Predictive Analytics Using Soft Computing: A Case Study on Forecasting For Indian Automobile Industry

Jewel Murel D'Souza[1], Sheik Sana Begam[2], Santhosh Rebello[3]

Department of Information Technology, AIMIT, Mangalore, India[1,2,3]

ABSTRACT: Predictive analytics is the branch of the advanced analytics, which encompasses a variety of statistical techniques from modeling, machine learning, statistics, and data mining that analyze current and historical facts to make predictions about unknown future events. Soft Computing will fit the predictive analytics using neural networks in a very efficient way by replacing all the other methods and produce forecasts as accurate as or better than those available from other statistical methods. Indian automobile industry has gone tremendous changein terms of sales, customer expectation and challenges.A quick decision making and a positive movetowards change is required as the market scenario can be highly dynamic and can cause change without giving much time to respond.This paper aims at providing a proposed neural network solution that can be used to predict the future sales on the base of historical data using Back Propagation algorithm. Especially we aim at finding which automobile company will be on a high scale of sales as per the customer's capacity and desire. This forecasting helps firms to take strategic decisions to achieve their goals.

KEYWORDS: Predictive analytics, SoftComputing, forecasting, Back propagation.

I. INTRODUCTION

Predictive analytics encompasses a variety of statistical techniques from modeling, machinelearning, and data mining that analyze current and historical facts in order to determine patterns and predict about future outcome and trends. Predictive analytics does not tell you what will happen in the future. It forecasts what might happen in the future with an acceptable level of reliability, in order to better understand customers, products and partners and to identify potential risks and opportunities for a company. Often the unknown event of interest is in the future, but predictive analytics can be applied to any type of unknown whether it be in the past, present or future.

Soft Computing is a mechanism by which we give a feasible solution to a problem. It is tolerant of imprecision, uncertainty, partial truth, and approximation. In effect, the role model for soft computing is the human mind. Soft computing is used when we don't have enough information about the problem itself. These kinds of problems originate in the human mind with all its doubts, subjectivity and emotions. Soft Computing is a term used in computer science to refer to problems whose solutions are unpredictable, uncertain and between 0 and 1. Neural Networks, including neural computing is one of the principle constituents of soft computing. Neuralnetworks are nonlinear sophisticated modeling techniques that are applied to problems of prediction, classification.

In Neural networks the exact nature of the relationship between inputs and output is not known. A key feature of neural networks is that they learn the relationship between inputs and output through training. There are three types of training in neural networks used by different networks, supervised and unsupervised training, reinforcement learning, with supervised being the most common one. Some examples of neural network training techniques are back propagation, quick propagation.

II. AUTOMOBILE SALES IN INDIA AND PROBLEMS IN SALES FORECASTING

The automobile industry is one of the core and vibrant industries of the Indian economy. The automobile industry is the largest industry in the world. Some call it the 'industry of industries'. With the huge number of new automobiles rolling out each year, on Indian roads, the industry is set to grow on a very high scale. Since it has the most leading product in the manufacturing sector, it can be said as the driver of the economic growth. The silent entry of the automobile industry in India was made in the nineteenth century. The industry accounts for 22 per cent of the country's manufacturing gross domestic product (GDP) which comprises passenger cars, two-wheelers, three-wheelers and commercial vehicles and is currently the seventh-largest in the world with an average annual production of 17.5 million vehicles, of which 2.3 million are exported. The Indian car industry has undergone tremendous change and has witnessed a phenomenal growth in recent times in terms of innovative designs, concepts and technology which therefore has attracted many global automotive players.

Increase in the paying capacity of the consumers which has led to an increase market demand has been a great benefit for the industry. The Indian car companies in like Toyota, Hyundai, Maruti Suzuki, and Ford are doing extremely well with great success along with other players such as M&M and Hindustan Motors are devising new techniques to accelerate growth. India is soon becoming a hub for car manufacturers not only to sell their cars but to setup manufacturing units. In April 2014, passenger car sales stood at 1,786,899 units while utility vehicles sales stood at 525,942 units, as per data from Society of Indian Mobile Manufacturers (SIAM).

Depending on the demand for the product the price fluctuates. The possibility of constant price for an automobile for a period of time is also high. The condition of the overall economy is still a primary determinant of general sales volume. Therefore having a sales forecast on the automobile industry will sure help firms to take strategic decisions to have a solid growth in the Indian automobile industry by taking cautious measures. This paper aims at providing a proposed neural network solution that can be used to predict the future sales on the base of historical data using Back Propagation algorithm.

There are countless issues, problems, and considerations in forecasting for new product. Understanding what a sales forecast is and what is designed to do is very important. A sales forecast is an educated guess of future performance based on sales and expected market conditions. The value of the forecast is that we can predict and prepare for the future objectively. The objective is to review the past, be focused in the present and follow the trends of the past and present to predict the future.Typically due to things that are out of our control, it is very difficult to forecast precisely. Mathematically, however, it is very possible to forecast sales with accuracy. The following are few of the major events that will affect the sales forecast: Seasonality, the economy, politics, fashion changes, average household income, demographic changes, and production requirements.It also includes:

A)*Lack of Sales History*
Sales forecast are based upon what the company has been able to achieve in the past. Early stage companies do not have significant revenue history to rely on and it is difficult to accurately forecast what sales might be without a track record.

B)*Difficulty Forecasting the Industry and Competition*
You never know how your industry sales would be in the next year. Whether the new competitors would come in and take away your sales. In this case sales forecasting would be a tedious job.

C) *Customer Behavior is hard to Predict*
The customers may not find your marketing campaign convincing compared to the other competitors. The customers taste may differ at any point of time.

D) *Wrong Predictions*
The problems faced by every company is that markets are unpredictable. This issue may cause the company's Sales fall down to a greater extent.

III.   SURVEY OF HISTORICAL RESULTS

As a test data we selected some of the fast growing Indian automobile industries secondary sales data through web browsing. We especially aimed at collecting the sales of 10 automobile industries based on the yearly sales basis. The data collected was present in crores.The data varied year by year which caused rise and fall in company's sale.

It was found that even though few of the companies lagged behind, the next year there was tremendous change in sales which brought growth for the companies. It was difficult to collect data for few companies as they recently emerged in Indian Automobile Industry but still made a great impact in the automobile sales. This variation and changes in sales shows how much faster the Indian Automobile Industry is growing in today's era. This project how solid the industry will grow in next few years but will have a cautious growth due to improved affordability, the rise in incomes and unpredictable market. This will be just another opportunity for the Indian market.

Company	2014	2013	2012	2011	2010	2009
Maruti Suzuki	43700.6	43587.9	35587.1	36618.4	29098.9	20453.7
Hyundai		25110	22103.48	19465.42	19622.54	15522.55
Tata	34288.11	44765.72	54306.56	47088.44	35024.66	25149.87
Renault				14.67		
Ford		6530	6604.18	5636.5	2144.38	1754.32
Fiat		1922.11	4978.03	3429.65	2753.92	695.81
BMW			2349.36			
Toyota	13216.31	14889.35	11452.21	7724.13	5603.7	3760.31
Volkswagen		4916.5	5375.6			
Honda		3959.91	3248.46	4001.03		

The conversion of sales data present in crores as inputs was done as follows:
- The data equal to and greater than10, 000 crores was converted as input with sales amount followed by the decimal point.
- Data less than 10,000 and greater than equal to 1000 was converted as input with a zero's succeedingthe decimal point followed by the sales amount.
- Data below 1000 was converted as input with two zero's succeeding the decimal point followed by the sales amount.

Company	2014	2013	2012	2011	2010	2009
Maruti Suzuki	0.43700	0.43587	0.35587	0.36618	0.29098	0.20453
Hyundai		0.25110	0.22103	0.19465	0.19622	0.15522

Tata	0.34288	0.44765	0.54306	0.47088	0.35024	0.25149
Renault				14.67		
Ford		0.06530	0.06604	0.05636	0.02144	0.01754
Fiat		0.01922	0.04978	0.03429	0.02753	0.00695
BMW			0.02349			
Toyota	0.13216	0.14889	0.11452	0.007724	0.005603	0.003760
Volkswagen		0..004916	0.005375			
Honda		0.003959	0.003248	0.004001		

IV. FRAMEWORK FOR PREDICTING ANALYTICS ON SALES FORECAST USING BACK PROPAGATION NETWORK

A sales forecast predicts the value of sales over a period of time. It becomes the basis of marketing mix and sales planning. Sales forecasting is crucial because without a proper sales forecast a company cannot program to attain the desired sales and marketing objectives. It is based on a number of assumptions regardingcustomer and competitor behavior as well as the market environment, and therefore, its reliability dependsupon a number of uncertain parameters. There are several topologies possible for artificial neural networks. Among which the Back Propagation neural network is commonly used. The back propagation (BP) neural network algorithm is a multi-layer feed forward network where input signals go forward, errors are calculated from respected outputs and these errors are sent to previous layers for weight adjustment.
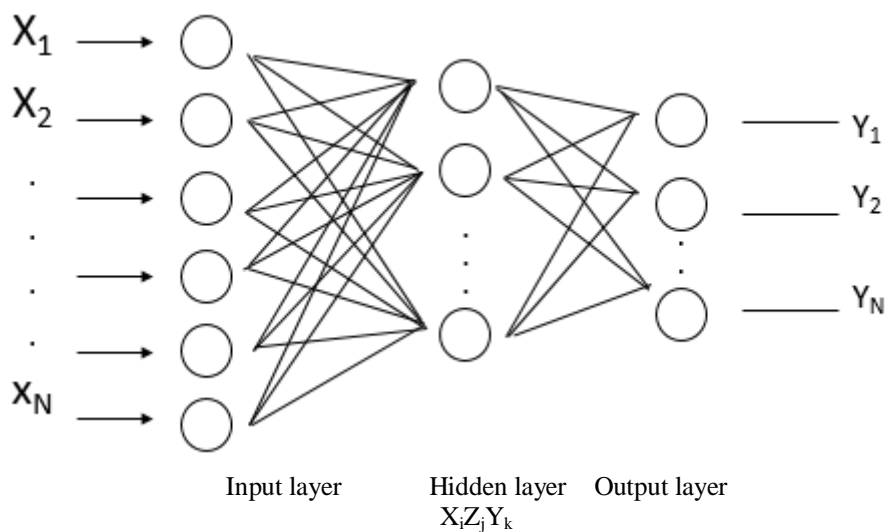

Figure 1:Back Propagation nueral network

Back Propagation Network Algorithm:
Back-propagation of error (Phase I)
Step 1:
Initialize weights and learning rate

Step 2:
Perform steps 2-9 when stopping condition is false

Step 3:
Perform Steps 3-8 for each training pair

Step 4:
Each input unit receives input signal X_i and sends it to the hidden unit (i=1 to n)

Step 5:
Each hidden unit Z_j (j=1 to p) sums its weighted input signals to calculate net input:

$$Z_{inj} = V_{0j} + \sum_i x_i v_{ij}$$

Calculate output of hidden unit by applying its activation functions over Z_{inj}
$$Z_j = f(Z_{inj})$$
And send the output signal from the hidden unit to the input of output layer units

Step 6:
for each output unit y_k (k=1 to m) calculate net input:

$$Y_{ink} = W_{0k} + \sum_{i=1}^{p} z_j w_{jk}$$

Applying its activation functions to compute output signal
$$Y_k = f(Y_{ink})$$

Back-propagation of error (Phase II)
Step 7:
Each output unit Y_k (k=1 to m) receives target pattern corresponding to the input training pattern and computes the error correction term:
$$\delta_k = (t_k - y_k)f'(y_{ink})$$
On the basis of the calculated error correction term, update the change in weights and bias
$$W_{jk} = \alpha \delta_k z_i$$
$$W_{0k} = \alpha \delta_k$$
Also send δ_k to the hidden layer backwards

Step 8:
Each hidden units $(Z_j, j=1$ to p) sums its delta inputs from the output units:

$$\delta_{inj} = \sum_{k=1}^{m} \delta_k w_{jk}$$

The term δ_{inj} gets multiplied with the derivative of f ($Zinj$) to calculate the error term:
On the basis of the calculated δ_j, update the change in weights and bias:
$$\triangle v_{ij} = \alpha \delta_j x_i$$

$$v_{0j} = \alpha \delta_j \qquad \triangle$$
Weight and bias updating (Phase III)

Step 9:

Each output unit (Y_k, k=1 to m) updates the bias and weights:

$$W_{jk(new)} = W_{jk(old)} + \quad W\triangle$$
$$W_{0k}(new) = W_{0k}(old) + \quad V\triangle$$

Each hidden unit (Z_j j=1 to p)updates its bias and weights:

$$V_{ij}(new) = V_{ij}(old) + \quad V\triangle$$

Step 10:

 Check for the stopping condition. The stopping condition may be certain number of epoch reached or when the actual output equals the target output

V.   LIMITATIONS

Sales forecasting if done incorrectly then the business will waste capital used for buying components for the products. The market is constantly changing this may affect the organization. Generally the assessment made from theses prediction give an approximation value but not the details.

VI.   FUTURE STUDY

This approach is able to determine the relationship between the historical data supplied to the system during the training phase and on that basis make automobile sales prediction for the future.

REFERENCES

1. Alekh Dwivedi, Maheshwari Niranjan, KalicharanSahu. 'A Business Intelligence Technique for Forecasting the Automobile Sales using Adaptive Intelligent Systems (ANFIS and ANN)', International Journal of Computer Applications, Volume 74, July 2013.
2. Rashmi Sharma,Ashok K. Sinha. 'Sales Forecast of an Automobile Industry',International Journal of Computer Applications,Volume 53–No.12, September 2012.
3. S.N. SIVANANDAM, S. D. (2014).Principles of Soft Computing. Wiley.
4. Pradeepta K. Sarangi, Shahin Bano, Megha Pant. 'Future Trend in IndianAutomobile Industry: AStatisticalApproach',Apeejay - Journal of Management Sciences and Technology, February 2014.
5. Tian-Syung Lan, Pin-Chang Chen, Chun-Hsiung Lan, Kai-Chi Chuang. 'A Study of Using Back-Propagation Neural Network for the Sales Forecasting of the Thin Film Sputtering Process Material',International Journal on Advancements in Computing Technology, Volume4, May 2012.
6. Sukhmeet Kaur,Kiran Jyoti. 'Predicting the future of car manufacturing industry using Naïve Bayse Classifier',International Journal for Science and Emerging Technologies with Latest Trends,July 2012.
7. Dr. M.Hanumanthappa1,Sarakutty.T.K. 'Predicting the Future of Car Manufacturing Industry using Data Mining Techniques',Int. J. on Information Technology, Vol. 01, September 2011.
8. P.C. Chen, C.Y. Lo, H.T. Chang and Y. Lo Cho, "A Study of Applying Artificial Neural Network and Genetic Algorithm in Sales Forecasting Model", Journal of Convergence Information Technology- 2011.
9. Automobile Industry in India,Indian Autommobile Industry. (n.d.). Retrieved from ibef.org: http://www.ibef.org/industry/india-automobiles.aspx
10. Beal, V. (n.d.). What is Predictive Analytics and Big Data? Webopedia. Retrieved from webopedia.com: http://www.webopedia.com/TERM/P/predictive_analytics.html
11. Business Intelligence:Predictive Analytics. (n.d.). Retrieved from bim-businessintelligence.blogspot.in: http://bim-businessintelligence.blogspot.in/2011/04/predictive-analytics.html
12. Citeseerx-A nonlinear forecasts combination method based on Takagi-Sugeno fuzzy systems (1998) . (n.d.). Retrieved from citeseerx.ist.psu.edu: http://citeseerx.ist.psu.edu/showciting?cid=5347434
13. Distinguish and bring out the major differences between sales planning and sales forecasting. . (n.d.). Retrieved from marketingformanagers.blogspot.in: http://marketingformanagers.blogspot.in/2009/06/distinguish-and-bring-out-major.html
14. Fact-Sheet-Automobile-. (n.d.). Retrieved from indusview.com: http://www.indusview.com/factsheet_9.html
15. IJCA-A Business Intelligence Technique for Forecasting the Automobile Sales using Adaptive Intelligent Systems (ANFIS and ANN) . (n.d.). Retrieved from ijcaonline.org: http://www.ijcaonline.org/archives/volume74/number9/12911-9383
16. International Journal for Science and Emerging Technologies with Latest Trends. (n.d.). Retrieved from http://ijsett.com/: http://www.oalib.com/journal/10944/1#.VCVeYCbhXIU
17. Jackson, J. D. (n.d.). Issues in Sales Forecasting-Essays-Jamesdjackson. Retrieved from studymode.com: http://www.studymode.com/essays/Issues-In-Sales-Forecasting-599018.html

18. Limitations of Sales Forecasting. (n.d.). Retrieved from http//yourbusiness.azcentral.com: I:\Report_PREDICTIVE ANALYTICS USING SOFT COMPUTING123.docx.html
19. Model Extremely Complex Functions, Neural Networks. (n.d.). Retrieved from www.statsoft.com: http://www.statsoft.com/Textbook/Neural-Networks/button/2
20. Predictive Analysis for Dummies-Big Sonata. (n.d.). Retrieved from bigsonata.com: http://bigsonata.com/predictive-analytics-for-dummies/
21. Predictive analytics-Ask in wiki. (n.d.). Retrieved from http://ask.inwiki.org: http://ask.inwiki.org/Predictive_analytics
22. Predictive analytics-Wikipedia, the free encyclopedia. (n.d.). Retrieved from inimino.org: http://inimino.org/~inimino/tests/js/tree_explorer/wikipedia/Predictive_analytics.html
23. Predictive Analytics-Wikipedia, the free encyclopedia. (n.d.). Retrieved from wikipedia.com: http://en.wikipedia.org/wiki/Predictive_analytics
24. Ravindran, D. (n.d.). Mkt-slideshare. Retrieved from slideshare.com: http://www.slideshare.net/devi.ravindran/mkt-11648447
25. Scifort-Neural Network. (n.d.). Retrieved from scifort.com: http://www.scifort.com/index.php?option=com_content&task=view&id=91&Itemid=139
26. Soft computing - Wikipedia, the free encyclopedia. (n.d.). Retrieved from Wikipedia.com: http://en.wikipedia.org/wiki/Soft_computing
27. Soft Computing-Intelligent Software Reliability. (n.d.). Retrieved from reliabilitymodeling.com: http://reliabilitymodeling.com/?page_id=27
28. Softcomputing-Data Mining and Soft computing techniques. (n.d.). Retrieved from weebly.com: http://dataminingzone.weebly.com/softcomputing.html
29. Vira, V. (n.d.). Final Project-scribd. Retrieved from scribd.com: http://www.scribd.com/doc/61758843/Final-Project
30. 30. Automobile Industry in India,Indian Automobile Industry,Sector,Trends,Statistics. (n.d.). Retrieved from www.ibef.org: http://www.ibef.org/industry/india-automobiles.aspx