Use of Neural Network in Data Mining

Mrs. Archana Kalia
Lecturer, Information Technology Department,
VPM's Polytechnic, Thane, India

Abstract: Neural networks have emerged as advanced data mining tools in cases where other techniques may not produce satisfactory predictive models. As the term implies, neural networks have a biologically inspired modeling capability, but are essentially statistical modeling tools. In this paper, we try to understand the basics of neural network modeling, some specific applications, and the process of implementing a neural network in data mining.
Keywords: Data mining, Neural network, weighing factors, Transfer function

I.INTRODUCTION

Data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified. So, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases.

Data mining software analyzes relationships and patterns in stored transaction data based on open-ended user queries. Several types of analytical software are available: statistical, machine learning, and neural networks. Generally, any of four types of relationships are sought:

- **Classes**: Stored data is used to locate data in predetermined groups. For example, a restaurant chain could mine customer purchase data to determine when customers visit and what they typically order. This information could be used to increase traffic by having daily specials.

- **Clusters**: Data items are grouped according to logical relationships or consumer preferences. For example, data can be mined to identify market segments or consumer affinities.

- **Associations**: Data can be mined to identify associations. The beer-diaper example is an example of associative mining.

- **Sequential patterns**: Data is mined to anticipate behavior patterns and trends. For example, an outdoor equipment retailer could predict the likelihood of a backpack being purchased based on a consumer's purchase of sleeping bags and hiking shoes.

Data mining consists of five major elements:

- Extract, transform, and load transaction data onto the data warehouse system.

- Store and manage the data in a multidimensional database system.

- Provide data access to business analysts and information technology professionals.

- Analyze the data by application software.

- Present the data in a useful format, such as a graph or table.

Different levels of analysis are available:

- **Artificial neural networks**: Non-linear predictive models that learn through training and resemble biological neural networks in structure.

- **Genetic algorithms**: Optimization techniques that use process such as genetic combination, mutation, and natural selection in a design based on the concepts of natural evolution.

- **Decision trees**: Tree-shaped structures that represent sets of decisions. These decisions generate rules for the classification of a dataset. Specific decision tree methods include Classification and Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID). They provide a set of rules that can be applied to a new (unclassified) dataset to predict which records will have a given outcome.

- **Nearest neighbor method**: A technique that classifies each record in a dataset based on a combination of the classes of the k record(s) most similar to it in a historical dataset (where k 1). Sometimes called the k-nearest neighbor technique.

- **Rule induction**: The extraction of useful if-then rules from data based on statistical significance.

- **Data visualization**: The visual interpretation of complex relationships in multidimensional data is done by graphics tools which are used to illustrate data relationships.

I.   BASIC CONCEPT OF NEURAL NETWORK:

Neural networks represent a brain metaphor for information processing. These models are biologically inspired rather than an exact replica of how the brain actually functions.

Neural networks have been shown to be very promising systems in many forecasting applications and business classification applications due to their ability to "learn" from the data, their nonparametric nature, and their ability to generalize.

Neural computing refers to a pattern recognition methodology for machine learning. The resulting model from neural computing is often called an artificial neural network (ANN) or a neural network.

Biological neural networks are composed of many massively interconnected primitive biological neurons. Each neuron possesses axons and dendrites, finger-like projections that enable the neuron to communicate with its neighboring neurons by transmitting and receiving electrical and chemical signals. More or less resembling the structure of their counterparts, neural network is composed of interconnected, simple processing elements called artificial neurons. In processing information, the processing elements in an neural network operate concurrently and collectively in a similar fashion to biological neurons. Neural network possesses some desirable traits similar to those of biological neural networks, such as the capabilities of learning, self-organization, and fault tolerance.

II.   ARTIFICIAL NEURON MODEL

An artificial neuron is a mathematical function conceived as a simple model of a real (biological) neuron. The McCulloch-Pitts Neuron is a simplified model of real neurons, known as a Threshold Logic Unit.
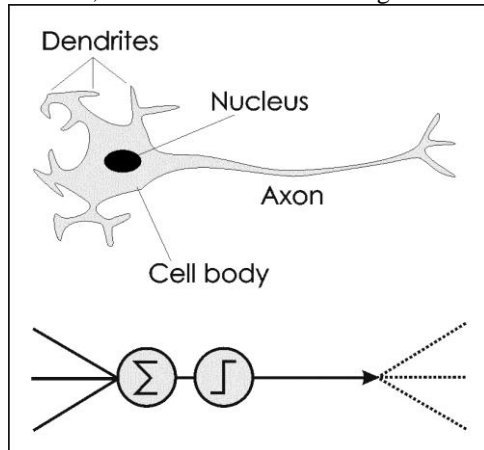
■A set of input connections brings in activations from other neurons.
■A processing unit sums the inputs, and then applies a non-linear activation function (i.e. squashing / transfer / threshold function).
■An output line transmits the result to other neurons.

In other words, the input to a neuron arrives in the form of signals. The signals build up in the cell. Finally the cell discharges (cell fires) through the output. The cell can start building up signals again.

III.   NEURAL NETWORK ARCHITECTURES

An Artificial Neural Network is a data processing system, consisting large number of simple highly interconnected processing elements as artificial neuron in a network structure that can be represented using a directed graph G, an ordered 2-tuple (V, E), consisting a set V of vertices and a set E of edges.

The vertices may represent neurons (input/output) and the edges may represent synaptic links labeled by the weights attached.

Example:
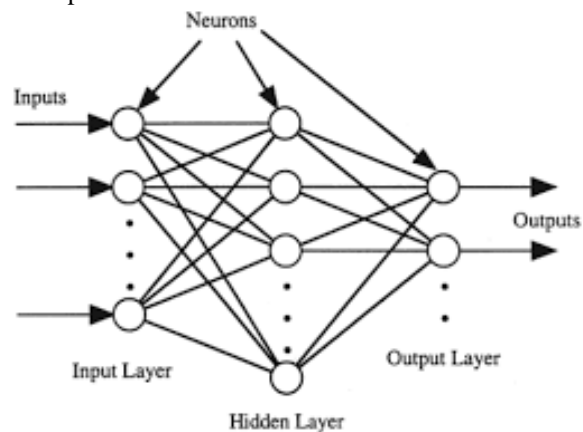
Seven major components make up an artificial neuron. These components are valid whether the neuron is used for input, output, or is in the hidden layers.

Component 1. Weighting Factors*:* A neuron usually receives many simultaneous inputs. Each input has its own relative weight, which gives the input the impact that it needs on the processing element's summation function. Some inputs are made more important than others to have a greater effect on the processing element as they combine to produce a neural response. Weights are adaptive coefficients that determine the intensity of the input signal as registered by the artificial neuron. They are a measure of an input's connection strength.

 Component 2*.* **Summation Function***:* The inputs and corresponding weights are vectors which can be represented as (i1, i2 . . . in) and (w1, w2 . . . wn). The total input signal is the dot product of these two vectors. The result; (i1 * w1) + (i2 * w2) +…….. + (in * wn) ; is a single number. The summation function can be more complex than just weight sum of products. The input and weighting coefficients can be combined in many different ways before passing on to the transfer function. In addition to summing, the summation function can select the minimum, maximum, majority, product or several normalizing algorithms. The specific algorithm for combining neural inputs is determined by the chosen network architecture and paradigm.

Component 3. Transfer Function*:* The result of the summation function is transformed to a working output through an algorithmic process known as the transfer function. In the transfer function the summation can be compared with some threshold to determine the neural output. If the sum is greater than the threshold value, the processing element generates a signal and if it is less than the threshold, no signal (or some inhibitory signal) is generated. Both types of response are significant. The threshold, or transfer function, is generally non-linear. Linear functions are limited because the output is simply proportional to the

input. The step type of transfer function would output zero and one, one and minus one, or other numeric combinations. Another type, the 'threshold' or ramping function, can mirror the input within a given range and still act as a step function outside that range. It is a linear function that is clipped to minimum and maximum values, making it non-linear.

Component 4. Scaling and Limiting: After the transfer function, the result can pass through additional processes, which scale and limit. This scaling simply multiplies a scale factor times the transfer value and then adds an offset. Limiting is the mechanism which insures that the scaled result does not exceed an upper, or lower bound. This limiting is in addition to the hard limits that the original transfer function may have performed.

Component 5. Output Function (Competition): Each processing element is allowed one output signal, which it may give to hundreds of other neurons. Normally, the output is directly equivalent to the transfer function's result. Some network topologies modify the transfer result to incorporate competition among neighboring processing elements. Neurons are allowed to compete with each other inhibiting 79 processing elements unless they have great strength. Competition can occur at one or both levels. First, competition determines which artificial neuron will be active or provides an output. Second, competitive inputs help determine which processing element will participate in the learning or adaptation process.

Component 6. **Error Function and Back-Propagated Value:** In most learning networks the difference between the current output and the desired output is calculated as an error which is then transformed by the error function to match particular network architecture. Most basic architectures use this error directly but some square the error while retaining its sign, some cube the error, other paradigms modify the error to fit their specific purposes. The error is propagated backwards to a previous layer. This back-propagated value can be either the error, the error scaled in some manner (often by the derivative of the transfer function) or some other desired output depending on the network type. Normally, this back-propagated value, after being scaled by the learning function, is multiplied against each of the incoming connection weights to modify them before the next learning cycle.

Component 7. Learning Function: Its purpose is to modify the weights on the inputs of each processing element according to some neural based algorithm.

IV.   DATA MINING PROCESS BASED ON NEURAL NETWORKS

Data mining process can be composed by three main phases: data preparation, data mining, expression and interpretation of the results, Data mining process is the reiteration of the three phases. The data mining based on neural network is composed by data preparation, rules extracting and rules assessment three phases:

 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 Clustering: 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. The simplest method is to establish a table with one-to-one correspondence between the sign data and the numerical data. The other more complex approach is to adopt appropriate Hash function to generate a unique numerical data according to given string. Although there are many data types in relational database, but they all basically can be simply come down to sign data, discrete numerical data and serial numerical data three logical data types.

B. Data Preparation: 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 rules extracting (BIO-RE), partial rules extracting algorithm (Partial-RE) and full rules extracting algorithm (Full-RE).

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; Detect the inconsistency between the extracted rules and the trained neural network.

VI. APPLICATION OF NEURAL NETWORK

Neural network has been applied in many domains. There have been several tests of neural networks in financial markets. Neural network model for commodity training would have resulted in significant profits over other trading strategies. Neural networks provided a more general framework for connecting financial information of a firm to the respective bond rating. Fraud prevention is another area of neural network application in business. Another significant area of statistical application of neural networks is in time series forecasting. Several studies have attempted to use neural networks for time-series prediction. Because neural networks have been a subject of intense study since late 1980s, there have been many applications of as well as experiments with applications.

VII.CONCLUSION

Neural Networks are suitable for problems, whose inputs are both categorical and numeric, and where the relationships

between inputs and outputs are not linear or the input data are not normally distributed. In such cases, classical statistical methods may not be reliable enough. Because neural network do not make any assumptions about the data distribution, their power is less affected than traditional statistical methods when data are not properly distributed. Finally, there are cases in which the neural networks simply provide one more way of building a predictive model for the situation at hand. Given the ease of experimentation using the available software tools, it is certainly worth exploring the power of neural networks in any data modeling situation.

ACKNOWLEDGMENT

REFERENCES
1. Data mining and Neural Networks from a   Commercial Perspective by Portia A. Cerny   Data Analyst & Modeller Aim Proximity, Auckland, New Zealand
2. Data Mining with Neural Networks by Svein Nordbotten
3. Introduction to Neural Networks by Kevin Swingler / Bruce Graham
4. Data Mining: Concepts and Techniques by Jiawei Han and Micheline Kamber Simon Fraser University.

BIOGRAPHY

Mrs. Archana Kalia ,Working as a Lecturer in Information Technology Department at V.P.M.'s Polytechnic, Thane.Professional qualifications,B.E. [Comp. Sc.], Seemanta Engineering College, Jharpokharia, Mayurbhanj, OdishaPaper Published,1.Integration of Big Data and Relational Database published in International Journal of Recent and Innovation Trends in Computing andCommunication ISSN: 2321-8169 Journal, Volume 4, Issue 1, January 2016,NoSQL Databases published in Think Quest International Journal of Technology and Management (TQIJTM) ISSN: 2320-4214 Volume 2, Issue 2, August 2014