International Journal of Engineering Research and Modern Education (IJERME)
ISSN (Online): 2455 - 4200
(www.rdmodernresearch.com) Volume I, Issue II, 2016

SHORT TERM PRICE FORECASTING OF A PRODUCT IN THE FIELD OF E-COMMERCE USING ARTIFICIAL NEURAL NETWORK APPROACH

A. M. Chandrashekhar*, H. P. Shivaraj**
& M. S. Jeevitesh***

*Assistant Professor, Department of Computer Science and Engineering, Sri Jayachamarajendra College of Engineering (SJCE), Mysore, Karnataka
** M.Tech Scholar, Department of Computer Science and Engineering, Sri Jayachamarajendra College of Engineering (SJCE), Mysore, Karnataka
*** M.Sc Biotechnology, Tech Lead, Unilog Content Solution, Mysore, Karnataka

Abstract:

E commerce is an evolving business and now there are around thousands of products in a single website. In each organization there are separate Business Units to monitor the price and variations of each and every product to maintain their revenue for the Organization keeping their profit per product. Based on the business needs product value may decrease and increase based on demand of the product in the market. Hence changing the product value based on the demand needed to be changed manually in the current scenario.This requires lots of human effort as there are many such products in the market which loses and gains demand often. Hence to overcome this scenario there needed an automated system which can predict the price based on the current demand which may increase their profit but never incur a loss. In this automated system the main aim is to monitor the price of each product based on their past sales, price variations and Demand. It can forecast the price of the product on regular intervals. In this paper we have used the artificial neural network concepts and machine learning for developing this automated system.

Index Terms: Corporate Managements, Forecasting Parameters, Forecasting Method & Artificial Neural network

1. Introduction:

Forecasting price is a challenging task and it has been fascinating the minds of people. Many different techniques have been proposed to solve the task of financial prediction. However, not many works have been dedicated to exploring the problem in relation to electronic environment, where the process of forming prices occurs dynamically. Examples of such environment include online shopping, online auction etc.

Many E Commerce Companies have turned to knowledge, processing and analyzing information in order to realize short-term or longterm forecasting based on powerful software applications.Information gained with these tools, by using time-series data and choosing the highest impact variables, can be processed with different technologies and will help detecting trends in the financial evolution of a company, a topic of high importance for top-level managers, helping them in taking the best decisions possible.

To forecast the Price of a product is mainly under the influence of important business environment factors and to make a new plan on the business following the prediction, the E Commerce company management includes three main financial ethics: profit and loss, balance sheet, cash flow. Obviously these are the main factors for those companies to grow economically and make mark in these competitive world, those factors should implement on each product they sell. And they cannot analyze those thousands of products 24X7 manually so they need an automation tool that look after

each & every product to fix the price in such a way, so they can achieve those three main ethics(profit and loss, balance sheet, cash flow).This is the main aim to do this research. This paper is devoted to forecast price of a product for a certain period of time. In traditional statistical and technical analysis methods, entrepreneurs tend to react to market irregularities easily. Neural network approach chooses to develop a number of predictive models which differ in the number of observations they consider and the methods of data transformation and normalization applied over inputs. To evaluate the performance of model accuracy measure like root mean square error is used.

2. Related Work:

The literature review gives few application of forecasting price using neural network. Each of them is carried out by a detailed study of their research paper. The summary of each of them is given below

Yusuf Perwej and Asif Perwej done on forecasting of US and Indian Currency exchange rate is to show the effect of number of input nodes and hidden nodes of NN in forecasting and to show large number of observations reduces forecast errors. They used the dataset used here is a daily currency exchange rate from beginning of 1989 to end of 2009. The detailed study of this research paper leads to the conclusion that numbers of input nodes have a greater impact on forecasting rather than number of hidden nodes. [7]Frank M. Tiesing and Oliver Vornberg said that neural network method of sales forecasting is far better than the traditional way of forecasting. The dataset used here Weekly observation on advertise spending, temporary price reduction, holiday, seasonal, late opening period for one full year. Over with four node layer with seven nodes each, one layer with four nodes each and one layer with one node is used for input layer, hidden layer and output layer respectively.

Back propagation algorithm is used from hidden, layer to input layer. Detailed study of this research paper leads to the conclusion that neural network approach outperforms the statistical approach. [8]Girish K. Jha and Kanchan Sinhamade a research that compare linear method or hybrid method performance better in forecasting agriculture price. The dataset used here is price of soyabean for 228 months (October 1991-September 2010) Price of mustard for 372 months (Jan 1980-Dec 2010) one layer with two nodes for ARIMA, one layer with two nodes for TDNN for input layer, one layer with three nodes for ARIMA, one layer with eight nodes for TDNN for hidden layer. The detailed study of this research paper leads to the conclusion that TDNN model provides better prediction accuracy in terms of RMSE and MAD than ARIMA model [9].

Marjan Niyati and Amir Masud wrote a research paper that introduces a new method of forecasting is proposed using Bayesian Analysis and neural network. The dataset used here is cost of raw material, cost of consumed raw material, Consumed material weight and cost of labor and equipment price per hour, Time needed for construction. One layer with seven nodes, one layer with three nodes and one layer with one node is used for input layer, hidden layer and output layer respectively. Detailed study of this research paper leads to the conclusion that the proposed method is precise, fast and without errors [10].

A. Architecture of Neural Network:

Neural network architecture consists of three layers: input layer, hidden layer and output layer. Each layer consists of one or more nodes. This is shown in the figure1. The line indicates the flow of information from one node to another node or more nodes.

Nodes at the input layer are passive i.e. they do not modify the data. They receive

single value as input and duplicates to multiple outputs.

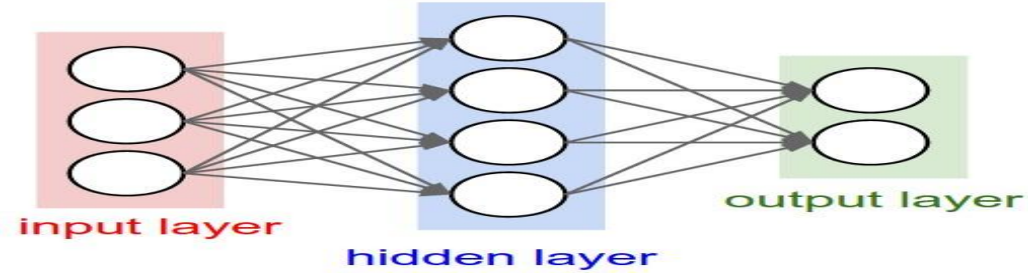

Figure1: Neural Network Architecture

The values entering hidden layer are multiplied by weights, a set of predetermined number stored in program. Before leaving the node this input is passed through a nonlinear mathematical function [11].

B. Back Propagation Algorithm:

The back-propagation algorithm is the most widely used method for calculating the error derivative of the weights, denoted by ΔE. ΔE can be computed as follows [12-15]:

$$\text{Error, } E = \tfrac{1}{2}\,\text{Err}^2 \approx \tfrac{1}{2}\,(X - Y)^2$$

Where, X= Actual Value and Y = Predicted Value

The squared error can be reduced by calculating partial derivative of error E with respect to every weight Wj,i as follows:

$$\Delta E = \frac{\partial E}{\partial W_{j,i}} = \frac{\partial}{\partial W_{j,i}}\left(\frac{1}{2}Err^2\right) = Err \times \frac{\partial Err}{\partial W_{j,i}}$$

$$= Err \times \frac{\partial}{\partial W_{j,i}}(X - Y)$$

[16-20]The process of training the three-layered network using back-propagation algorithm to reduce the errors at each layer by updating the weights is as follows:

- ✓ Compute the errors ΔE* at the output units which is simply the difference between desired output (X) and actual output (Y).
- ✓ Update the weights between the hidden layer and the output layer by using the errors ΔE*.
- ✓ Propagate the errors ΔE* back to the hidden layer to find the errors ΔE, - at the hidden layer units.
- ✓ Update the weights between the input layer and the hidden layer by using the errors ΔE, determined in step (iii).

The activation function/controls the amplitude of the output of the units and we used tanh function as activation function.

$$f(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1$$

C. Forecasting:

Forecasting is a tool used for predicting future demand based on past information. Forecasting can be used for: [3].Financial accounting: Financial accounting is concerned with summary, analysis and reporting of financial transactions pertaining to business.Production and operation: Production and operation management is concerned with transformation of production and operation inputs to outputs.

Key Features of Forecasting: [3]

- ✓ Demand for product and service are usually uncertain. Thus forecasting is also uncertain.

- ✓ Forecasting is more accurate for groups of items and for shorter time period.
- ✓ Forecast is no substitute for calculated errors.
- ✓ Every forecast value should include an error estimate.

D. E-Commerce:

Electronic commerce is the term given to a type of business which involves transaction of goods or service over the internet. E-commerce allows consumers to electronically exchange information over the network with no barrier or distance. Electronic commerce has expanded rapidly over the last decade. In the future the boundaries between the traditional and electronic commerce will increase rapidly as more entrepreneurs will move their transactions on the internet [4].

3. Design of the Proposed Model:

In our model, for determining the price used a back propagation neural network which contains three layers: input layer, hidden layer and output layer. The network we designed consists of five units in the input layer across three companies. The input layer also consists of one extra unit u0 as the bias unit. We set the value of a0 to the production cost of the product. Usually, sellers are not willing to sell their products below the production cost of the corresponding products. Hence, we considered the production cost of the product as the output of the bias unit

Our model can accept more attributes. One additional unit in the input layer needs to be added for each new attribute.On the other hand, in order to remove an attribute from the network the corresponding unit from the input layer, along with all the links that are connected to the unit, has to be eliminated. This implies that our model will work for any number of attributes [12].

A. Train Network:

Assumption is made that sellers use their historical data as the training patterns to the network. A training pattern consists of a set of inputs with desired output. We began our simulation by training the network of our model with 10 sets of training patterns so that errors can be minimized as much as possible by using back propagation algorithm. We trained our network for nine different numbers of epochs or iterations: 10, 50, 100, 500, 1000, 5000, 10000, 50000 and 100000.

As the training continues, after each iteration or epoch, the network calculates amount of error. The calculated error is then used to update the weights of the links by using back propagation algorithm so that error is minimized in the next iteration. Practically the value of error never becomes zero, but approaches to zero. We let our network to tolerate an error of amount 0.01 and 0.001. We run our network with five different learning rates: 0.01, 0.005, 0.001, 0.0005 and 0.0001. Analysis of the training process in the following sub-section indicates that the model performs better if we use 50000 epochs with 0.005 learning rate during training the network.

Dataset is taken from the e commerce companies like Savoy website for a product based and observed and collected the each and every variation for past few months according to the research parameters as shown in below Table 1.

Table 1: Dataset Prepared according to the variations of a Single Product

Product Price	Discount	Items In Stock	Sales	Selling Price
155	5	300	3	147.25
156	8	297	5	143.52
160	12	292	8	140.8
165	12	284	9	145.2
168	15	275	10	142.8

168	13	265	15	146.16
155	10	250	7	139.5

B. Algorithm:

The price of the product determined by the network (Fig. 1) can be found by using final output a9. The value of a9 can be calculated with the aid of equation (1) as follows:

$$\text{Final output } a9 = f(W6,9a6 + W7,9a7 + W8,9a8) \text{ . (2)}$$

$$\text{Where, } a6 = f(W0,6a0 + W1,6a1 + W2,6a2 + W3,6a3 + W4,6a4 + W5,6a5)$$

$$a7 = f(W0,7a0 + W1,7a1 + W2,7a2 + W3,7a3 + W4,7a4 + W5,7a5)$$

$$a8 = f(W0,8a0 + W1,8a1 + W2,8a2 + W3,8a3 + W4,8a4 + W5,8a5)$$

The phase is sub-dividing the process of price forecasting by our model of neural network into two phases: training phase and price determination phase. In the training phase we train our network with a set of training pattern. A training pattern consists of a set of inputs with desired output. A typical set of training pattern for our model each row represents a training pattern which contains a set of inputs with corresponding desired output. Initially assumptions that the buyers have equal preference on all the five attributes that are considered.

Therefore, link between input units and hidden units with equal weights. The purpose of the training process is to adjust the weights between the links such that the errors are minimized. To obtain this goal we feed units of the input layer of our network with the corresponding input values (Inputij) from each training pattern. We then determine the output from our network and compare it with the corresponding desired output (Output) of the training pattern to calculate error. Finally, then update weights between the links depending on the calculated errors. In our model, during the process of training, the errors between the links are minimized by using back-propagation technique.

The training process can be portrayed by the following steps:

- ✓ Input values from a training pattern to units of the input layer of the network.
- ✓ If the current training pattern is the first training pattern of the training set, then associate the links between input units and hidden units with equal weight. Also, associate the links between hidden units and output unit equally.
- ✓ Determine the value from the output layer.
- ✓ Compute the error, i.e., the difference between desired output of the training pattern and the value obtained in step iii.
- ✓ If the error is more than zero then go to step viii.
- ✓ If the error is approximately zero and there is more training pattern left, then take the next training pattern and go to step i.
- ✓ If the error is approximately zero and there is no more training pattern left, then terminate the training process.
- ✓ Update the weights of the links using back-propagation technique to minimize the error.
- ✓ Go to step iii.

The training phase updates weights between the links of the network as needed so that it can provide better output. Once the training process is complete, our model of network is ready to determine a competitive price for a product, P, from the price determination phase as follows:

- ✓ Set the production cost of the product, P as the input to bias unit of the input layer and set the weights of the links associated with bias unit to 1.

- ✓ Set the values (ai) of the input units for the corresponding purchase attributes of product by using prior knowledge about the prices of product offered by other competing sellers.
- ✓ Run the network and derive the price from the output layer.
- ✓ Set the price from the output layer as the product price.[13]

4. Results:

An ANN with back-propagation algorithm was trained and the training epoch (cycles) set for each network is 10,000. The purpose of the training is to minimize the mean squared error (MSE) and RMSD represents the sample standard deviation of the differences between predicted values and observed values. The RMS of the pairwise differences of the two data sets can serve as a measure how far on average the error is from 0 as shown in below table 2.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2}$$

Where, n = number of observed values, yj = input values, yj^ = predicted output values

We got RMSE as 0. 04323860 which is almost nearer to 100% accurate &the proposed model has a neural network structure and actual and predicted graphs shown in the following Fig 3-5.

Table 2: The forecasting results for the curretn dataset

Trial No.	No. of Hidden Layers	Threshold	Learning Rate	Algorithm	Activation Function	RMSE
1	2,3	0.05	0.01	Rprop-	Tanh	0.0654466
2	2,3	0.05	0.01	Rprop+	logistic	0.1152260
3	3,2	0.05	0.005	Backprop	Tanh	0.1132199
4	3,2	0.04	0.01	Backprop	Tanh	0.1058355
5	3,2	0.05	0.01	Rprop-	Tanh	0.0710285
6	2,3	0.04	0.01	Rprop-	Tanh	0.04323860
7	2,3	0.04	0.01	Backprop	Tanh	0.04623860
8	1,2	0.18	0.015	Rprop-	Tanh	0.1826710
9	1,2,3	0.2	0.01	Rprop-	logistic	0.08209940
10	1,3,2	0.22	0.025	Backprop	Tanh	0.19756710

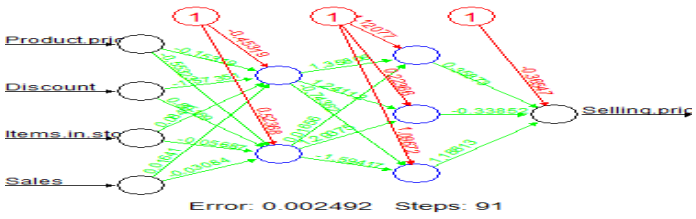

Figure 3: Generated neural network structure

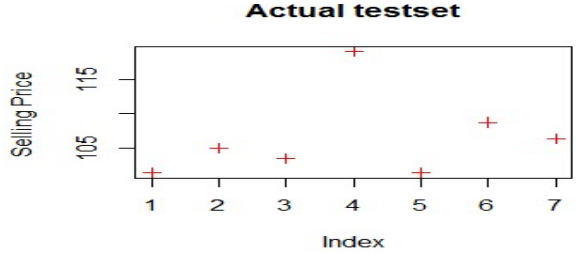

Figure 4: Actual Test Set

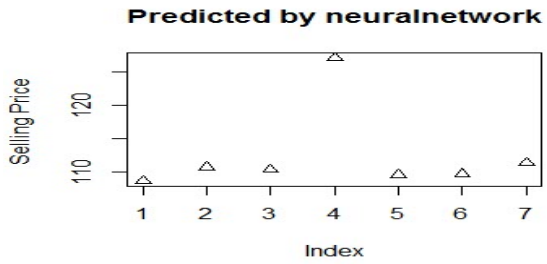

Figure 5: Predicted Test Set

As we can see that actual and predicted are almost similar hence we trained our model well that can forecast the price very accurately.

The actual forecasted price for the next three successive terms has shown in the figure-6.

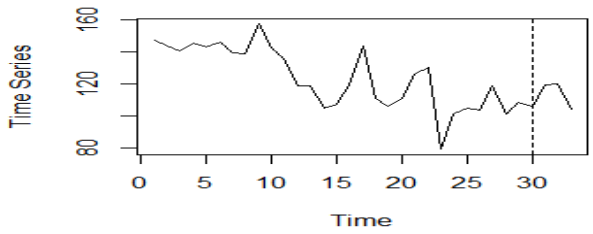

Figure 6: Price forecasted for a next 3 terms

5. Conclusion and Future Work:

Configuring prices dynamically is one of the most challenging and interesting research problems. With the development of ecommerce, it has become a vital task. This paper investigates the problem of winning product price over the dynamic and uncertain ecommerce environment. This paper gives a detailed picture of neural network concepts, ecommerce and forecasting. Different data transformation and normalization for preparing neutral network input is investigated in predictive model. The best performance is achieved by threshold activation function, reverse propagation algorithm among all others.

This project work leads to the conclusion that neural network approach of forecasting price is one of the premium solutions to handle the dynamic changes of the product price on all ecommerce business scenarios. As a part of future work we can aim to use different neural network architecture and different learning rates & combine those generated models to earn more accuracy to our pricing algorithm.

6. References:

1. Simon Haykin, McMaster University, Hamliton, Canada, Neural Networks-A Comprehensive Foundation
2. Arnold F. Shapiro Penn State University, Artifical Neural Network, Penn State Smeal College of Business.
3. G. Peter Zhang and Douglas M. Kline – Quarterly time series forecasting using neural network.
4. http://www.tutorialspoint.com/e_commerce/e_com merce_tutorial.pdf
5. https://www.rstudio.com/about
6. https://cran.rstudio.com/web/packages/neuralnet/ne uralnet.pdf
7. Yusuf Perwej and Asif Perwej – Forecasting of Indian Rupee (INR)/ US Dollar (USD) currency exchange rate using artificial neural network. International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol.2, No.2, April 2012.

8. Frank M. Tiesing and Oliver Vornberg – Forecasting sales using neural network. Proceedings ICNN `97 houston, Texas, 9-12 June 1997, vol.4 pp.2125-2128, IEEE, 1997.
9. Girish K. Jha and Kanchan Sinha – Agriculture price forecasting using neural network model: An innovative information delivery system. Agricultural Economics Research Review, Vol. 26 (No.2) July-December 2013 pp 229-239
10. MarjanNiyati and Amir Masud – Estimation of products final price using Bayesian analysis generalized Poisson model and artificial neural network.Journal of IndustrialEngineering, 2, (2009), 55-60, Received 3 Nov., 2008; revised 20 Dec., 2008; accepted 12 Feb. 2009
11. A. M. Chandrashekhar and K. Raghuveer - An Effective Technique for Intrusion Detection Using Neuro-Fuzzy and Radial SVM Classifier. Computer networks and communications (NETCOM). Part VI. pp 499-507. 2013, 10.1007/978-1-4614-6154-8_49, 978-1-4614-6153-1.
12. A. M. Chandrashekhar and K. Raghuveer - Fortification of hybrid intrusion detection system using variants of neural networks and support vector machines, International Journal of Network Security & Its Applications (IJNSA), Vol.5, No.1, January 2013. DOI: 10.5121/ijnsa.2013.5106.
13. A. M. Chandrashekharand K. Raghuveer - Amalgamation of K-means Clustering Algorithm with Standard MLP and SVM Based Neural Networks to Implement Network Intrusion Detection System. Advanced Computing, Networking and Informatics-volume 2. pp 273-283. 2014, 10.1007/978-3-319-07350-7_31, 978-3-319-07349-1, 2190-3018.
14. A. M. Chandrashekhar and K. Raghuveer -Performance evaluation of data clustering techniques using KDD Cup-99 Intrusion detection data set.International Journal of Information & Network Security (IJINS) Vol.1, No.4, October 2012, pp. 294~305 ISSN: 2089-3299
15. A. M. Chandrashekhar and K. Raghuveer – Fusion of multiple data mining techniques for effective network intrusion detection- A contemporary approach, Fifth International Conference on Security of Information and Networks(SIN2012) ,178-182 , 978-1-4503-1668-2.
16. A. M. Chandrashekhar and K. Raghuveer – Diverse and Conglomerate Modi Operandi for Anomaly Intrusion Detection Systems, International Journal of Computer Application (IJCA), ISSN: 0975 – 8887, IJCA Special Issue on "Network Security and Cryptography (NSC)" Number5, 2011, 10.5120/4353-054.
17. A. M. Chandrashekhar and K. Raghuveer –Intrusion Detection Technique by using K-means, Fuzzy Neural Network and SVM classifiers, 2013 IEEE International Conference on Computer Communication and Informatics (ICCCI - 2013), IEEE Xplore Digital Libirery. Catalog Number: CFP1308R-ART, ISBN Number: 978-1-4673-2907-1.
18. A. M. Chandrashekhar and K. Raghuveer – Confederation of FCM Clustering, ANN and SVM Techniques of Data mining to Implement Hybrid NIDS Using Corrected KDD Cup Dataset, 2014 International Conference on Communication and Signal Processing (ICCSP), IEEE xplore.
19. A. M. Chandrashekhar and K. Raghuveer – Hard Clustering Vs. Soft Clustering: A Close Contest for Attaining Supremacy in Hybrid NIDS Development, International Conference on Communication and Computing (ICC - 2014), Elsevier science and Technology Publications ISBN: 978-9-3510-7244-7, Page no 497-507.

20. A. M. Chandrashekhar and K. Raghuveer – Impact of Data Normalization on Machine learning Techniques Used in intrusion detection system development, Conference Proceedings of National Conference on Computing and Communication Engineering (NCCCE-14). Agni College of Technology, Chennai, India.
21. A. M. Chandrashekhar and K. Raghuveer – Improvising Intrusion detection precision of ANN based NIDS by incorporating various data Normalization Techniques – A Performance Appraisal. International Journal of Research in Engineering & Advanced Technology (IJREAT), Volume 2, Issue 2, Apr-May, 2014, ISSN: 2320 – 8791.