Data Mining with Neural Network: A Perspective

Foram kalyani
Computer Department , Atmiya Institute of Technology and Science , Rajkot , Gujarat , India

Abstract - With the increasing applications of database management systems, large amount of important data is stored and much of its information is hidden. The technology used to extract information from data is Data Mining. Various tools are available to forecast the trends that will support decision of people.

Neural Networks or Artificial Neural Networks (ANN) have been a promising system in many applications due to their learning ability from data and generalization ability. Neural Networks are used for prediction, classification, forecasting, and pattern recognition. This paper provides a brief overview of Data Mining with the Neural Network, its tools and process.

Keywords: Neural Networks, Parallel distributed processing, Neural computing, Neuroscience, Data Mining, Knowledge Mining

I. INTRODUCTION

Data Mining describes process of extracting values from database. Data Mining is named as "Knowledge mining from data" or "Knowledge mining" [1]. Some data mining tools solve problems taking more time and remaining unnoticed by the experts.

The Neural Network is a common technique used to solve problems related to data mining [2]. Neural Network is parallel distributed processing networks which have tightly bound network formation.

Neural Network has complex structure, long training time and poor interpretability, the main reasons for their defect in data mining. But the advantages of neural network are high affordability for the noise data and low error rate [3].

A. Artificial Neural Network

Artificial Neural Network is similar to human brains. It is a biological neuron model consisting of number of neurons modeled after human nerve cell.

Each neuron takes number of inputs and performs operation on it. The output of this neuron is applied as input to other neuron.
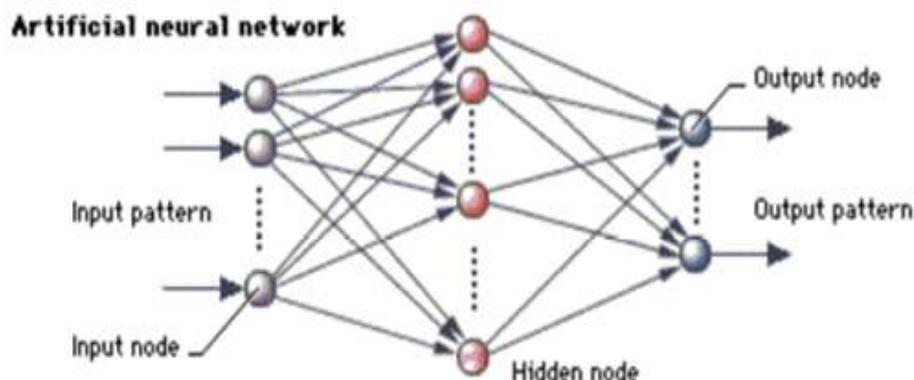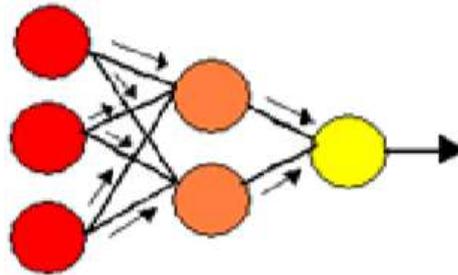

*Figure 1. Artificial Neural Network [4]*

There are two types of Neural Network formations:

- FeedForward Neural Network: The information can pass only in one direction i.e. in forward direction, from input node to output node via hidden nodes (if present) with no loops.

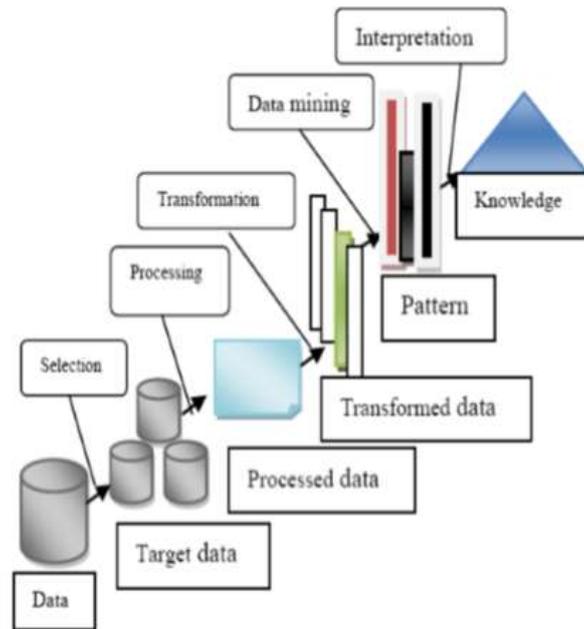*Figure 2. FeedForward Neural Network example [5]*

- Recurrent Network: The information can propagate bi – directionally in linear manner from input node to output node.
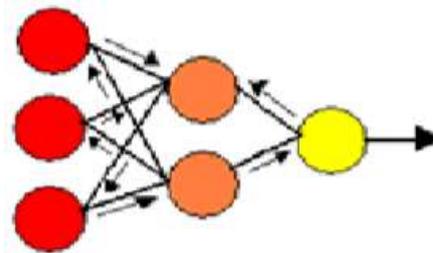
Figure 3. Recurrent Neural Network example [5]

B.   Data Mining using Neural Network

Data Mining is used to extract rules and patterns from large datasets. Analyzing the data contains several stages

Figure 4. Data Processing [3]

- **Cleaning:** Removing Noise from data.

- **Integration:** Combining multiple data into single one.

- **Selection:** Pre – processed data is retrieved from database.

- **Transformation:** Data is transformed into Standard format using various aggregation and summarization operation.

- **Mining:** Various tools and techniques are applied to extract pattern from data.

- **Evaluation of pattern:** Identifies tree pattern for knowledge representation.

- **Knowledge Representation:** Knowledge representation techniques are used for interpretation of data mining.
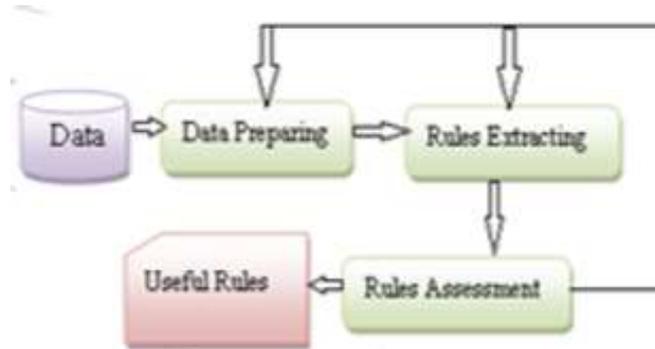

C.   Data Mining Task using Neural Network

There are two types of tasks:

- **Predictive:** allows user to enter records whose values are unknown. There are two models :

    - **Classification**: Recognizes patterns which describe the group of items which are predefined.

    - **Prediction**: Assess class and value of the object

- **Descriptive:** Provides knowledge of the data. There are two models in Descriptive:

    - **Clustering:** Recognizes patterns which describe items of no predefined groups.

    - **Association Rule:** Ensures completeness of mining using constraints and measures.

### D. Data Mining Process Based on Neural Network

Neural Network is one of the methods used in data mining to useful information from the data and generates error free result. Data mining process based on Neural Network is composed of three phases: Preparation of Data, Rules Extraction and Rules Assessment.

Figure 5. Data Mining Process based on Neural Network [3]

Preparation of Data: Specific data mining method must be applied for mining data. It is a necessary step in mining data based on Neural Network and plays an important role in the entire process. It consists of four sub – processes such as:

a) **Cleaning of data:** Removes noise from data and fills the vacant value of the data.

b) **Option of data:** Selected data is arranged in a row.

c) **Pre-processing of data:** Cleans the selected data collected in option of data for better improvement.

d) **Expression of data:** Repeated data must be accepted and clustered into the shape acceptable by data mining using Neural Network.
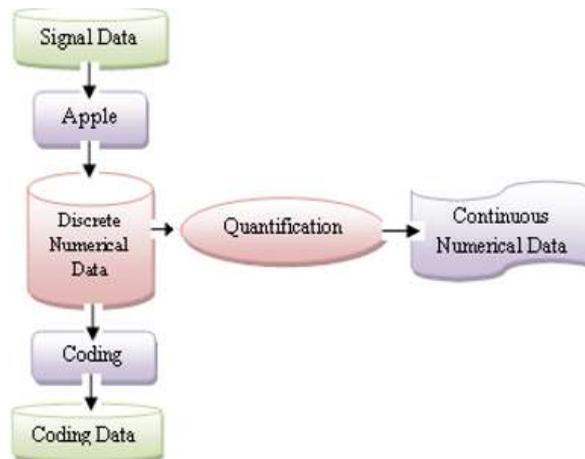
Figure 6. Data Expression process in data mining [3]

Rules Extraction: Various methods to extract rules are applied such as LRE method of fuzzy rule extraction, method for extracting rules from recursive network, black – box method. Binary input and output rules extracting formula (BIO - RE), full rules extracting formula (Full - RE), partial rules extracting formula (Partial - RE) [3].

Rules Assessment: Rules assessment is based on some objectives such as:

a) Finding the sequence which is suitable for extracting rules.

b) Testing the accuracy of rules.

c) Observing and detecting data not being extracted from the neural network.

d) Observing and detecting the inconsistency between the trained neural network and extracted rules.

II. LITERATURE REVIEW

Before the computers were invented, McCulloch and Walter Pitts in 1943studied human brain and proposed "computing element" model. Hebb in 1949 proposed a law of fundamental learning in neural network literature. In 1958 , Rosenblatt proposed the perceptron model. Windrows and Hoff proposed an ADALINE (adaptive linear element) model and LMS(least mean square) learning algorithm used to adjust the weights of an ADALINE model in 1960[6]. In 1969, Marvin Minsky and Seymour published the machine learning research paper. In 1975, Paul Werbos proposed a circuit which can process mathematical computation. In the 1990s, Neural Networks were become more popular in machine learning because of support vector machines and other. Renewed interest in the 2000s in Neural Networks was sparked because of the advent of deep learning [1].

III. PROPOSED WORK

In this paper, Neural Network techniques for data mining, Preparation of Data, Mining of Data, Expression and Interpretation of the results are presented. Method of parallel processing used in neural network for data mining and non-linear mapping predicts new outcomes from past trends. Classification, prediction and noise reduction provides accurate solution.

IV. CONCLUSION

For solving data mining problem, Neural Network is very useful and suitable due to good robustness, high degree of fault tolerance and distributed storage, self-organizing adaptive, parallel processing, Error detection and auto correction. The combination of Neural Network and Data Mining Method improves the efficiency of the process. We can use Neural Networks due to its capability of self error correction.

REFERENCES

[1] Nikita Jain "DATA MINING TECHNIQUES: A SURVEY PAPER", International Journal of Research in Engineering and Technology e-ISSN: 2319-1163, Nov 2013

[2] H Lu, R Setiono, H Liu. "Effective Data Mining Using Neural Network", IEEE Transactions on Knowledge and Data Engineering, 1996, 8(6): 957-961….

[3] Rasika P. Ghom, N. R. Chopde "Survey Paper on Data Mining Using Neural Network", International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064

[4] M.Charles Arockiaraj, "Applications of Neural Networks In Data Mining", International Journal Of Engineering And Science , Issn(e): 2278-4721, Issn(p):2319-6483

[5] Sonalkadu, Prof. Sheetal Dhande "Effective Data Mining Through Neural Network", ISSN: 2277 128X.

[6] Amrender Kumar "Artificial Neural Networks for Data Mining", I. A. S. R. I., Library Avenue, Pusa, New Delhi.