Retail Forecasting using Neural Network and Data Mining Technique: A Review and Reflection

Archana Kumari¹, Umesh Prasad², Pradip Kumar Bala³

¹²Department of Computer Science & Engineering, BIT Ext Lalpur, Ranchi-834001, India
³Indian Institute of Management, Ranchi Ranchi-834001; India

Abstract: Retail Forecasting is a challenging problem in modern world. At the organizational level, forecasts of sales are essential inputs to many decision activities in various functional areas such as marketing, sales, and production/purchasing, as well as finance and accounting. Sales forecasts also provide basis for regional and national distribution and replenishment plans. The importance of accurate sales forecasts for efficient inventory management has long been recognized. A good forecasting model leads to improve the customer’s satisfaction, reduce cost, increase sales revenue and make production plan efficiently. In this paper, we focus on exploring the concept of soft computing and data mining techniques to solve retail problem. This paper describes data mining in the context of retail application from both technical and application perspective by comparing different data mining techniques. This paper also discuss soft computing techniques viz. neural network, genetic algorithm etc. in sales forecasting and inventory management.

Keywords: Retail Forecasting, Data Mining, Soft Computing, Neural Network.

1. INTRODUCTION

Data mining techniques have been used to uncover hidden patterns and predict future trends and behaviours in retail markets. The competitive advantages achieved by data mining include increased revenue, reduced cost and much improved marketplace responsiveness and awareness. Data mining has been applied to a number of retail applications, including development of trading models, investment selection, inventory control and so on.

Knowledge provides power in many retail context enabling and facilitating the preservation of valuable heritage, learning new things, solving intricate problems, creating core competencies and initiating new situations for both individuals and organizations now and in the future [1]. In most sectors’ retailing is extremely competitive and the financial margins that differentiate between success and failure are very tight, with most and established industries needing to compete, produce and sell at a global level. To master these trans-continental challenges, a company must achieve low cost production yet still maintain highly skilled, flexible and efficient workforces who are able to consistently design and produce high quality, low cost products. In higher-wage economies, this can generally only be done through very efficient exploitation of knowledge [2-3]. However knowledge can take many forms and it is necessary to identify the kind of knowledge to be mined when examining the huge amount of data generated during retail.

Understanding the underlying demand patterns for a particular product mean that outlets can be re-stocked in sufficient time to cope with changes in consumer demand. For products those are perishable or have a short shelf-life, this issue is more critical than the slower moving products with smaller demand requirements. The super market chains face additional problems in terms of number of stores, slow and erratic sales for many items at the store level, assortment instability, and promotional activity and price changes [4].

Modern demand forecasting systems provide new opportunities to improve retail performance. Although the art of deterministic demand estimation of individual merchant may never be replaced, it can be augmented by an efficient, objective and scientific approach to forecasting demand. Nevertheless it remains a difficult task. Point of sales automation through large-scale system may be an aid to handle the mass of retail data - organizing it, mining it and projecting it into future customer behavior. An ecommerce network facilitates this automated data collection, which is further facilitated by statistical analyzer.

2. DATA MINING TECHNIQUES

2.1 Classification and Issues of Data mining in Retail Application

The objective of data mining is to discover hidden knowledge, unknown patterns, and new rules from large databases that are potentially useful and ultimately understandable for making crucial decisions. Based on the type of knowledge that is mined, data mining can be mainly classified into the following categories
2.1.1. Association rule mining finds frequent patterns, associations, correlations, or causal structures among sets of items or objects in transaction databases, relational databases, and other information repositories. Applications are basket data analysis, cross-marketing, catalog design, etc. A typical example is market basket analysis, which analyzes purchasing habits of customers by finding association between different items in customers’ shopping baskets.

2.1.2. Classification and prediction is the process of classifying data (constructs a model) based on the training set and the values (class labels) in a classifying attribute and uses it in classifying new data. The models are used to predict the class of objects whose class label is unknown. It is a two step process. In the first step, a model is built. This step is also known as supervised learning. The learned model is represented in the form of classification rules, decision trees, or mathematical formulae. In the second step, based on classifier accuracy, the model is used for classification of the future data or test data. General techniques used for classification are decision tree induction, Bayesian classification, Bayesian belief network, and neural network. Other techniques such as K Nearest neighbour, case based reasoning, GA, Rough Set Theory, fuzzy logic and various hybrid methods are also used for classification. Typical Applications are credit approval, target marketing, medical diagnosis, treatment effectiveness analysis.

2.1.3. Clustering analysis segments a large set of data into subsets or clusters. Clustering is also known as unsupervised learning. Clustering maps a data item into one of several clusters, where clusters are natural grouping of data items based on similarity metrics or probability density models. Within the same cluster data objects are similar to one another and dissimilar to the objects in other clusters. Generally clustering techniques are classified in to following categories as: partitioning methods, hierarchical methods, density based methods, grid based methods, and model based methods. Liao and Wen reviewed the application of artificial neural network for clustering and classification. A detailed review and study of clustering techniques, application areas are mentioned in [7].

2.1.4. Sequential pattern and time series mining example clustering techniques can be used to identify stable dependencies for risk management and investment management looks for patterns where one event (or value) leads to another later event (or value). One example is that after the inflation rate increases, the stock market is likely to go down.

2.2 Existing Data mining Techniques

Data mining techniques (DMT) have formed a branch of applied artificial intelligence (AI), since the 1960s. During the intervening decades, important innovations in computer systems have led to the introduction of new technologies, for web-based education. Data mining allows a search, for valuable information, in large volumes of data. The explosive growth in databases has created a need to develop technologies that use information and knowledge intelligently. Therefore, DMT has become an increasingly important research area.

2.2.1 Neural Networks:
The term, neural network, is traditionally used to refer to a network, or circuit of biological neurons. Modern use of the term often refers to artificial neural networks, which are composed of artificial neurons, or nodes. As well as electrical signaling, other forms of signaling arise from neural transmitter diffusion, which have an effect on electrical signaling. As such, neural networks are extremely complex. Some applications for neural networks are radial basis function networks, neural classification, Bayesian confidence propagation neural networks, gene regulatory networks, fuzzy recurrent neural networks, neural nets, back-propagation artificial neural networks, Bayesian networks, general regression neural networks and flow networks.

2.2.2 Genetic Algorithm:
The basic idea genetic algorithm is that given a problem, the genetic pool of a specific population potentially contains the solution, or a better solution. Based on genetic and evolutionary principles, the genetic algorithm repeatedly modifies a population of artificial structures through the application of initialization, selection, crossover and mutation operators in order to obtain an evolved solution.

2.2.3 Statistical Inference:
Statistics provides a solid theoretical foundation for the problem of data analysis. Through hypothesis validation and/or exploratory data analysis statistical techniques give asymptotic results that can be used to describe the likelihood in large sample. The basic statistical exploratory methods include such techniques as examining distribution of variables, reviewing.

2.2.4 Rule induction:
Rule induction models belong to the logical, pattern distillation based approaches of data mining. Based on data sets, these techniques produce a set of if-then rules to represent significant patterns and create prediction models. Such models are fully transparent and provide complete explanations of their predictions. One commonly used and well-known type of rule induction is the family of algorithm that produces decision trees.

2.2.5 Data Visualization – "Seeing" the Data: Data are difficult to interpret due to its overwhelming size and complexity. In order to
achieve effective data mining, it is important to include people in the data exploration process and combine the flexibility, creativity, and general knowledge of people with the enormous storage capacity and computation power of today’s computers. Data visualization is the process of analyzing and converting data into graphics, thus taking advantage of human visual systems, large number of variables while still presenting useful information.

3. Back Propagation Artificial Neural Network

Soft computing was first proposed by Zadeh [9] to construct new generation computationally intelligent hybrid systems consisting of neural networks, fuzzy inference system, approximate reasoning and derivative free optimization techniques. It is well known that the intelligent systems, which can provide human like expertise such as domain knowledge, uncertain reasoning, and adaptation to a noisy and time varying environment, are important in tackling practical computing problems. In contrast with conventional artificial intelligence techniques which only deal with precision, certainty and rigor the guiding principle of hybrid systems is to exploit the tolerance for imprecision, uncertainty, low solution cost, robustness, partial truth to achieve tractability, and better rapport with reality.

Artificial neural networks consist of an inter-connection of a number of neurons that try to resemble the way the human brain works. Generally, the neural network is the multilayered network topology, including the input layer, hidden layer and output layer. Neural network is defined as massively parallel interconnected networks of simple elements (processing elements) and their hierarchical organizations which are intended to interact with the objects of real world in the same way as the biological neuron system does [10]. A properly modeled network has the potential to generate the acceptable results from the information what might be incomplete and non-specific. Their parallelism, speed and trainability makes neural network fault tolerant, as well as fast and efficient for handling large amount of data [11].

By training a network with a particular pattern of data, it subsequently recognizes the similar patterns with generalization [12]. The special features of learning and generalization makes neural network different from conventional algorithm. Neural network models are specified by the network topology and training procedure. In dozens of neural network models that were put forward, researches often use the Hopfield network, B P network.

**Hopfield** is the most typical feedback network model; it is one of the models which are most commonly studied now. The Hopfield network is the monolayer constituted by the same neuron, and is also a symmetrically connected associative network without learning function. It can implement the restriction optimization and associative memory BP network is the back-propagation network. It is a multi-layer forward network, learning by minimum mean square error. It is one of the most widely used networks. It can be used in the field of language integration, identification and adaptive control etc.

**Back propagation network** is semi supervised learning. First of all, artificial neural network needs to learn a certain learning criteria, and then it can work.

The Back-propagation algorithm [13] is probably the most widely used learning rule among the current available neural network systems. The back-propagation algorithm is a multi layer feed forward network with different transfer functions in the artificial neurons. The back-propagation is a kind of gradient descent technique with backward error propagation, which is the back-propagation rule, varies the connection weights in a manner that reduces the error as rapidly as possible. This is an interactive gradient algorithm designed to minimize the mean square error between the actual output and the desired output of a multilayer feed forward network. The network has been trained with the input vectors and their corresponding output vectors until it approximates a function, associate an input vectors with specific output vectors, or classify input vectors in an appropriate way as defined by the user. A properly trained back-propagation network tends to give reasonable results. Typically a new input leads to an output similar to the correct output for the input vectors used in training that are similar to the new input generated. This generalization property makes it possible to train a network on a representative set of input/target pairs and get good results without training the network on all possible input/output pairs.

The back-propagation neural network consists of one input layer, one output layer and one or more hidden layers. The number of neurons in the input and output layer are selected on the basis of the number of parameters affecting the process and the complexity of the relation existing between them. There is no rigid rule for calculating the number of hidden layers and the number of hidden neurons in each layer. Normally number of neurons or the processing elements in the hidden layer is selected taking into account the number of data points available for training the network and also the complexity of the relation existing between the input parameters and the output parameters.

4. CONCLUSION

This paper address the problem of retail forecasting using different data mining techniques. Even though different existing techniques can be used for retail forecasting but back propagation neural network has been found most appropriate technique. In future this tool may bring major benefits in this area.
References

AUTHOR

Archna Kumari received the Master in Computer Application (MCA) degrees from IGNOU in 2004 and pursuing M. Tech. in Computer Science from Birla Institute of Technology, Lalpur Ext Centre, Ranchi from 2011-14. Currently, she is working as faculty Member in Department of Computer Application at Marwari College, Ranchi. His research interest is in Data Mining and Soft Computing applications in IT and Business.

Umesh Prasad, received the Master in Computer Application (MCA) from Birla Institute of Technology, Mesra University, Ranchi in 2005 and the Ph.D.(Engg.) degree from Jadavpur University, Kolkata, in 2012. Currently, he is working as faculty Member in Department of Computer Science and Engineering at Birla Institute of Technology, Mesra University, Ranchi. His research interest is in computer applications to power systems and multi-objective optimization methods.

Pradip Kumar Bala, received his B. Tech. M. Tech. and Ph.D. from Indian Institute of Technology (IIT) Kharagpur. Currently, he is working as Associate Professor and Dean at Indian Institute of Management, Ranchi. He has published several papers in International Journal. His research interest is in Business Analytics and Operations Management.