

Survey Paper on Data Mining Using Neural Network

Rasika P. Ghom¹, N. R. Chopde²

¹Student of Master of Engineering in (CSE), Rasoni College of Engineering and Technology, Amravati, India

²Professor, Head of Department (CSE), Rasoni College of Engineering and Technology, Amravati, India

Abstract: Data Mining means mine or sorted important data from huge amount of data. It is most useful in every field like business, engineering, web data etc. Classification of data is most difficult task in data mining which is used to solve with the help of different algorithms. Data mining include classification, clustering, rule generation and knowledge discovery. There are many methods used for data mining, including Neural Networks, Regression, and Decision Trees. Neural Networks are explaining the Neural Network Method in Data mining also data mining task using Neural Network which is helpful to classification of data. Also the Data Mining Process based on Neural Network including three main phases: Preparation of Data, Mining of Data, Expression and Interpretation of the Results. In this paper we study and research the data mining using Neural Networks.

Keywords: Data Mining, Neural Networks, Data Mining Process.

1. Introduction

The very fast and rapid development of database technology and the various applications of database management system, the large amount of data store in database increases very quickly in the world. And much important information is hidden in the large amounts of data. Data mining refers to extracting or mining the knowledge from large amount of data. The term data mining is named as 'Knowledge mining from data' or "Knowledge mining" [6]. If the information or knowledge can be selected from the database that they will created potential profit and useful value for the companies and the technology. Few data mining tools can used to resolve some traditional problems which take much time and find some useful information which is experts unnoticed. The Neural Network is a very common and useful technique to solve problems related to data mining [10]. Neural Network is a parallel processing network, which are gather together for tightly interconnected network purpose, based on some Biological Neural Network features. Biological Neural Network or human brain observed and study from its surrounding, and its past experiences. The parallel processing method and the Neural Network structure used to express the input output using associated knowledge. Neural Network has complex structure, long training time and poor interpretability, this is the main reasons that they defect in data mining. But the advantages of neural network are high affordability for the noise data and low error rate.

1.1 Neural Network

The Neural Network actually called as an artificial neural network. It is a biological model based on the Biological Neural Networks structure and functions. Data or information that passes through the network affects Artificial Neural Network structure. Because a Neural Network learns from its surrounding, and its past experiences and same wrong condition occurs in future, it will already give appropriate solution for that [4]. Neural Network is useful in various ways.

1.2 Data Mining Using Neural Network

Data mining is the method of exploration of large data sets, in order to discover meaningful pattern and rules. The aim of data mining is to design and work more efficiently with large data sets [9]. Data Mining is the process of analyzing data or proper information from the huge amount of data. In which several steps are involved which are shown below [3].

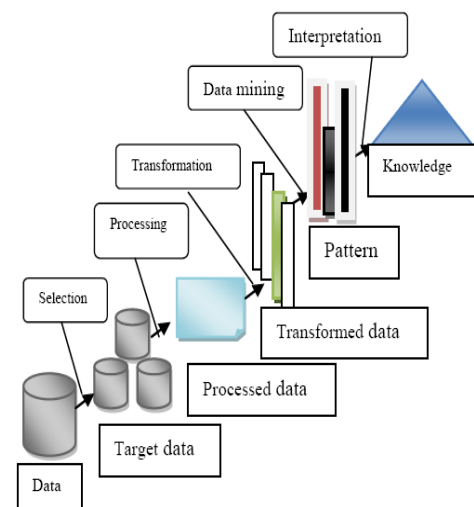


Figure 1: Data Processing

- 1) Cleaning of Data: In this step, we remove noise data from collected raw data.
- 2) Integration of Data: the combination of multiple data into only one single data and store it called as target data. This is happened in this step.
- 3) Selection of Data: In this step, from data base data retrieved for analyzing task as pre-processed data.
- 4) Transformation of Data: in this step, for mining purpose using aggregated and summarizing operation and data is transform into standard formats.
- 5) Mining of Data: Here, we apply various tools and smart techniques are applied for rules or extract data pattern.

- 6) Evaluation of Pattern: Here, used this step to identify tree patterns for representation of knowledge.
- 7) Representation of Knowledge: Here, to help the users for understand and interpretation of data mining and also their result we used the knowledge representation technique.

1.3 Data Mining Tasks Using Neural Network:

There are two tasks of Data mining using Neural Network.

1. Descriptive
2. Predictive Data Mining [2].

A. Descriptive data mining provides knowledge or data to understand inside the data, what is happening there without any predetermine idea.

B. Predictive data mining, in predictive data mining, users allows to submitting records which having field values which is unknown, and after that based on previous patterns system will able to predict the unknown values discovered form the database.

According to the tasks that the system performs Data mining models can be categorized:

- A. Classification
- B. Clustering
- C. Prediction
- D. Association Rules.

The predictive model include Classification and prediction but descriptive models include clustering and association rules

- a) **Classification:** classification is the most common action in data mining. It is used to recognizes patterns which describe the given item belongs to group By examining existing items it does this that already has been classified.
- b) **Clustering:** clustering is Similar to classification. The major difference between both is that no groups have been predefined.
- c) **Prediction:** Prediction is the construction associated which is used for class assess of an unlabeled object and value assess of a given object is likely to own.
- d) **Forecasting:** forecasting is the next application. It is different from predictions because inside the data based on patterns, it able to used the longer term price of continuous variables. Classifications, clusters, prediction and forecasting used for data mining,

1.4 Neural Network Method in Data Mining:

Here, we focus on Neural Network method in Data Mining. For classification, clustering, prediction and pattern recognition we used the concept of Neural Network [2]. The Neural Network model divided into the following three main types:

- a) **Feed-Forward Network:** It relates to the back-propagation model and the function network must be representation. It is used mostly in prediction and pattern recognition [1].
- b) **Feed-back Network:** It relates Hopfield discrete model and as representatives used in continuous model. For

associative memory and optimization calculation it is mostly used [2].

- c) **Self-Organization Network:** It relates to adaptive resonance theory (ART) model as representatives used Kohonen model, it is mostly used in cluster analysis [2].

1.5 Data Mining Process Based On Neural Network:

A Neural Network is an artificial representation of human brain that tries to simulate its learning process. An Artificial Neural Network is known as Neural Network. Neural network is one of the method used in data mining not only to extracted the useful data from huge amount of data but also to given accurate and efficient result without having any error in the resulted data [5]. Data Mining process using Neural Network can be composed by three main phases: Preparation of Data, Mining of Data, Expression and Interpretation of the Results [1].

The details are as shown in Fig. 2

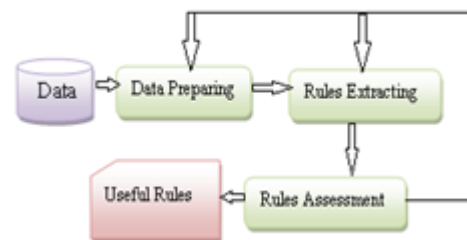


Figure 2: DM process based on neural network Approach

A. Preparation of Data

Preparation of data is a process which mining data for this data must be fit into specific data mining method. The first valuable or necessary step in the data mining using neural network is preparation of data. It plays an important role in the whole in its process. It mainly includes four processes as follows.

a) Cleaning of Data

Cleaning of Data is used to fill the data's vacancy value, and remove noise data from collected raw data.

b) Option of Data

Option of Data is to collect the selected data and arrange it properly. In data option mining, we used the row.

c) Preprocessing of Data

Preprocessing of Data is process to clean selected data to improvement.

d) Expression of Data

Expression of Data is to work repeated for data must be accepted when preprocessing into the shape which accepted by the data mining formula using Neural Network.

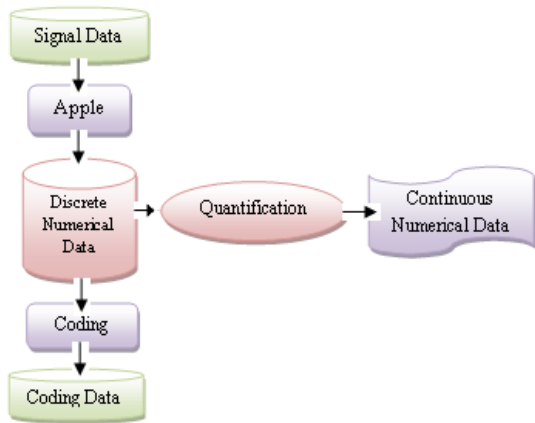


Figure 3: Data Expression and conversion in data mining based on Neural Network

B. Rules Extracting

There are various methods to extract rules, where most common used methods are the method of extracting fuzzy rules LRE method, the method of 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).

C. Rules Assessment

Rules assessment objective depends on almost every specific application, but actually rules can be assessed in generally in accordance with the some objectives [2].

- 1) Firstly find the sequence which is suitable for extracting rules. In the given data set, obtains the effective results.
- 2) Test the rules extraction accuracy.
- 3) Observe and detect how much data has not been extracted from the neural network.
- 4) Observe and detect the inconsistency between the trained neural network and extracted rules.

1.6 Learning Methods of Artificial Neural Network in Data Mining:

A **Neural Network** is configured in such a way that the given application of a inputs set are produces the set of desired outputs. No of methods are used to set the strengths of the connections that is exist. The priori knowledge is used to set the weights explicitly. One Another way by feeding teaching patterns to '**train**' the neural network and change its weights according to learning rule. There are various methods of learning are as follows:

- a) **Supervised learning** or Associative learning: In supervised learning, neural network is trained by providing input and as well as matching output patterns.
- b) **Unsupervised learning** or Self-organization: In Unsupervised learning an output unit is trained with the respond of clusters of pattern which is present within the input. Here system must develop its own input stimuli of representation. For classified the pattern, no set of priori

categories are used. Set of priori categories used in supervised learning.

2. Literature Review

In the recent development field Genesis of ANN model and its applications appear. However, before the advent of computer this field was established. McCulloch and Walter Pitts in 1943 the functions of a human brain studied, proposed a model of "computing element" called McCulloch and Pitts which present weighted sum of the inputs to the element used by threshold logic operation. Combination of some of these computing elements was used to realize several logical computations. Hebb in 1949 proposed a law becomes a fundamental learning in neural network literature. The perceptron model was proposed by Rosenblatt in 1958. Windrows and Hoff proposed for computing element an ADALINE (adaptive linear element) model and LMS(least mean square) learning algorithm is used to adjust the weights of an ADALINE model in 1960[11]. Neural Network research started after the publication of machine learning research paper by Marvin Minsky and Seymour in 1969. They discovered two key issues using the computational machines that processed Neural Networks. With mathematical notation, circuitry also described by Rosenblatt, such as the circuit, only after the back propagation algorithm a circuit whose mathematical computation can be processed, it was proposed by Paul Werbos in 1975 [12]. Under the name connectionism The parallel distributed processing of the mid-1980s became popular. The statement by David E. Rumelhart and James McClelland in 1986 provided a full exposition in computers to simulate neural depend on the use of connectionism network processes. In artificial intelligence we used the Neural Networks. 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. In CMOS, Computational devices have been created, to both neuromorphic computing and biophysical simulation. In the computational devices had big improvement from 2006 [8].

In the research group, current neural networks and deep feed forward neural networks developed by Jürgen Schmidhuber Between 2009 and 2012, at the Swiss AI Lab which have won in pattern recognition and machine learning eight international competitions [7].

3. Proposed Work

In this paper, the Data Mining using Neural Network technique based on the Preparation of Data, Mining of Data, Expression and Interpretation of the Results. The idea of the non-linear mapping used by the Data Mining using Neural Network, and the method of parallel processing used in neural network in data mining, can predict new outcomes from past trends. Also support to give appropriate solution based on past experiences and able to gain knowledge or data from huge amount of data. Especially Neural Network used in data mining for classification, prediction and noise reduction.

4. Conclusion

At present, for research data mining is important and new area. For solving data mining problem Neural Network itself is very useful and suitable, because some characteristics like good robustness, high degree of fault tolerance and distributed storage, self-organizing adaptive, parallel processing. Error detection and auto correction are also done using neural network. The combination of Neural Network and Data Mining Method, model can greatly improve the efficiency of the process. And it has been used over a large area. We can use Neural Networks because it has the capability of self error correction also they adjusting weights as required. It also receives more and more attention.

Reference

- [1] Sanjesh Ghore, Data Mining used of Neural Networks Approach, IJISSET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 6, August 2014
- [2] Rasika Ashokrao Dugane, Review On Data Mining using Neural Network, International Journal of Science and Research (IJSR), ISSN (Online): 2319-7064, Paper ID: 020131544, April 2014
- [3] Parvez Rahi et al Int. Journal of Engineering Research and Applications www.ijera.com ISSN: 2248-9622, Vol. 4, Issue 4(Version 7), April 2014, pp.36-42
- [4] D.Prashanth Kumar et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (6), 2014, 7041-7044
- [5] Priyanka Gaur, Neural Networks in Data Mining, International Journal of Electronics and Computer Science Engineering, ISSN- 2277-1956, 2014
- [6] Nikita Jain DATA MINING TECHNIQUES: A SURVEY PAPER, International Journal of Research in Engineering and Technology eISSN: 2319-1163, Nov-2013
- [7] Kurzweil AI Interview with Jürgen Schmidhuber on the eight competitions won by his Deep Learning team 2009–2012, 2012
- [8] Yang, J. J.; Pickett, M. D.; Li, X. M.; Ohlberg, D. A. A.; Stewart, D. R.; Williams, R. S. Nat. Nanotechnol. 2008, 3, 429–433
- [9] Fayyad, U., Piatetsky-Shapiro, G. & Smyth, P. (1996).From Data Mining to Knowledge Discovery in Databases. (AI Magazine, 17(3), 37-54.)
- [10] H Lu, R Setiono, H Liu., Effective Data Mining Using Neural Network. IEEE Transactions on Knowledge and Data Engineering, 1996, 8(6): 957-961....
- [11] Amrender Kumar, Artificial Neural Networks for Data Mining, I. A. S. R. I., Library Avrnue, Pusa, New Delhi
- [12] Werbos, P.J. (1975).Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences