and efficient techniques for handling and analyzing such large data sets [8]. These techniques have been become the concern of data miners. This does not mean that statistics has no role in the study of large data sets. Many of the data mining techniques were invented by statisticians and most of these techniques are integrated form of standard statistical techniques.

It may be interesting to note that statistics is a discipline which is as old as human society and it plays vital role in almost every field of human activity. Statistics is a branch of applied mathematics and is immensely useful in business, industry, medical field, weather forecast etc. There are some areas where statistical methods and techniques have much to offer to data mining. For example, statistics can be useful to answer some important questions about the data.

- How the data is distributed?
- What is the nature of the pattern?
- Which patterns are significant for a particular problem?

Visualization of the data and its structure is one of the most important themes in data mining. Visualization of quantitate data is a higher level form of analyzing the collected data. One of the most commonly used methods in the visualization of data is histogram.

1) *Histogram Example:* Histograms illustrate the distribution of data designed to show centering, dispersion and shape of the data. It can also be used to make predictions about a stable process. To understand the application of histogram, consider a simple example. Weight of 50 students in an honors class is as shown as below.

TABLE I
Database of student weight

Individual weights measured in kilogram				
60	59	57.5	67.6	54.1
59.2	60.8	56.2	72.7	50
60.2	58.4	60	42.4	74
55.3	61.7	56.5	46.1	63.4
56	61	59	57	60
57.1	62.8	45.3	74.2	55.4
58	63	55.7	68.7	65
40.8	58.3	45	66.2	62.9
61	52.7	49	64	67.5
54	51	53.7	53.8	68

Here, there are only 50 observations, but still it is very difficult to draw any specific conclusion from the tabular form of data. A histogram can provide more useful information.

This gives a quick summary of the data. From the histogram it is easy to see the average weight of a student is about 57.5. If this sample represents the whole population of students in the university, we can say that the probability of finding students whose weight is below 40 kg or above 75 kg is very low.

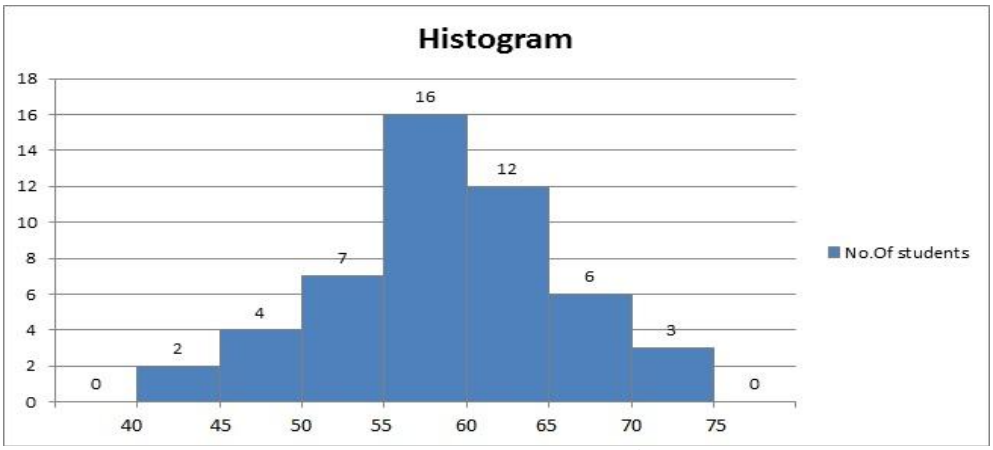

Fig. 1 Histogram that shows the weight of students

2) *Prediction using regression analysis:* In statistics, we use regression and prediction synonymously even though the terminologies are different. Basically regression deals with numerical datasets and develops a mathematical formula that fits the data. When we want to predict future behavior of a new data we plug our new data into already developed formula and we would get the prediction immediately. It is a very powerful tool in data mining with large datasets. Regression analysis is a process of determining how a variable y (prediction) is related to another variable x (predictor). The relationship can be mapped on a two-dimensional plane where predictor (independent variable x) along the X-axis and the prediction (dependent variable y) along the y-axis.

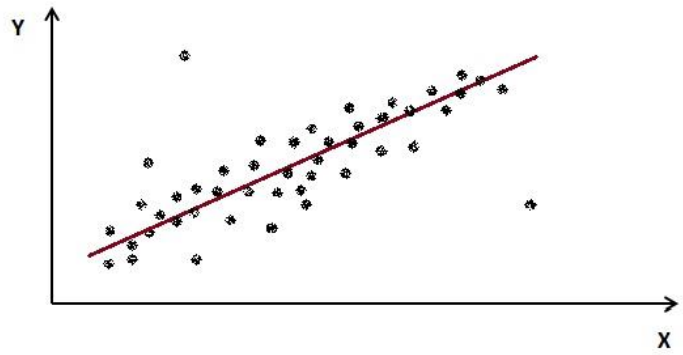

Fig. 2 Linear regression is the process of finding a straight line that fits the given sets of data.

There are different types of regression methods. The simplest is linear regression which uses the formula of a straight line, $y = mx+c$ and determines appropriate values for m and c to predict the value of y for a given value of x. For instance, a company's profit can be predicted using an equation y=\$200000 + 0.2* annual income. Adding more input values $x_1, x_2, \ldots\ldots x_k$ can produce more lines that take more values and a more generalized equation can be as follows.

$y_i = \alpha + \sum_{i=1}^{k} \beta_i \, x_i$, where α and β_i s are some constants.

A. Clustering:

A number of definitions can be given for clustering. Clustering is an automated process for grouping the related records together. Records having similar values for the attributes are grouped together [9]. The objective of clustering analysis is to find segments or clusters and to examine their attributes and values. Numerous algorithms are used for clustering. The clustering technique defines the classes and puts objects in each class accordingly. The objective of cluster analysis is to assign observations to groups (clusters) so that observations within each group are similar to one another with respect to variables or attributes of interest and the groups themselves stand apart from one another.

Clustering is the method by which like records are collected together. Usually this is done to give the end user a high level view of what is going on in the database [10]. Clustering is occasionally used to mean segmentation - which most marketing people use for coming up with a bird's eye view of the business. Once this is done the business user can get a quick high level view of what is happening within the cluster. Once the business user has worked with these codes for some time they also begin to build intuitions about how these different customer clusters will react to the marketing offers particular to their business.

1) *An instance of clustering:* A simple explanation of clustering can be made from the following scenario. In a pharmacy, there is a wide range of medicines available. The challenge is how to keep those medicines in a way that pharmacist can take from several medicines in a particular order without hassle. By using clustering technique, we

can keep medicines that have some kinds of similarities or alphabetical order in one cluster or one shelf and label it with alphabets. If pharmacists want to take medicines in that alphabet, they would only have to go to that shelf instead of looking for entire pharmacy.

2) *Types of Clustering: Hierarchical and Non-Hierarchical:* Clustering techniques are mainly of two different types- Hierarchical and Non-Hierarchical. The hierarchical method creates a hierarchy of clusters from smaller to bigger. This is usually viewed as a tree where the smallest clusters can be merged together to create the next higher level of clusters which can further form highest levels. The main advantage of this method is that it allows the end user to choose either from many clusters from many clusters or only a few.

This hierarchy is created using two algorithms namely Agglomerative and Divisive

- Agglomerative - The method starts with as many clusters as there are records where each cluster contains just one record. The clusters that are nearest each other thus merged together to form the next largest cluster. The merging thus continues until a hierarchy of clusters is constructed with just a single cluster comprising all the records at the top of the hierarchy.
- Divisive – The technique take the opposite approach from agglomerative techniques. They start with all the records in one cluster and then split that cluster into smaller pieces and then in turn to try to split those smaller pieces.

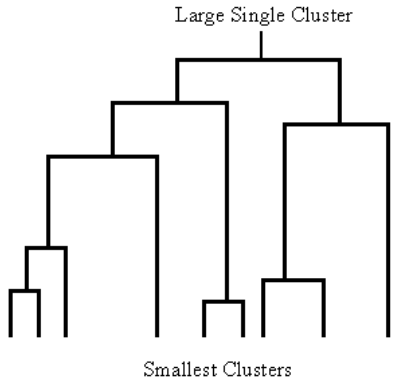

Fig 3 Hierarchy of clusters [11]

Non- hierarchical method often starts off with arbitrary or random clustering. This is done by shuffling the records to form new clusters thus improving clusters iteratively. Non-hierarchy is created using two algorithms namely single pass method and reallocation method.

- Single pass method- In this method the database must be passed through only once in order to create a cluster where each record will be visited only once.
- Reallocation – The method uses multiple passes through the data base. The method is comparatively faster than hierarchical method.

B. Nearest Neighbor

Nearest neighbor is a prediction technique that is pretty similar to clustering - its basic principle is that in order to forecast what a prediction value is in one record look for records with similar forecaster values in the historical database and use the forecast value from the record that it "nearest" to the unclassified record. Nearest Neighbor techniques are among the easiest to use and understand because they work in a way similar to the way that people think - by detecting closely matching examples. The nearest neighbor prediction algorithm works in very much the same way except that "nearness" in a database may consist of a variety of features [11].

The nearest neighbor prediction algorithm simply stated is:

Objects that are "near" to each other will have similar prediction values as well. Thus if you know the prediction value of one of the objects you can predict it for its nearest neighbors.

1) *An instance of Nearest Neighbor:* The explanation can be made clearer with the following example. If you look at the people in your locality (in this case those people that are in fact geographically near to you), you may notice that, in general, you all have somewhat similar earnings. Thus if your neighbor has earnings greater than $80,000 chances are good that you also have similar high income. Definitely the chances that you have a high income are more when all of your neighbors have earnings over $80,000 than if all of your neighbors have incomes of $10,000. Within your neighborhood there may still be a wide range of earnings likely among even your "closest" neighbors but if you had to forecast someone's income based on only knowing their neighbors, you're best chance of being right would be to predict the incomes of the neighbors who live closest to the unknown person. We may also consider certain features that are important to know, which school someone attended and what degree they attained when predicting income.

2) *Difference between the techniques: Clustering and Nearest Neighborhood:* The main difference between clustering and nearest neighbor technique is that clustering is an unsupervised learning technique whereas nearest neighbor is supervised learning technique. Unsupervised learning techniques are used for creation of models and supervised learning techniques are used to perform prediction.
   - Clustering is mainly used for consolidating data and grouping of records into like behaviors whereas nearest neighborhood is used for prediction and consolidation.
   - In case of clustering space is defined by user or by past experience whereas in nearest neighborhood space is defined by the problem.

C. Decision Trees

Decision tree is a predictive model that can be viewed as a tree where each branch of the tree is a classification question and leaves represent the partition of the data set with their classification. Investopedia defines a Decision Tree as a schematic tree-shaped diagram used to determine a course of action or show a statistical probability [12]. Decision trees can be viewed from the business perspective as creating a segmentation of the original data set. Thus marketing managers make use of segmentation of customers, products and sales region for predictive study. These predictive segments derived from the decision tree also come with a description of the characteristics that define the predictive segment. Because of their tree structure and skill to easily generate rules the method is a favored technique for building understandable models.

1) *An Instance of Decision Tree:* For instance if we were going to classify churners in the Cellular Telephone Industry, a decision tree might look something like that found in Figure 4.

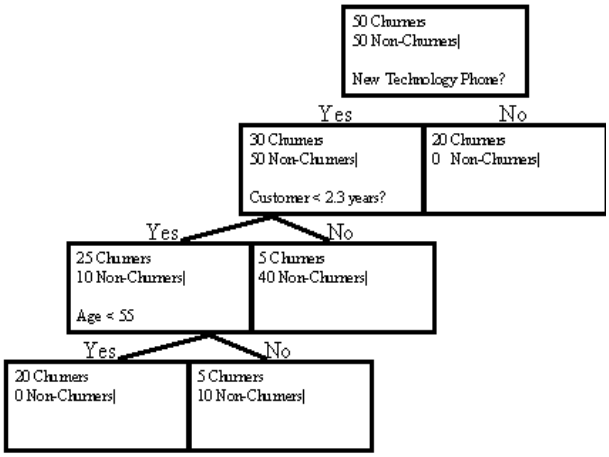

Fig 4 A decision tree is a predictive model that makes a prediction on the basis of a series of decision much like the game of 20 questions.[11]

It should be noticed that the segmentation above is done for the purpose of prediction and easy to understand the number churners coming under various categories like smart phone users, customers using the same network for certain number of years or various age groups. But there is always a conflict on the question selected to make the decision tree whether it is a good question or a bad question. For example in this case, change the likelihood of a prepaid or postpaid customer appearing in the customer segment. Decision tree models can be easily translated into SQL with their high level automation for deployment in relational databases helps enable the technology to be easy to integrate with the existing IT process with little preprocessing and cleansing of the data for data mining.

1) *Decision Tree Algorithms:* Numerous algorithms have been developed in the area of Decision Trees. Rose Quinlan introduced a decision tree algorithm in the late 1970s namely ID3. It was one of the first decision tree algorithms which was used for tasks like playing strategies for chess. Then it has been applied to various problems in different areas like industries and academia. ID3 was later improved and introduced as a version called C4.5 which can be used in different areas where predictors with missing or continuous values can be used or pruning or rule derivation is required. Two popular decision tree algorithms in practice today are CART and CHAID [13].

   - *Cart:* CART stands for Classification and Regression Trees and is a data analysis and prediction algorithm. This algorithm was developed in 1984 by Leo Breiman, Jerome Friedman, Charles Stone and Richard Olshen. CART tree chooses each predictor on how well it teases the other records with different predictions. The greatest advantage of CART is the validation of the model. This can be accomplished by building a very complex tree and pruning it back to the optimally general tree based on the results of validation. CART algorithms can be used to predict on new data since it is relatively robust. CART will make use of the maximum information it has, to make a decision for selecting the best possible split.
   - *Chaid:* Another popular decision tree algorithm is CHAID or Chi-Square Automatic Interaction Detector. Instead of choosing the optimal splits CHAID relies on chi square test to determine which predictor is furthest

from independence with prediction values. All predictors must be categorized because CHAID depends on the contingency tables to form its test of significance for each predictor.

D. Neural Networks

Neural Network or an artificial neural network is a biological system that detects patterns and makes predictions. The greatest breakthroughs in neural network in recent years are in their application to real world problems like customer response prediction, fraud detection etc. Since the neural networks are automated to a greater level it is really easy for a user to use this. Data Mining techniques such as neural networks are able to model the relationships that exist in data collections and can therefore be used for increasing business intelligence across a variety of business applications [14]. This powerful predictive modeling technique creates very complex models that are really difficult to understand by even experts. Neural Networks are used in a variety of applications. Some of the application areas include detecting the fraudulent use of credit card, automated driving of unmanned vehicle for military use etc.

Neural networks can also be used for clustering. Clusters are created by forcing the system to compress the data by creating prototypes. One important problem in data mining determination of predictors that is most relevant and accurate at prediction. These predictors are used in conjunction with other predictors to form features. The neural network functions by receiving predictor vales on the input nodes and executing calculations on these values to generate a new value. This value will denote the prediction from the neural network model. In neural network there is a row of nodes called hidden nodes which exist between the input and output nodes. These are invisible to the end user and have no predefined meaning. The neural network shown in fig: 4 below is used to extract features by the network to recreate the input data at output nodes by using just 5 hidden nodes.

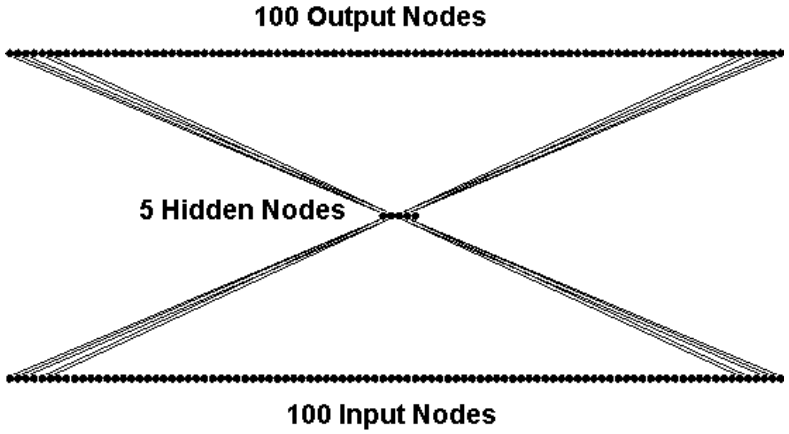

Fig 5 Neural networks can be used for data compression and feature extraction [11]

1) *Neural Networks - Different types:* Back propagation is a method of training that takes place in the neural networks. It stands for the propagation of error backwards from the output nodes through the hidden layers to the input nodes. Back propagation uses a search technique called gradient descent to hunt for the best probable improvement in the link weights to diminish the error. There are some other methods of high dimensional search namely Newton's methods and conjugate gradient. Since it is simple, easy to understand and works in large number of problem domains, back propagation learning procedure is most widely accepted by multiple users. Two other neural architectures that are more often used include Kohonen feature maps and Radial Basis Function Networks. Kohonen feature maps are used for clustering and unsupervised learning whereas Radial Basis function networks are used for supervised learning. Kohonen feature maps were developed in 1970's to simulate certain brain function. These are feed forward neural networks with no hidden layer.

There are two main structures of consequences in neural network:

- The Node – it corresponds to the neuron in human brain
- The Link – it corresponds to the connections between neurons (like axons or dendrites in human brain).

E. Rule Induction

Data mining may be viewed as the extraction of patterns and models from observed data [15]. Rule induction is one of the major forms of data mining in the unsupervised category. The method bring to mind the process that people think about when whenever they think about data mining as something like mining for gold from a huge database. The method takes significance when all possible patterns are checked and pulled out accordingly for which accuracy and significance is added to find out how strong the pattern is and their likeliness to occur again. These rules that are drawn from the database are extracted and ordered to be open to the user based on the percentage of times that they are correct and how often they apply.

1) An Instance of Rule Induction: In broad these rules are moderately simple such as for a market basket database of items scanned in a shopper market basket one might find motivating correlations in the database such as:

- If rice-noodle is purchased then celery is purchased 90% of the time and this pattern occurs 3% of all shopping baskets.

- If plants are purchased from a nursery then plant fertilizer is purchased 60% of the time and these two items are bought together in 6% of the shopping baskets in the nursery.

These rules that are drawn from the database are extracted and ordered to be open to the user based on the percentage of times that they are correct and how often they apply. Sometimes the strength of the system can be botheration as well as at times users get stunned with a large number of rules through all of them. This may lead to conflicting predictions that may seem to be similarly interesting. Ultimately this requires a second checking of the rules to select the best pattern. This has led the rule induction system to be highly automated and considered the one of the best tools for finding predictive patterns in a database. So to an extent they also act like a bench of counselors having different opinions on a particular issue but with a good number of explanations from each. For the rules to be useful there are two pieces of information that must be supplied as well as the actual rule:

- Accuracy - How often is the rule correct?
- Coverage - How often does the rule apply?

Only because the pattern in the data base is expressed as rule, it does not mean that it is true always. So like data mining algorithms it is equally important to identify and make obvious the uncertainty in the rule. This is called accuracy. The coverage of the rule means how much of the database it "covers" or applies to. Examples of these two measure for a variety of rules is shown in Table 2.

Sometimes accuracy is called the *confidence* of the rule and coverage is called the *support*. Accuracy and coverage appear to be the preferred ways of naming these two measurements.

TABLE II

Examples of Rule Accuracy and Coverage

Rule	**Accuracy**	**Coverage**
If breakfast cereal purchased then milk purchased.	85%	20%
If bread purchased then butter purchased.	15%	6%

Generally a rule consists of two values. A left hand antecedent and a right hand consequent. An antecedent can have one or multiple conditions which must be true in order for the consequent to be true for a given accuracy whereas a consequent is just a single condition. Thus while mining a rule from a database antecedent, consequent, accuracy, and coverage are all targeted. Sometimes "interestingness" is also targeted used for ranking. The situation occurs when rules have high coverage and accuracy but deviate from standards. It is also essential to note that even though patterns are produced from rule induction system, they all not necessarily mean that a left hand side ("if "part) should cause the right hand side ("then") part to happen. Once rules are created and interestingness is checked they can be used for predictions in business where each rule performs a prediction keeping a consequent as the target and the accuracy of the rule as the accuracy of the prediction which gives an opportunity for the overall system to improve and perform well.

TABLE III

Accuracy and Coverage in Rule Antecedents and Consequents

Antecedent	Consequent	Accuracy	Coverage
Breakfast Cereal	milk	80%	30%
Potato Chips	Dip	80%	5%
Potato Chips	Fizzy drinks	40%	3%
Potato Chips	Juice	40%	2%
Bread	Butter	65%	20%
eggs	milk	35%	15%
cheese	milk	40%	8%

Thus in general rules are used to find interesting patterns in the database but they are also used at times for prediction. The two main factors that are important to understand a rule is:

- Accuracy – This speak of the probability that if the antecedent is true then the precedent will be true. High accuracy means that this is a rule that is highly dependable.
- Coverage – This speak of the number of records in the database that the rule applies to. High coverage means that the rule can be used very often and also that it is less possible to be a fake artifact of the sampling technique.

2) *An instance of rule evaluation:* An instance of a shopping market database, in order to calculate the accuracy and coverage for a simple rule, we can consider T = 100 = Total number of shopping baskets in the database, N = 30 = Number of baskets with potato chips in them, C = 40 = Number of baskets with dip in them, B = 20 = Number of baskets with both potato chips and dip in them.

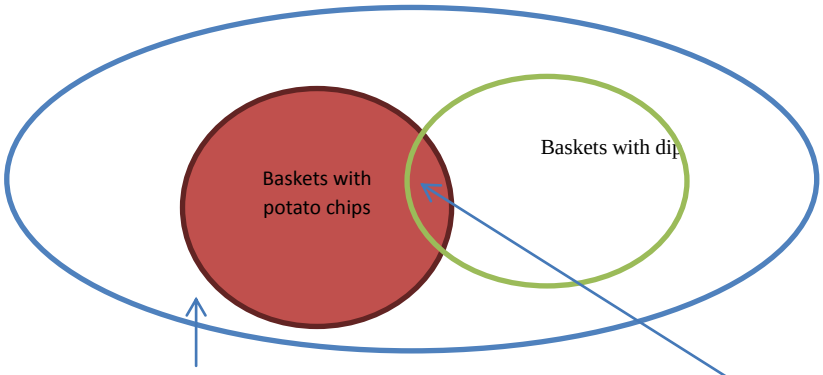

Fig 6 Graphically the total number of shopping baskets can be represented in a space and the number of baskets containing potato chips or dip can be represented by the area of a circle. The coverage of the rule "If Potato Chips then Dip" is just the relative size of the circle corresponding to Potato Chips. The accuracy is the relative size of the overlap between the two to the circle representing Potato Chips purchased.

Thus rules from rule induction system are generally generated from a database by adding more constraints to it until the coverage becomes very small and meaningless. This means rules can have families or specialty where a general rule becomes the parent of many specific rules.

IV. CONCLUSION

Undoubtedly one of the toughest things to do when deciding to device a data mining technique is the determination of which technique to use and when. Most of the time the technique to be used are determined by trial and error. There are certain differences in the kinds of problems that are most favorable to each technique but the authenticity of real world data and the dynamic way in which markets, customers and data that represents them is formed which means that the data is continuously changing. These dynamics mean that is not always possible to build a "perfect" model on the historical data since whatever was known in the past cannot adequately predict the future. But sometimes the situation is very crucial for business person who is waiting for all information to come in before they make their decision. Since business, economy and even the world are changing in unpredictable and even chaotic ways data mining techniques are not always predictable. So it's safe to select a robust model that may under-perform when compared to best data mining tools for analysis and execution at the earliest to take business decisions before it's too late.

REFERENCES

[1]. Ian H. Witten and Eibe Frank, *Data Mining- Practical Machine Learning Tools and Techniques*- second Edition.
[2]. Data mining definitions available at http://cplus.about.com/od/glossar1/g/Datamining.htm-
[3]. Data Mining definition available at http://www.gartner.com/it-glossary/data-mining/
[4]. Hand David, Manilla Heikki and Smyth Padhaic. (2001) *Principles of Data Mining*, The MIT Press, pp.9-10.
[5]. Alex Berson, Stephen Smith and Kurt Thearling- *Building Data Mining Applications for CRM -An Overview of Data Mining Techniques.*
[6]. Loveland Kumar Grover and Rajni Mehra , Guru Nanak Dev University, India, BBK DAV College for Women, India, *The Lure of Statistics in Data Mining,Journal of Statistics Education* Volume 16, *Number* 1 (2008).
[7]. Yoav Benjamini,Department of Statistics, School of Mathematical Sciences, Sackler Faculty for Exact, Sciences Tel Aviv University, Moshe Leshno, Faculty of Management and Sackler Faculty of Medicine Tel Aviv University, *Statistical methods for data mining.*
[8]. Siva Ganesh, Massey University, New Zealand, and ICOTS6, 2002: Ganesh, *Data mining: should it be included in the 'statistics' curriculum?*
[9]. DSstar Clustering and Classificatin –Data Mining Approaches by Ed Colet.
[10]. Cluster analysis available at http://www.yorku.ca/ptryfos/f1500.pdf
[11]. Alex Berson, Stephen Smith and Kurt Thearling- *Building Data Mining Applications for CRM -An Overview of Data Mining Techniques.*
[12]. Decision Trees available at http://www.investopedia.com/terms/d/decision-tree.asp#ixzz2KrflvpyR
[13]. Han, J. & Kamber, M.2001. *Data Mining: Concepts and techniques. Morgan-Kaufman Series of Data Management Systems.* San Diego: Academic Press.
[14]. Smith Kate and Gupta Jatinder. (2002) *Neural Networks in Business: Techniques and Applications*, IGI Publishing.
[15]. Berry, M. J. A., & Linnof, G, *Data mining Techniques*, New York: Wiley, (1997).

