ISSN:0975-9646
Vinodhini Katiyaar et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (3) , 2012,4041 - 4049

Use of Data Mining & Neural Network in Commercial Application

Dr Vinodini Katiyaar
Shri Ramswaroop Memorial College of Engineering & Management

Ina Kapoor Sharma
B.N College of Engineering & Technology

Abstract: Although neural networks are applied in a wide range of learning and commercial applications even though this method is not commonly used for data mining tasks
Companies have been collecting data for decades and have build up massive data warehouses and used data mining techniques to extract the value of data but only few practitioners have used neural networks in data mining though this method has proven successful in many situations.
This paper will try to prove how neural networks can be preferred as an effective tool for data mining practitioners and how both these technologies together can be very successful form a commercial perspective.

1. INTRODUCTION:

1.1 Data mining:

Data mining helps end users extract useful business information from large databases. It is the exploration and analysis of large quantities of data in order to discover meaningful patterns and rules. The type of data stored depends largely on the type of industry and the company.
With the continuous development of database technology and the extensive applications of database management system, the data volume stored in database increases rapidly and in the large amounts of data much important information is hidden. If the information can be extracted from the database they will create a lot of potential profit for the companies,
Data mining tools can forecast the future trends and activities to support the decision of people. Through analyzing the whole database system of the company the data mining tools can answer the many problems. Some data mining tools can also resolve some traditional problems which consumed much time, this is because that they can rapidly browse the entire database and find some useful information experts unnoticed.
The goal of data mining is to allow a company to improve its marketing, sales, and customer support operations through a better understanding of its customers.
Data mining are of two types directed and undirected. Directed data mining attempts to explain some particular target field such as income or response. Undirected data mining attempts to find patterns or similarities among groups of records without the use of a particular target field or collection of predefined classes

1.2 Tasks that can be performed with Data Mining

Many tasks can be performed with data mining to solve the problems of intellectual, economic and business interest.
Following are the few tasks that can be performed:

- Classification
- Estimation
- Prediction
- Affinity grouping
- Clustering
- Description and Profiling

Classification is the process of examining the features of a newly presented object and assigning it to one of a predefined set of objects
This task is characterized by a well defined definition of the classes, and a training set consisting of pre classified examples
Estimation deals with continuously valued outcomes. In this some input data is given and estimation comes up with a value for some unknown continuous variable.
In real life, estimation is often used to perform a classification task. This approach has the great advantage that the individual records can be rank ordered according to the estimate.
Prediction is use to classify the records according to some predicted future behavior or estimated future value. The main reason for treating prediction as a separate task from classification and estimation is that in predictive modeling there are additional issues regarding the temporal relationship of the input variables or predictors to the target variable
Affinity grouping is the task is to determine which things go together. It is the simple approach for generating rules from data
Clustering is the task of segmenting a heterogeneous population in to a number of more homogeneous subgroups or clusters. Clustering is like classification but the groups are

not predefined, so the algorithm will try to group similar items together

Clustering is often done as a prelude to some other form of data mining or modeling

Profiling is used when the purpose of data mining is simply to describe what is going on in a complicated database in a way that increases our understanding of the people , products, or processes that produced the data in the first place A good enough description of a behaviour will often suggest an explanation for it as well.

1.3 Approaches to Data Mining Problems

The approaches to data mining problems are based on the type of information / knowledge to be mined.

The **classification** task maps data into predefined group of classes. The class of a tuple is indicated by the value of a user-specified goal attribute. Tuples consists of a set of predicating attributes and a goal attribute. The task, is to discover, some kind of relationship between the predicating attributes and the goal attribute, so that, the discovered information can be used to predict the class of new tuple.

The task of **clustering** is to group the tuples with similar attributes values into same class. Given a database of tuples and an integer value k, the clustering is to define a mapping, such that, tuples are mapped to different cluster.

The principle is to maximize the intra class similarity and minimize the interclass similarity. In clustering, there is no goal attribute. So, classification is supervised by the goal attribute, while clustering is an unsupervised classification.

The task of **association rule** mining is to search for interesting relationships among items in a given data set. Its original application is on “market basket data”. The rule has the form x->y, where x and y are sets of items and they do not intersect. Each rule has two measurements, support and confidence. Given the user-specified minimum support and minimum confidence, the task is to find, rules with support and confidence above, minimum support and minimum confidence.

Decision tree learning, used in data mining and machine learning, uses a decision tree as a predictive model which maps observations about an item to conclusions about the item's target value. More descriptive names for such tree models **are classification trees** or regression trees. In these tree structures, leaves represent classifications and branches represent conjunctions of features that lead to those classifications.

In decision analysis, a decision tree can be used to visually and explicitly represent decisions and decision making. In data mining, a decision tree describes data but not decisions; rather the resulting classification tree can be an input for decision making.

Data mining tools can answer business questions that traditionally were very time consuming to resolve. They scour databases for hidden patterns, finding predictive information that experts may miss because it lies outside their expectations.

Most companies already collect and refine massive quantities of data. Data mining techniques can be implemented rapidly on existing software and hardware platforms to enhance the value of existing information resources, and can be integrated with new products and systems as they are brought on-line. When implemented on high performance client/server or parallel processing computers, data mining tools can analyze massive databases to deliver answers to questions such as, "Which clients are most likely to respond to my next promotional mailing, and why?"

. The process of data mining consists of three stages:

(1) The initial exploration,

(2) Model building or pattern identification with validation/verification, and

(3) Deployment (i.e., the application of the model to new data in order to generate predictions).

Data mining is primarily used today by companies with a strong consumer focus - retail, financial, communication, and marketing organizations. It enables these companies to determine relationships among "internal" factors such as price, product positioning, or staff skills, and "external" factors such as economic indicators, competition, and customer demographics. And, it enables them to determine the impact on sales, customer satisfaction, and corporate profits. Finally, it enables them to "drill down" into summary information to view detail transactional data.

With data mining, a retailer could use point-of-sale records of customer purchases to send targeted promotions based on an individual's purchase history. By mining demographic data from comment or warranty cards, the retailer could develop products and promotions to appeal to specific customer segments.

While data mining can be used to uncover patterns in data samples, it is important to be aware that the use of non-representative samples of data may produce results that are not indicative of the domain. Similarly, data mining will not find patterns that may be present in the domain, if those patterns are not present in the sample being "mined". There is a tendency for insufficiently knowledgeable "consumers" of the results to attribute "magical abilities" to data mining, treating the technique as a sort of all-seeing crystal ball. Like any other tool, it only functions in conjunction with the appropriate raw material: in this case, indicative and representative data that the user must first collect. Further, the discovery of a particular pattern in a particular set of data does not necessarily mean that pattern is representative of the whole population from which that data was drawn. Hence, an important part of the process is the verification and validation of patterns on other samples of data.

Data mining is a powerful new technology with great potential to help companies focus on the most important information in the data they have collected about the behaviour of their customers and potential customers. It discovers information within the data that queries and reports can't effectively reveal.

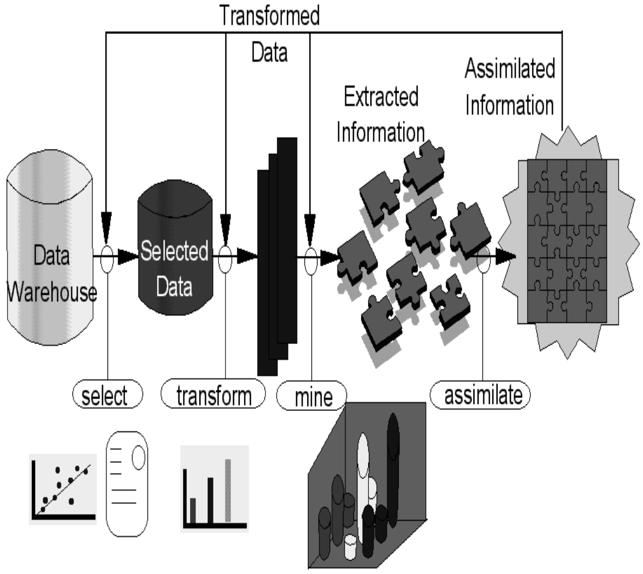

Figure 1

In addition to these improved data management tools, the increased availability of information and the decreasing costs of storing it have also played a role. Over the past several years there has been a rapid increase in the volume of information collected and stored, with some observers suggesting that the quantity of the world's data approximately doubles every year. At the same time, the costs of data storage have decreased significantly from dollars per megabyte to pennies per megabyte.

Similarly, computing power has continued to double every 18-24 months, while the relative cost of computing power has continued to decrease.

Data mining has become increasingly common in both the public and private sectors. Organizations use data mining as a tool to survey customer information, reduce fraud and waste, and assist in medical research. However, the proliferation of data mining has raised some implementation and oversight issues as well. These include concerns about the quality of the data being analyzed, the interoperability of the databases and software between agencies, and potential infringements on privacy.

Data mining is a powerful new technology with great potential to help companies focus on the most important information in the data they have collected about the behavior of their customers and potential customers (Rao and Vidyavathi, 2010). Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data set. These tools can include statistical models, mathematical algorithm and machine learning methods. It discovers information within the data that queries and reports can not effectively reveal.

The explosive growth in data stored in databases and data warehouses has generated an urgent need for new techniques that can intelligently transform this huge amount of data into useful knowledge. Consequently, data mining has become an important research area (Chen *et al*., 1996). Data mining differs from other data analysis techniques in that the system takes the initiative to generate patterns by itself. Data mining is concerned with the algorithmic means by which patterns, changes, anomalies, rules and statistically significant structures and events in data are extracted from large data sets (Grossman *et al*., 1999). Data mining studies can be classified into two generations. Studies in the first generation have focused on which kinds of patterns to mine. Studies in the second generation have focused on how mining can interact with other components in the framework like DBMS (Johnson *et al*., 2000).

1.4 Association Rule Mining (ARM):

Association rule mining is the discovery of associations or connections among objects. Since its inception, association rule mining has become one of the core data-mining tasks and has attracted tremendous interest among researchers and practitioners. ARM is undirected or unsupervised data mining over variable-length data and it produces clear, understandable results. It has an elegantly simple problem statement. Association rule mining has a wide range of applicability such market basket analysis, medical diagnosis/research, Website navigation analysis, homeland security and so on. An association rule is in the form of A1 ^ ....^ Ai → B1 ^ .... ^ Bj which means objects B1; ......;Bj tend to appear with objects A1; ……..; Ai in the target data. Association rules at multiple conceptual levels will reveal such kind of association in the relevant set (s) of data in a database. For example one may discover that a set of symptoms often occur together with another set of symptoms and then further study the reasons behind this association.

The conventional algorithm of association rules discovery proceeds in two steps. All frequent item sets are found in the first step. The frequent item set is the item set that is included in at least minsup transactions. The association rules with the confidence at least minconf are generated in the second step. According to Hipp *et al*. (2000), there is a border that separates the frequent item sets from the infrequent ones thus the problem is restricted on finding that border.

Modern organizations are geographically distributed. Typically, each site locally stores its ever increasing amount of day to day data. Using centralized data mining to discover useful patterns in such organizations data is not always feasible because merging data sets from different sites into a centralized site incurs huge network communication costs. Data from these organizations are not only distributed over various locations but also vertically fragmented making it difficult if not impossible to combine them in a central location. Distributed data mining has thus emerged as an active sub-area of data mining research. Therefore, this study proposes an agent-based architecture for a distributed Association Rule Mining in performing the mining process.

The field of distributed data mining has therefore gained increasing importance in the last decade. The Apriori algorithm by Agrawal and Srikant (1994) has emerged as one of the best Association Rule mining algorithms. It also serves as the base algorithm for most parallel algorithms. The enormity and high dimensionality of datasets typically available as input to problem of association rule discovery, makes it an ideal problem for solving on multiple processors in parallel. The primary reasons are the memory and CPU speed limitations faced by single processors.

The formal definition of association rule mining is: Let I = {i1,i2,…,im} be a set of literals called items and D be a set of transactions where each transaction T is a set of items such that T□I. Associated with each transaction is a unique identifier, called its TID. We say that a transaction T contains X, a set of some items in I, if X□T.

Association rule mining process could be decomposed into two main phases to enhance the implementation of the algorithm. The phases are:

Frequent item generation: This is to find all the item sets that satisfy the minimum support threshold. The item sets are called frequent item sets.

Rule generation: This is to extract all the high confidence rules from the frequent item sets found in the first step. These rules are called strong rules.

Association rules: An association rule is an implication expression of the form X□Y where X□I, Y□I and X and Y are disjoint item sets, i.e., X∩Y = φ. The strength of an association rule can be measured in terms of its support and confidence. The rule X□Y holds in the transaction set D with confidence c and support s, if c% of the transactions in D that contains X also contains Y and s% of transactions in D contains X□Y. Both the antecedent and the consequent of the rule could have more than one Item. The formal definitions of these two metrics are:

$$\text{Supports, s } (X \Rightarrow y) = \frac{\Sigma(x \cup y)}{n}$$

$$\text{Confidence, c } (X \Rightarrow y) = \frac{\Sigma(x \cup y)}{\Sigma x}$$

There have been many algorithms developed for mining frequent patterns which can be classified into two categories:

- Pattern-growth methods
- Candidate-generation-and-test

1.5 What Can Data Mining Do?

Although data mining is still in its infancy, companies in a wide range of industries - including retail, finance, health care, manufacturing transportation, and aerospace - are already using data mining tools and techniques to take advantage of historical data. By using pattern recognition technologies and statistical and mathematical techniques to sift through warehoused information, data mining helps analysts recognize significant facts, relationships, trends, patterns, exceptions and anomalies that might otherwise go unnoticed.

For businesses, data mining is used to discover patterns and relationships in the data in order to help make better business decisions. Data mining can help spot sales trends, develop smarter marketing campaigns, and accurately predict customer loyalty. Specific uses of data mining include:

- **Market segmentation -** Identify the common characteristics of customers who buy the same products from your company.
- **Customer churn -** Predict which customers are likely to leave your company and go to a competitor.
- **Fraud detection -** Identify which transactions are most likely to be fraudulent.
- **Direct marketing** - Identify which prospects should be included in a mailing list to obtain the highest response rate.
- **Interactive marketing -** Predict what each individual accessing a Web site is most likely interested in seeing.
- **Market basket analysis -** Understand what products or services are commonly purchased together; e.g., beer and diapers.
- **Trend analysis -** Reveal the difference between a typical customer this month and last.

1.6 Data Mining Process

Data mining is an iterative process that typically involves the following phases:

Problem definition

A data mining project starts with the understanding of the business problem. Data mining experts, business experts, and domain experts work closely together to define the project objectives and the requirements from a business perspective. The project objective is then translated into a data mining problem definition.

In the problem definition phase, data mining tools are not yet required.

Data exploration

Domain experts understand the meaning of the metadata. They collect, describe, and explore the data. They also identify quality problems of the data. A frequent exchange with the data mining experts and the business experts from the problem definition phase is vital.

In the data exploration phase, traditional data analysis tools, for example, statistics, are used to explore the data.

Data preparation

Domain experts build the data model for the modeling process. They collect, cleanse, and format the data because some of the mining functions accept data only in a certain format. They also create new derived attributes, for example, an average value.

In the data preparation phase, data is tweaked multiple times in no prescribed order. Preparing the data for the modeling tool by selecting tables, records, and attributes, are typical tasks in this phase. The meaning of the data is not changed.

Modelling
Data mining experts select and apply various mining functions because you can use different mining functions for the same type of data mining problem. Some of the mining functions require specific data types. The data mining experts must assess each model.
In the modeling phase, a frequent exchange with the domain experts from the data preparation phase is required.
The modeling phase and the evaluation phase are coupled. They can be repeated several times to change parameters until optimal values are achieved. When the final modeling phase is completed, a model of high quality has been built.
Evaluation
Data mining experts evaluate the model. If the model does not satisfy their expectations, they go back to the modeling phase and rebuild the model by changing its parameters until optimal values are achieved. When they are finally satisfied with the model, they can extract business explanations and evaluate the following questions:

- Does the model achieve the business objective?
- Have all business issues been considered?

At the end of the evaluation phase, the data mining experts decide how to use the data mining results.
Deployment
Data mining experts use the mining results by exporting the results into database tables or into other applications, for example, spreadsheets.

1.7Artificial Neural Networks
Artificial Neural networks commonly referred to as **"neural networks"** is a mathematical model based on biological neural system, it an algorithm for optimization and learning based loosely on concepts inspired by research into the nature of the brain.
The brain is a highly complex, nonlinear and parallel computer. It has the capability to organize its structural constituents, known as neutrons, so as to perform certain computations many times faster than the digital computer In the general form neural network is a machine that is designed to model the way in which the brain performs a particular task or function of interest
According to **Simon Haykin** " A neural network is a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:

(1) Knowledge is acquired by the network from its environment through a learning process
(2) Interneuron connection strengths, known as a synaptic weights are used to store the acquired knowledge"

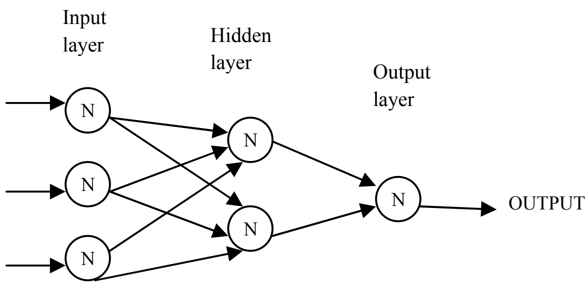

Figure 2: Neural Network

1.8 Neural Network Architecture
A neural network can be viewed as s weighted directed graph in which neurons are nodes and directed edges represent connection between neurons neuron network architecture can be classified in three classes:
Single layer feedforward networks is a layered neuron network in which neurons are organized in the form of layers, in this we have an input layer of source nodes that projects onto an output layer of neurons but not vice versa. This network is strictly a feedforward or acyclic type.
Multilayer feedforward networks are network which has one or more hidden layer and whose computation nodes are correspondingly called hidden neurons. The function of hidden neurons is to intervene between the external input and network output in some useful manner.
Recurrent networks has at least one feedback loop. It may consist of a single layer of neurons with each neuron feeding its output signal back to the inputs of all the other neurons. Recurrent networks also no hidden neurons
The model era of neural networks began with the pioneering work of McCulloch and Pitts (1943).Since then, many other significant contributions (Hebb 1949; Rosenblatt 1958; Minsky & Papert 1969; von der Malsburg 1973; Grossberg 1976; Kohonen 1982; Hopfield 1982; Rumelhart *et al.* 1986) have been made in the field. One of the most important developments in the field of neural networks was the development of an algorithm called *backpropagation algorithm* (Rumelhart *et al.* 1986) to train multilayered networks. Details of the back propagation algorithm will be presented later in the section.
Neural network models are inspired by natural physiology and are an attempt to mimic the neurons and synaptic connections of the brain (Hertz *et al.* 1991). Biological neurons transmit electrochemical signals through neural pathways. Each neuron receives signals from other neurons through special junctions called *synapses*. Some inputs tend to excite the neurons, others inhibit them. When the cumulative effect exceeds a threshold, the neuron fires and sends a signal to other neurons. The notion of an artificial neuron, created by the artificial intelligence (AI) researchers and simulated on a computer, models these simple biological

characteristics. Each artificial neuron is connected to a set of inputs; each input value is multiplied by a weight analogous to a synaptic strength, and the combination of these weighted values is what is acted upon by the artificial neuron.

Haykin (1999) has offered the following definition of a neural network viewed as an adaptive machine: A neural network is a parallel distributed processor made up of simple processing units, which has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:

- Knowledge is acquired by the network from its environment through a learning process;
- Interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge.

As data sets grow to massive sizes, the need for automated processing becomes clear. With their “model-free” estimators and their dual nature, neural networks serve data mining in a myriad of ways. (Kosko (1992)). But first, it is necessary to examine the architecture, the learning algorithms, and the activation functions in more detail.

Before we proceed it is interesting to note that several neural network models are similar or identical to statistical models.

- “Information processing occurs at many simple elements called neurons [also referred to as units, cells, or nodes].
- Signals are passed between neurons over connection links.
- Each connection link has an associated weight, which, in a typical neural net, multiplies the signal transmitted.
- Each neuron applies an activation function (usually nonlinear) to its net input (sum of weighted input signals) to determine its output signal.” (Faussett (1994), p. 3)

According to **Simon Haykin** “ A neural network is a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:

(3) Knowledge is acquired by the network from its environment through a learning process

(4) Interneuron connection strengths, known as a synaptic weights are used to store the acquired knowledge”

A neural network is a powerful data modeling tool that is able to capture and represent complex input/output relationships. The motivation for the development of neural network technology stemmed from the desire to develop an artificial system that could perform "intelligent" tasks similar to those performed by the human brain. Neural networks resemble the human brain in the following two ways:

1. A neural network acquires knowledge through learning.
2. A neural network's knowledge is stored within inter-neuron connection strengths known as synaptic weights.

The true power and advantage of neural networks lies in their ability to represent both linear and non-linear relationships and in their ability to learn these relationships directly from the data being modeled. Traditional linear models are simply inadequate when it comes to modeling data that contains non-linear characteristics.

1.9 Uses of Neural Network

Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyze. This expert can then be used to provide projections given new situations of interest and answer "what if" questions. Other advantages include:

1. **Adaptive learning:** An ability to learn how to do tasks based on the data given for training or initial experience.
2. **Self-Organization:** An ANN can create its own organization or representation of the information it receives during learning time.
3. **Real Time Operation:** ANN computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability.
4. **Fault Tolerance via Redundant Information Coding**: Partial destruction of a network leads to the corresponding degradation of performance. However, some network capabilities may be retained even with major network damage.

1.10 Training an Artificial Neural Network

Once a network has been structured for a particular application, that network is ready to be trained. To start this process the initial weights are chosen randomly. Then, the training, or learning, begins.

There are two approaches to training - supervised and unsupervised. Supervised training involves a mechanism of providing the network with the desired output either by manually "grading" the network's performance or by providing the desired outputs with the inputs. Unsupervised training is where the network has to make sense of the inputs without outside help.

The vast bulk of networks utilize supervised training. Unsupervised training is used to perform some initial characterization on inputs. However, in the full blown sense of being truly self learning, it is still just a shining promise that is not fully understood, does not completely work, and thus is relegated to the lab.

Supervised Training

In supervised training, both the inputs and the outputs are provided. The network then processes the inputs and compares its resulting outputs against the desired outputs. Errors are then propagated back through the system, causing the system to adjust the weights which control the network. This process occurs over and over as the weights are continually tweaked. The set of data which enables the training is called the "training set." During the training of a network the same set of data is processed many times as the connection weights are ever refined.

The current commercial network development packages provide tools to monitor how well an artificial neural network is converging on the ability to predict the right answer. These tools allow the training process to go on for days, stopping only when the system reaches some statistically desired point, or accuracy. However, some networks never learn. This could be because the input data does not contain the specific information from which the desired output is derived. Networks also don't converge if there is not enough data to enable complete learning. Ideally, there should be enough data so that part of the data can be held back as a test. Many layered networks with multiple nodes are capable of memorizing data. To monitor the network to determine if the system is simply memorizing its data in some no significant way, supervised training needs to hold back a set of data to be used to test the system after it has undergone its training. (Note: memorization is avoided by not having too many processing elements.)

If a network simply can't solve the problem, the designer then has to review the input and outputs, the number of layers, the number of elements per layer, the connections between the layers, the summation, transfer, and training functions, and even the initial weights themselves. Those changes required to create a successful network constitute a process wherein the "art" of neural networking occurs.

Another part of the designer's creativity governs the rules of training. There are many laws (algorithms) used to implement the adaptive feedback required to adjust the weights during training. The most common technique is backward-error propagation, more commonly known as back-propagation. These various learning techniques are explored in greater depth later in this report.

Yet, training is not just a technique. It involves a "feel," and conscious analysis, to insure that the network is not over trained. Initially, an artificial neural network configures itself with the general statistical trends of the data. Later, it continues to "learn" about other aspects of the data which may be spurious from a general viewpoint.

When finally the system has been correctly trained, and no further learning is needed, the weights can, if desired, be "frozen." In some systems this finalized network is then turned into hardware so that it can be fast. Other systems don't lock themselves in but continue to learn while in production use.

Unsupervised or Adaptive Training

The other type of training is called unsupervised training. In unsupervised training, the network is provided with inputs but not with desired outputs. The system itself must then decide what features it will use to group the input data. This is often referred to as self-organization or adaptation.

At the present time, unsupervised learning is not well understood. This adaption to the environment is the promise which would enable science fiction types of robots to continually learn on their own as they encounter new situations and new environments. Life is filled with situations where exact training sets do not exist. Some of these situations involve military action where new combat techniques and new weapons might be encountered. Because of this unexpected aspect to life and the human desire to be prepared, there continues to be research into, and hope for, this field. Yet, at the present time, the vast bulk of neural network work is in systems with supervised learning. Supervised learning is achieving results.

One of the leading researchers into unsupervised learning is Tuevo Kohonen, an electrical engineer at the Helsinki University of Technology. He has developed a self-organizing network, sometimes called an auto-associator, that learns without the benefit of knowing the right answer. It is an unusual looking network in that it contains one single layer with many connections. The weights for those connections have to be initialized and the inputs have to be normalized. The neurons are set up to compete in a winner-take-all fashion.

Kohonen continues his research into networks that are structured differently than standard, feedforward, back-propagation approaches. Kohonen's work deals with the grouping of neurons into fields. Neurons within a field are "topologically ordered." Topology is a branch of mathematics that studies how to map from one space to another without changing the geometric configuration. The three-dimensional groupings often found in mammalian brains are an example of topological ordering.

Kohonen has pointed out that the lack of topology in neural network models make today's neural networks just simple abstractions of the real neural networks within the brain. As this research continues, more powerful self learning networks may become possible. But currently, this field remains one that is still in the laboratory.

2. DATA MINING BASED ON NEURAL NETWORKS

In more technical terms neural networks are non-linear statistical data modeling tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data. Using neural networks as a tool, data warehousing firms are harvesting information from datasets in the process known as data mining. The difference between these data warehouses and ordinary databases is that there is actual manipulation and cross-fertilization of the data helping users makes more informed decisions.

The difference between general data mining and data mining based on neural networks is explained with the help of an example.

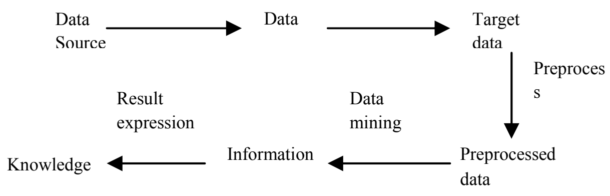

Figure 3: Data mining without neural networks

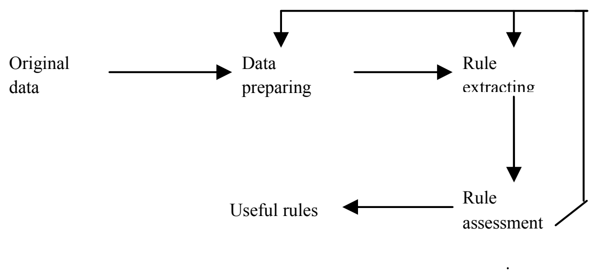

Figure 4: Data mining based on neural networks

A. Data Preparation

Data preparation is to define and process the mining data to make it fit specific data mining method. Data preparation is the first important step in the data mining and plays a decisive role in the entire data mining process. It mainly includes the following four processes.

1) Data cleaning

Data cleansing is to fill the vacancy value of the data, eliminate the noise data and correct the inconsistencies data in the data.

2) Data option

Data option is to select the data arrange and row used in this mining.

3) Data pre processing

Data pre processing is to enhanced process the clean data which has been selected.

4) Data expression

Data expression is to transform the data after preprocessing into the form which can be accepted by the data mining algorithm based on neural network. The data mining based on neural network can only handle numerical data, so it is need to transform the sign data into numerical data.

B. Rules Extracting

There are many methods to extract rules, in which the most commonly used methods are LRE method, black-box method, the method of extracting fuzzy rules, the method of extracting rules from recursive network, the algorithm of binary input and output rule extracting (BIO-RE), partial rules extracting algorithm (Partial-RE) and full rules extracting algorithm

C. Rules Assessment

Although the objective of rules assessment depends on each specific application, but, in general terms, the rules can be assessed in accordance with the following objectives.
(1) Find the optimal sequence of extracting rules, making it obtains the best results in the given data set;
(2) Test the accuracy of the rules extracted;
(3) Detect how much knowledge in the neural network has not been extracted;
(4) Detect the inconsistency between the extracted rules and the trained neural network

3. ADVANTAGES OF USING NEURAL NETWORKS IN DATA MINING

There are several advantages of using neural networks in data mining in many commercial applications:

(1) A neural network can be both linear and non linear so in data mining data can be distributed throughout the network and physical mechanism responsible for generation of data can also be maintained
(2) Neural networks have a built in capability to adapt their synaptic weights to changes in the surrounding environment which can be very useful in data mining
(3) Neural networks can be designed to provide information not only about which particular patter to select, but also the confidence in decision made which is very important aspect for data mining.
(4) In neural networks knowledge is represented in a very structural way which is also a desirable characteristic in data mining
(5) A neural network, implemented in hardware form has the potential to be inherently fault tolerant, or capable of robust computation which is an added advantage for data mining.
(6) Neural networks enjoy universality in analysis and design of data which is the same feature in data mining.
(7) Neural networks can be implemented in parallel hardware which can be very beneficial in data mining.
(8) If neural networks used in data mining than they can highly automate the performance and minimizing the human involvement
(9) Neural networks can easily be updated with fresh data that makes then very useful in dynamic environment and can be very helpful in data mining
(10) Neural networks can be very useful in case of noisy, missing and incomplete data and if it is used in data mining it can be very advantageous.

4. CONCLUSION:

There is rarely one right tool to use in data mining; it is a question as to what is available and what gives the "best" results. Many articles, in addition to those mentioned in this paper, consider neural networks to be a promising data mining tool. In most cases neural networks perform as well or better than the traditional statistical techniques to which they are compared. Resistance to using these "black boxes" is gradually diminishing as more researchers use them, in particular those with statistical backgrounds. Thus, neural networks are becoming very popular with data mining practitioners. This is because they have proven their predictive power through comparison with other statistical techniques using real data sets. Due to design problems neural systems need further research before they are widely accepted in industry. As software companies develop more sophisticated models with user-friendly interfaces the attraction to neural networks will continue to grow.

The key to gaining a competitive advantage can be found in recognizing that customer databases, if properly managed, analyzed, and exploited, are unique, valuable corporate assets. Firms can unlock the intelligence contained in their customer databases through modern data mining technology. Data mining uses predictive modelling, database segmentation, market basket analysis, and combinations thereof to more quickly answer crucial business questions with greater accuracy. New products can be developed and marketing strategies can be implemented enabling the insurance firm to transform a wealth of information into a wealth of predictability, stability, and profits.

REFERENCES:

[1]. Data Mining and Neural Networks from Commercial Point of View by Portia A Cerny

[2]. Neural networks in Data Mining by Dr Yashpal Singh, Alok Singh Chauhan

[3]. Database Mining: A Performance Perspective by Agrawal, R., Imielinski, T., Swami, A

[4]. Data Mining Techniques, by Berry, J. A., Lindoff, G

[5]. Data Mining News, Volume 1, No. 18, May 11, 1998

[6]. Data Mining '98, http://www.twocrows.com/

[7] Neural Networks in Data Mining by Cynthia Krieger