

# Price forecasting in e-commerce using neural network approach

**Nagashree.S. ; Dr. M.N.Jayaram &Mr. Jeevitesh M.S.**

Student, Final semester M.tech in Network and Internet Engineering Department of Electronics and Communication Sri Jayachamarajendra College of Engineering Mysore, India

<sup>1</sup>nagashree4767@gmail.com

<sup>2</sup>Associate Professor Department of Electronics and communication Sri Jayachamarajendra College of engineering Mysore, India <sup>2</sup>melkotejayaram@yahoo.com

<sup>3</sup>Tech Lead, Unilog Content Solution, Mysore, India  
jeevitesh.ms@gmail.com

**Abstract**— *The advancement of internet and better communication technology has led to the development of e-commerce. E-commerce is buying and selling of products or services over the computer network. With the incorporation of e-commerce in daily life, enterprises no longer rely on traditional business tactics. They have to cope up with dynamic and uncertain electronic environment, especially when developing cost planning. In such a situation prices are calculated dynamically through an aggressive bidding process depending on the market position, competitor policy and customer preferences. The capability to predict the price of any product is a crucial task. This paper introduces neural network approach of predicting future value of a product. Basics of neural network, different activation function and different network architecture are also explored. Training with different datasets taken from different online shopping website show the potential of proposed algorithm for forecasting price in the competing and uncertain environment.*

**Keywords**—neural network; price forecasting; styling; e-commerce

## I. INTRODUCTION

Forecasting price is a challenging task and it has been captivating the minds of many business professionals. 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.

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.

This paper is organized as follows: Firstly background work is provided. Secondly the proposed model for forecasting price is described followed by experimental results. This paper closes with conclusion and a discussion of future works. [9]

## II. BACKGROUND WORK

The background work of this paper gives a big picture of basics of neural network and price forecasting.

### A. Neural Network

A neural network is an automaton i.e. created to model the way in which human brain performs a desired task or function. The network is usually implemented by using electronic widgets or is simulated in software on a computer. To obtain good performance, neural network incorporates large interconnection of simple computing cells referred to as neurons or processing units. [1]

In the book, Neural network and learning machine, author Simon Haykin defines neural network as follows,

“A neural network is a massive parallel distributed processors made up of a simple processing units which has a natural propensity for storing experimental knowledge and making it available for use.”

It resembles human brain with respect two concepts:

- Knowledge is conquered by the network from its surroundings through leaning process.
- The acquired knowledge is stored in interneuron connection strength known as synaptic weight.

Neural network obtains its computing capability first from its large distributed structure. Secondly by its ability to learn through generalization. Generalization refers to the neural network producing outputs to inputs.

**Structure of neural network architecture:** Neural network architecture consists of input layer, hidden layer and an 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 or more nodes. [2]  
 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.

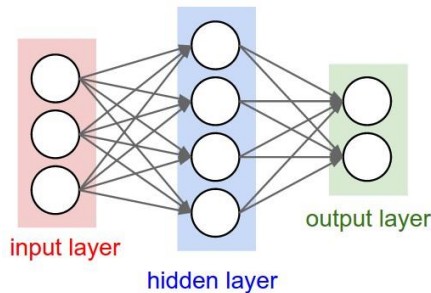


Fig 1:- Neural network architecture.

The values entering hidden layer are multiplied by weights, a set of predetermined number stored in program. The weighted inputs are then added to produce a single number. Before leaving the node this input is passed through a non linear mathematical function called sigmoid.

**Structure of a neuron:** A neuron is the structural and functional unit of artificial neural network. Figure 2 shows the structure of a neuron. There are four basic elements of a neuron.

- A set of synapse or connection links. Each as its own characteristics. Each signal is multiplied by weight  $w$ .
- An adder, which sums all the input signal.
- An activation function is used for normalizing the neuron. The output range of the neuron should be  $[0,1]$  or  $[-1,1]$ .
- A bias function which normalizes the net input of summation value of input signals. [1]

In mathematical term, a neuron  $k$  is described by the following pair of equation.

$$U_k = \sum_{j=0}^n W_{kj}x_j \quad (1)$$

and

$$y_k = \Phi(u_k + b_k) \quad (2)$$

$y_k$  is the output of a neuron.  
 $W_{kj}$  is the weight of the neuron  $k$  with input signal  $x_j$ .  
 $b_k$  is the bias function.

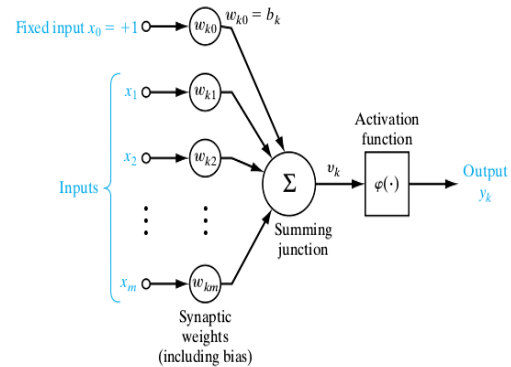


Fig 2: Structure of a neuron

**Analogy with human nervous system:** The human nervous system is viewed as a three stage system as shown in Figure 3:

- Central to the nervous system is the brain which continuously receives information. This is similar to the adder in Figure 2.
- There are two sets of arrows in the figure 3. The arrow pointing from left to right indicate transmission of signals to the brain. The arrow pointing from right to left indicate feed back in the system. This is similar to the synapse or connection links in Figure 2.
- The receptors convert stimulus from other nerve cell or from the surroundings to electrical impulse and sends to brain. This is similar to the arrows entering the neurons in Figure 2.
- The effectors convert the electrical signals to response. This is the output  $y_k$  in Figure 2.

**Activation function:** Activation function is the time varying component of neural network. It defines the output of a neuron in terms a local field  $v$ . It is denoted by  $\Phi(v)$ . Three basic types of activation functions have been identified. [2]

1. Threshold based activation function is defined by equation (3)

$$\Phi(v) = \begin{cases} 1 & \text{if } v \geq 0 \\ 0 & \text{if } v < 0 \end{cases} \quad (3)$$

2. Piecewise activation function is defined by equation. (4)

$$\Phi(v) = \begin{cases} 1 & \text{if } v \geq +1/2 \\ v & \text{if } v + \frac{1}{2} > v > -1/2 \\ 0 & \text{if } v \leq -1/2 \end{cases} \quad (4)$$

3. Sigmoid activation function is by far the most common type of activation function used in artificial neural network. The graph is s shaped. It is defined by equation (5)

$$\Phi(v) = \begin{cases} 1 & \text{if } v > 0 \\ 0 & \text{if } v = 0 \\ -1 & \text{if } v < 0 \end{cases} \quad (5)$$

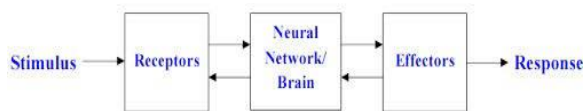


Fig 3:- Block representation of nervous system.

**Neural network architecture:** ANN can be seen as a directed graph where neurons are represented by nodes and weights are represented by edges. Based on connection pattern ANN can be grouped into two categories. [2]

- **Feed forward network:** These networks do not have loops. Feed forward network are static in the sense that they produce only one set of output values rather than a sequence of values from a given input. Feed forward networks are memory-less i.e. the response to an input is independent of the previous network state.
- **Feed back or recurrent network:** These networks contain loops. Feedback network are dynamic in nature i.e. when a new input pattern is presented a new output is calculated.

Figure 4 shows a feed forward and a feedback networks.

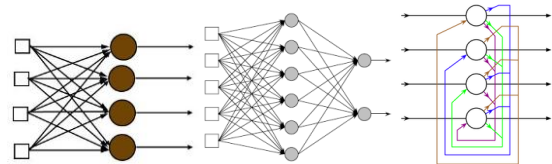
**Learning rate:** The capability to learn is the most basic quality of intelligence. A learning process is applied to neural network to update network architecture and to connecting weights so that network can work in full potential. Few learning rates are: [1]

- **Error correction rule:** During the learning process, the actual output  $y$  generated by the network may not equal the desired output  $d$ . The basic principle of error-correction learning rules is to use the error signal  $(d - y)$  to modify the connection weights to gradually reduce this error.
- **Boltzmann rule:** Boltzmann rule is used for symmetric feedback networks. Here the weights from unit  $i$  to  $j$  is equal to unit  $j$  to  $i$ .

**B. Forecasting:**

Forecasting is a tool used for predicting future demand based on past information. Forecasting can be used for: [3]

- **Strategic planning:** Strategic planning is organization's process of defining rules or decision and making decision on allocating resources to pursue its goals.
- **Financial accounting:** Financial accounting is concerned with summary, analysis and reporting of financial transactions pertaining to business.



(a) Single layer feed forward (b) Multilayer feed forward (c) feed back

Fig 4:- Different types of neural network architecture

- **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.

**C. 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]

As a part of background work, few application of forecasting price using neural network are investigated. Each of them is carried out by a detailed study of their research paper. The summary of each of the research paper is given below.

1. Forecasting of INR/USD exchange rate using neural network. [7]

The objective of this paper is to show the effect of number of input nodes and hidden nodes of neural network on forecasting and to show large number of observations reduces forecast errors. The dataset used here is a daily currency exchange rate from beginning of 1989 to end of 2009. They have chosen one layer with ten nodes each for input layer, one layer with five nodes for hidden layer and one layer with one node for output layer. The detailed study of this research paper leads to the conclusion that number of input nodes has a greater impact on forecasting rather than number of hidden nodes.

2. Forecasting sales using neural network.[8]

The objective of this paper is to show that neural network method of sales forecasting is far better than the traditional way of forecasting. The dataset used here is weekly observation on advertise spending, temporary price reduction, holiday, seasonal, late opening period for one full year. One 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.

3. Agriculture price forecasting: Mustard and soyabean crops in India. Hybrid model= Linear (time series-ARIMA\* algorithm) + Non linear (neural network-TDNN\*\* algorithm)[9]

The objective of this paper is to compare linear method and hybrid method performance 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 root mean square error and mean absolute deviation than ARIMA model.

4. Estimation of product final price using Bayesian Analysis and neural network.[10]

This research paper introduces a new method of forecasting called Bayesian Analysis. The objective of this paper is predicting future price of a product using Bayesian analysis and neural network methods. 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.

### III. PROPOSED METHOD

This work of forecasting price is carried out by a statistical tool called R studio. R studio is open source and commercial IDE for R programming language.

R Studio is powerful software used for data manipulation, calculation and graphical display. R Language comes with a large library of predefined function that can be used to perform high value data analysis.[5]

R is a simple and effective programming language which contains different data types, condition loops, user defined functions, recursive function etc. It also includes operators for calculation on array and matrix. It provides effective data handling and storage facility.[5]

The parameters of the current data set used are advertisement spending, promotional expense and quarterly sales. Table 1 shows the current data set.

\*ARIMA – Auto Regression Integrated Moving Average.

\*\*TDNN – Time Delay Neural Network.

Date	advertisin	promotioi	quarterly	predict	furure
12-Feb	494	740	4350	1679	
12-Mar	536	790	2020	2245	
12-Apr	614	928	4764	2590	
12-May	665	991	3456	2559	
12-Jun	811	819	3658	3439	
12-Jul	844	871	5220	3867	
12-Aug	1047	945	7678	3460	
12-Sep	1050	919	4210	3527	
12-Oct	1153	1014	7890	3913	
12-Nov	1271	1040	7328	4659	
12-Dec	1374	1080	6187	5702	
13-Jan	1616	1155	8526	5972	
13-Feb	1603	1109	2858	6280	
13-Mar	1611	1097	2692	6190	
13-Apr	1886	1211	9445	6596	
13-May	1927	1131	17488	7286	
13-Jun	2153	1270	9982	7196	
13-Jul	2078	1241	4015	7491	
13-Aug	2191	1326	6549	8103	
13-Sep	2087	1317	3763	8668	
13-Oct	2150	1321	1942	9708	

Table 1: Current data set

No.	Price	Customer rating	Delivery time-min	Delivery time-max	Seller ratings
1	135	0	216	576	99
2	73.9	5	72	120	96
3	119.9	5	72	120	99
4	217.9	5	72	240	0
5	105.9	53	0	12	99
6	185.9	55	0	11	100
7	130.9	895	0	9	100
8	115.9	755	0	8	99
9	95.9	51	0	144	99.6
	339.95	815	24	48	100
11	339.9	725	24	48	93
12	209.95	555	48	96	99
13	209.95	865	336	384	100
14	239.95	25	72	120	99
15	239.95	865	0	11.49	99
16	239.95	715	24	48	100
17	119.95	715	72	120	99
18	97.32	715	24	72	100
19	409.3	565	24	48	98
20	178.5	735	0	13.22	99
21	230	0	192	240	93
22	112.3	0	0	10.32	92
23	230	0	48	120	100
24	112.3	505	72	120	100
25	354.9	65	48	120	100

Table 2: Live dataset

The live data set is taken from different online shopping website and same is converted to numerical values. Live data set is used to train the algorithm, thus it shows potential of the algorithm. The live data set parameters includes Price of the product, Customer rating, Delivery time and seller rating. Table 2 shows the live data set.

The algorithm uses neuralnet R package to build the code. The following function describes the forecasting of price.

```
neuralnet (formula, data, hidden = 1, threshold = 0.01,
stepmax = 1e+05, rep = 1, start weights = NULL, learning
rate. Limit = NULL, learning rate. Factor = list (minus = 0.5,
plus = 1.2), learning rate=NULL, life sign = "none",
lifesign.step = 1000, algorithm = "rprop+", err.fct = "sse",
act.fct = "logistic", linear. Output = TRUE, exclude = NULL,
constant. Weights = NULL, likelihood = FALSE) [6]
```

With many trails and error method the following arguments for the neuralnet function is used.

- Formula: a symbolic description of the model to be fitted.
- Data: a data frame containing the variables specified in formula.
- Hidden: a vector of integers specifying the number of hidden neuron (vertices) in each layer
- Threshold: a numeric value specifying the threshold for the partial derivatives of the error function as stopping criteria.
- Rep: the number of repetitions for the neural network's training.
- StartWeights: a vector containing starting values for the weights. The weights will not be randomly initialized.
- Learning rate: numeric value specifying the learning rate used by traditional back propagation.
- Algorithm: a string containing the algorithm type to calculate the neural network. The following types are possible: 'backprop', 'rprop+', 'rprop-', 'sag', or 'slr'. 'backprop' refers to backpropagation, 'rprop+' and 'rprop-' refer to the resilient backpropagation with and without weight backtracking, while 'sag' and 'slr' induce the usage of the modified globally convergent algorithm (grprop).
- Err.fct: a differentiable function that is used for the calculation of the error. Alternatively, the strings 'sse' and 'ce' which stand for the sum of squared errors and the cross-entropy can be used.

#### IV. EXPERIMENTAL RESULTS

The proposed model is used for price forecasting using neural network approach. The output for the code shows price forecasting for the next 20 months.

Live data set is used to train the algorithm. The live dataset is taken from few online shopping websites which sells house lightings and fans. Different types of lights and fans are chosen from minkragroups.net – a leading decorative lighting company in America. The same products are searched in few other websites like amazon.com, ebay.com, lamps.com, houzz.com, 1800lightings.com, southernlightsinc.com, creative-lighting, homedepot.com, wayfair.com, ceilingfans.com, delmarfans.com, lightology.com, and hanswholesales.com. The corresponding price of the product taken from these websites is named as competitor price in the live dataset. Competitor price is one of the parameter in live dataset. The other parameters include customer ratings, delivery time and seller ratings. These parameters are also taken from the same online websites which are listed above. Each parameter is then converted to numerical representations and the fed to the code as input. Training the

code with live data set gives accuracy of 0.28. The table 1 shows the accuracy of current and live data set.

Accuracy is the measure of closeness between the input and output values. Accuracy measure used here is root mean square error. The root mean square error is defined by equation (6)

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2} \quad (6)$$

Where,

n = number of observed values,

y<sub>j</sub> = input values

y<sub>j</sub><sup>^</sup> = predicted output values

Table 3 shows the accuracy of the algorithm with live and current data set.

Dataset	Accuracy
Current data set	0.28
Live data set	0.4

Table 3: accuracy of algorithm with current and live dataset

#### V. 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 neural 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 best 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.

#### REFERENCES

- [1] Simon Haykin, McMaster University, Hamilton, Canada, Neural Networks-A Comprehensive Foundation
- [2] Arnold F. Shapiro Penn State University, Artificial Neural Network
- [3] G. Peter Zhang and Douglas M. Kline – Quarterly time series forecasting using neural network.
- [4] [http://www.tutorialspoint.com/e\\_commerce/e\\_commerce\\_tutorial.pdf](http://www.tutorialspoint.com/e_commerce/e_commerce_tutorial.pdf)
- [5] <https://www.rstudio.com/about/>
- [6] <https://cran.rstudio.com/web/packages/neuralnet/neuralnet.pdf>
- [7] Yusuf Perwej and Asif Perwej – Forecasting of Indian Rupee (INR)/ US Dollar (USD) currency exchange rate using artificial neural network.
- [8] Frank M. Tiesing and Oliver Vornberg – Forecasting sales using neural network.
- [9] Girish K. Jha and Kanchan Sinha – Agriculture price forecasting using neural network model: An innovative information delivery system.
- [10] Marjan Niyati and Amir Masud – Estimation of products final price using Bayesian analysis generalized Poisson model and artificial neural network.