A Business Intelligence Technique for Forecasting the Automobile Sales using Adaptive Intelligent Systems (ANFIS and ANN)

Alekh Dwivedi
Department of Information Technology
Jaypee University, Guna

Maheshwari Niranjan
Department of Information Technology
TIT, Bhopal

Kalicharan Sahu
School Computing Science
Galgotias Uni., Greater Noida

ABSTRACT
Today, Sales forecasting plays a key role for each business in this competitive environment. The forecasting of sales data in automobile industry has become a primary concern to predict the accuracy in future sales. This work addresses the problem of monthly sales forecasting in automobile industry (maruti car). The data set is based on monthly sales (past 5 year data from 2008 to 2012). Primarily, we used two forecasting methods namely Moving Average and Exponential smoothing to forecast the past data set and then we use these forecasted values as a input for ANFIS (Adaptive Neuro Fuzzy Inference System). Here, MA and ES forecasted values used as input variable for ANFIS to obtain the final accurate sales forecast. Finally we compare our model with two other forecasting models: ANN (Artificial Neural Network) and Linear Regression. Empirical results demonstrate that the ANFIS model gives better results out than other two models.

Keywords
Sales Forecasting, ES, MA, Adaptive Neuro Fuzzy Inference System, ANN, Linear Regression.

1. INTRODUCTION
Automobile industry plays an important role in modern manufacturing in India. Although, sales forecasting helps the industry to drive sales revenue, improve efficiency, increase customer retention and reduce costs, it is the hardest part of management [1-2]. However, a automobile market depends on the index of key factors including wholesale price, population, income, and oil price causes the demand in automobile sales instability in this current market. In forecasting researches, the common forecasting methods are Moving Average, Exponential Smoothing, so first we have forecasted our sales data using these methods and then use these forecasted values as a input variable for popular non linear forecasting method such as ANFIS (Adaptive Neuro Fuzzy Inference System). After getting forecasted results by ANFIS method, we compared these result with two other forecasting method such as ANN and Linear Regression. However, the forecasting's methods are not limited to the mentioned methods. For instance, ANFIS (Adaptive Neuro Fuzzy Inference System) as one of the newest methods is applied in different areas such as medical, industry, geography, econometric etc. In these areas, variety of tasks such as analysis of decision, forecasting, pattern recognition, system control, inventory management, logistic systems and operations management are used by ANFIS [3]. ANFIS is extended by an expert system which is capable to use combined fuzzy reasoning with the pattern recognition capability of neural networks.

The ANFIS is a neuro-fuzzy system, which uses a feed-forward network to search for fuzzy decision rules that perform well on a given task. Using a given input/output data set, ANFIS creates a fuzzy inference system whose membership function parameters are adjusted using a back-propagation algorithm alone or combination of a back-propagation algorithm with a least mean squares (LMS) method (hybrid learning). This allows the fuzzy systems to learn from the data being modeled. ANFIS provides a method for the fuzzy modeling procedure to learn information from the data set, followed by creation of the membership function parameters that best performs the given task. The ANFIS can simulate and analyze the mapping relation between the input and output data through a learning algorithm to optimize the parameters of a given Fuzzy Inference System (FIS) [4]. The aim of this research is the application of ANFIS as powerful features of fuzzy inference system model in automobile industry sales data forecasting (2008-2012).

2. SALES FORECASTING
The sales forecasting for automobile industry provide the best way to predict the future sales market demand [5]. An efficient forecasting system can help to improve machine utilization, achieve greater flexibility to change and increase profits [6]. Sales forecasting is particularly important because its outcome affects many functions in the organization [7].

In this paper, the popular technique called ANFIS is used to predict the future sales accuracy. Other statistical forecasting approaches such as exponential smoothing and moving average have recently used to predict the forecasting value for short period [8].

Since the sales are dependent on many factors, the sale forecasting is not an easy job. However, there are some methods to increase the accuracy sales forecasting such as Adaptive Neuro Fuzzy Inference Systems (ANFIS) method. In this paper we forecast the sales of Automobile Industry in India based on forecasted sales produced by two forecasting method moving average and exponential smoothing as input variables to the ANFIS method. In addition, we consider Root mean squared error (RMSE) as a scale to compare ANFIS with ANN and Linear Regression. By using some data from Maruti car Industry in India. it is seen that ANFIS method is better than ANN & Linear Regression based Method.

3. METHODOLOGY
3.1 Moving Average (MA)
Moving Average is simply a time series analysis forecasting method. The advantage of MA is that it is simple and easy to use. It also gives a basic and efficient tendency index. Based on this, this method is commonly used to make forecasts from historical data [9].

Moving Average (MA) is a popular method for averaging the results of recent sales history to determine a projection for the short term. The MA forecast method lags behind trends. This method works better for short range forecasts of mature

products than for products that are in the growth or obsolescence stages of the life cycle.

3.2 Exponential Smoothing

In Exponential Smoothing, the system assigns weights that exponentially decay. The forecast is a weighted average of the actual sales from the previous period and the forecast from the previous period [8-9]. Alpha is the weight that is applied to the actual sales for the previous period. You should assign a value for the smoothing constant, alpha. If you do not assign a value for the smoothing constant, the system calculates an assumed value that is based on the number of periods of sales history that is specified in the processing option.

Forecast specifications:

- α equals the smoothing constant that is used to calculate the smoothed average for the general level or magnitude of sales.

Values for alpha range from 0 to 1.

- n equals the range of sales history data to include in the calculations.

3.3 Linear Regression

Linear regression is a form of regression analysis and can be used to calculate a mathematical relationship between two (or more) sets of data. In forecasting, you would use this if you thought one set of data could be used to predict another set of data. A typical approach towards regression analysis involves using regression to determine the mathematical relationship, but also to help give you an idea of how valid that relationship is (that's the analysis part). The Forecast Function skips the analysis, and just calculates a relationship and automatically applies it to your output. This makes things easier for the user, but it assumes your relationship is valid. So essentially, the Forecast function uses linear regression to predict a value based on a relationship between two sets of data [10].

Regression analysis is a statistical tool for the investigation of relationships among variables. Regression analysis is widely used for prediction and forecasting. The focus is on the relationship between a dependent variable and one or more independent variables. More specifically, regression analysis helps one understand how the typical value of the dependent variable changes when any one of the independent variables is varied, while the other independent variables are held fixed. Linear regression including the simple linear or polynomial regressions, consider only one predictor as a linear or nonlinear form. Linear regression has been used widely in practical applications.

3.4 ANN (Artificial Neural Network)

Neural networks are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems [11]. As in nature, the connections between elements largely determine the network function. You can train a neural network to perform a particular function by adjusting the values of the connections (weights) between elements. Typically, neural networks are adjusted, or trained, so that a particular input leads to a specific target output. There, the network is adjusted, based on a comparison of the output and the target, until the network output matches the target. Typically, many such input/target pairs are needed to train a network. Neural networks can also be trained to solve problems that are difficult for conventional computers or human beings.

3.5 Adaptive Neuro Fuzzy Inference System (ANFIS)

Adaptive Neuro Fuzzy Inference System (ANFIS) has been suggested by (1993). Adaptative Neuro-Fuzzy Inference System (ANFIS) is one of the hybrid neuro-fuzzy inference expert systems and it works in Takagi-Sugeno-type fuzzy inference system. The technique provide a method for the fuzzy modeling procedure to learn information about a data set, in order to compute the membership function parameters that best allow the associated fuzzy inference system to track the given input/output data [12-13]. This learning method works in a manner similar to that of neural networks. Many types of FIS have been proposed in literature such as Mamdani, Sugeno and etc (Sugeno 1988). However, in the implementation of an ANFIS for financial predictions and estimation the most suitable model is the Sugeno model. The Sugeno model makes use of the rules to produce an output for each rule. Rule outputs consist of the linear combination of the input variables plus a constant term; the final output is the weighted average of each rule's output.
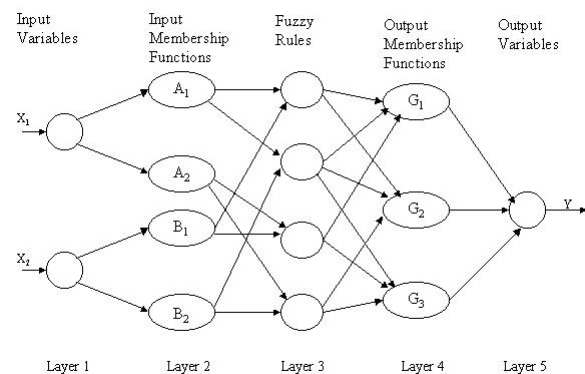

Figure 1:- general anfis architecture

4. SIMULATION RESULTS
ANFIS Based Forecasting Model

In Neuro-fuzzy model, The membership function of each input was tuned using the hybrid method consisting of back propagation for the parameters associated with the input membership function and the least square estimation for the parameters associated with the output membership functions. The computations of the membership function parameters are facilitated by a gradient vector which provides a measure of how well the FIS system is modeling the input/output data. For a given set of parameters, the numbers of nodes in the training data were found to be 40. The numbers of linear parameters and non-linear parameters were found to be 27 and 18 respectively. The hypothesized FIS model is trained to emulate the training data by modifying the membership function parameters according to the chosen error criterion. A suitable configuration has to be chosen for the best performance of the network.

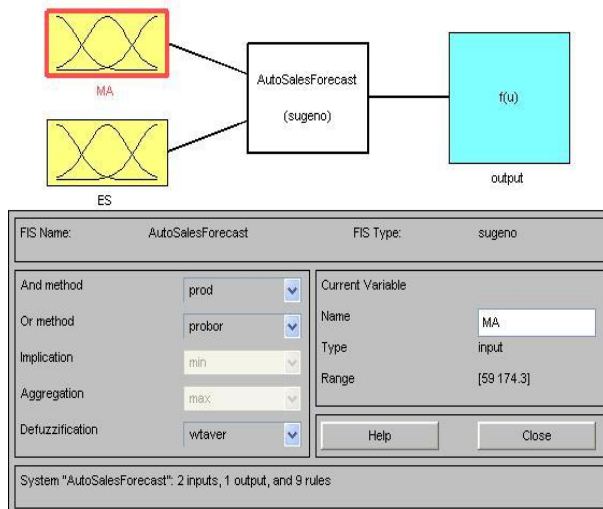**Sugeno Model for Sales Forecasting:-**
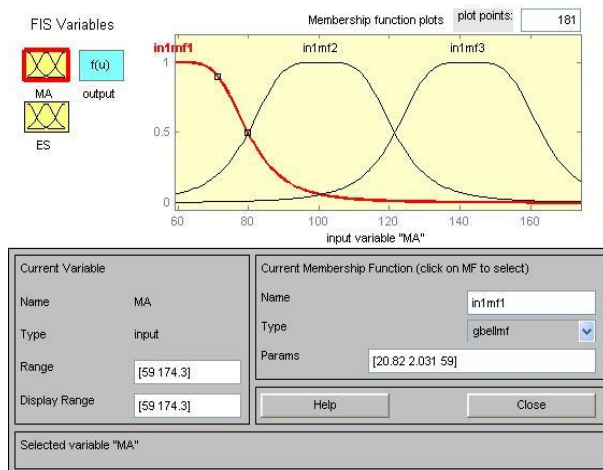

Figure 2:- sugeno fis model for sales forecasting
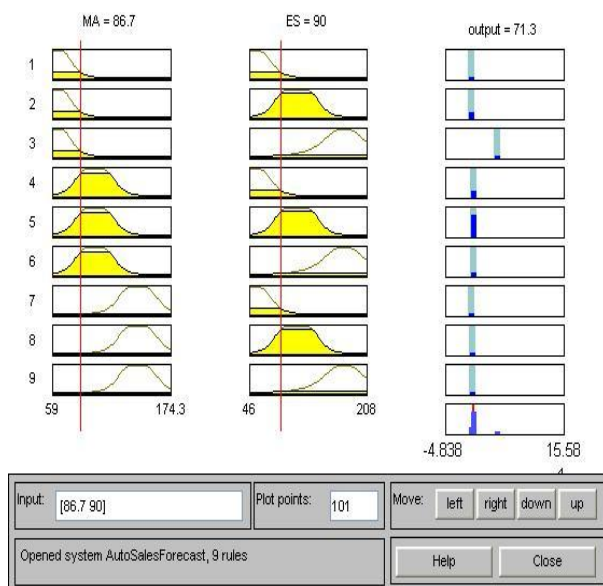

Figure 3:- membership function editor

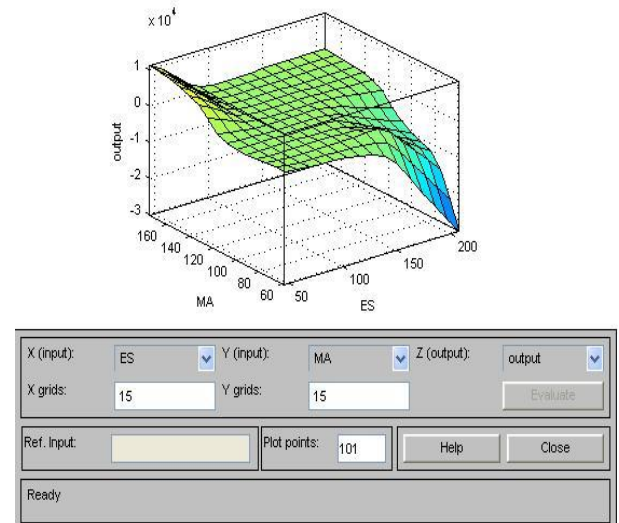**Figure 4:- rule base for sales forecasting model**

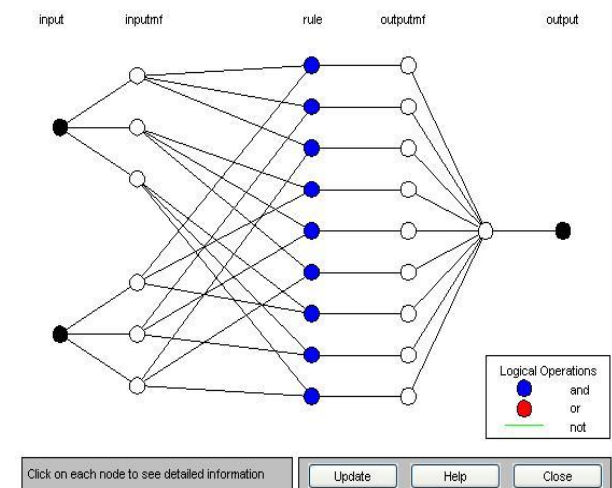**Figure 5:- surface viewer for sales forecasting model**

Figure 6:- ANFIS structure for sales forecasting analysis, Neuro-fuzzy, with 2 input parameters and 9 rules. Layer-1 represents inputs, layer-2 input membership function, layer-3 rules, layer-4 output membership function, and layer-5 weighted sum output and layer-6 output.

The error was set to be 0 and number of training epochs was given 10. After the training (with 10 epochs) was complete final configuration for the FIS are Number of output membership functions = 9 Number of fuzzy rules = 9 Neuro-fuzzy adaptive networks for sales forecasting accuracy:

1. No. of input = 2

2. No. of membership functions for each input = 3

3. Type of membership functions for input = gbell

4. No. of rules = 9

5. Type of membership functions for each output = Linear

6. No. of output membership function = 9

7. No. of training epochs = 10

8. No. of training data set pairs = 40

9. No. of checking data set pairs = 20

10. Error tolerance = 0

11. Error achieved = 5.6591

Neuro-fuzzy, with two input parameters and 9 rules is shown in figure. ANFIS info: Number of nodes: 35 Number of linear parameters: 27 Number of nonlinear parameters: 18 Total number of parameters: 45 Number of training data pairs: 40 Number of checking data pairs: 20 Number of fuzzy rules: 9

Start training ANFIS...

1.   5.6732   239.2481

2.   5.6591   238.7083

Designated epoch numbers reached -> ANFIS training completed at epoch 2.
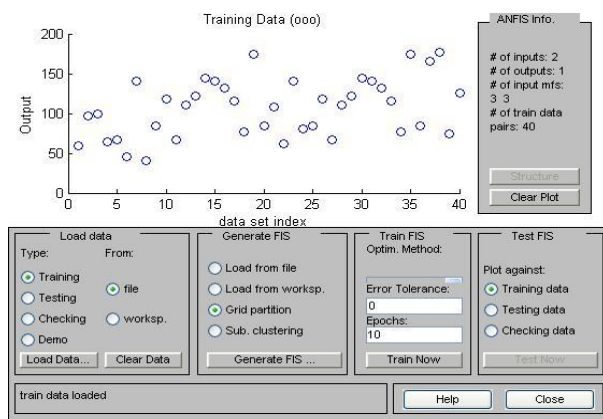
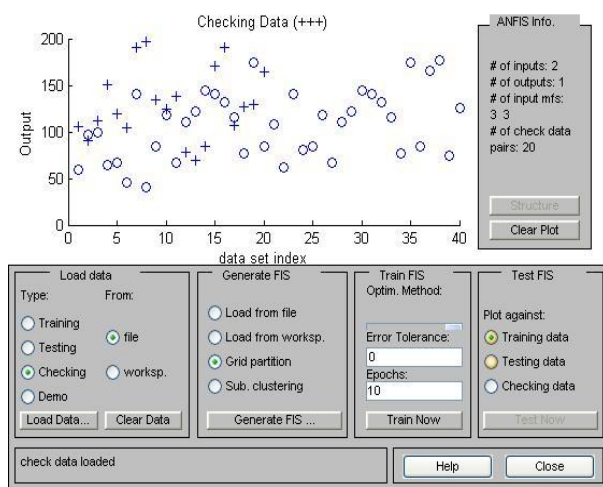**Figure 7:- training dataset for sales forecasting model**

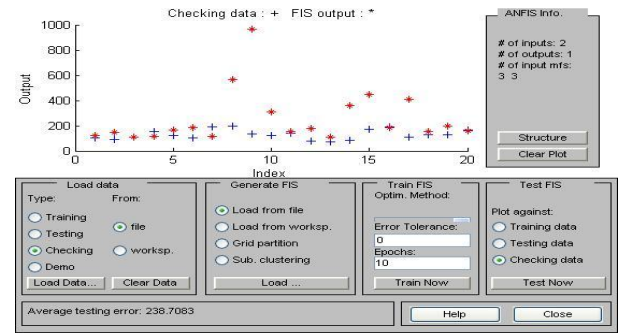**Figure 8:- checking dataset for sales forecasting model**

Figure 9:- testing error for sales forecasting model
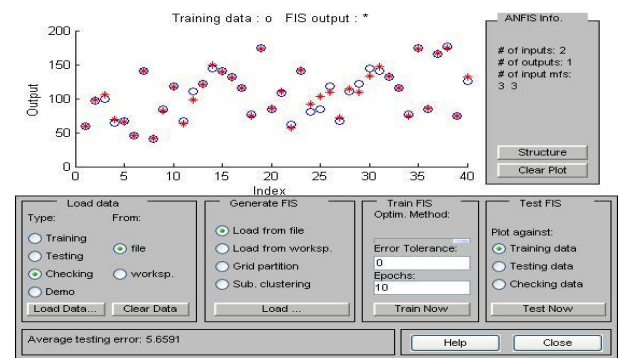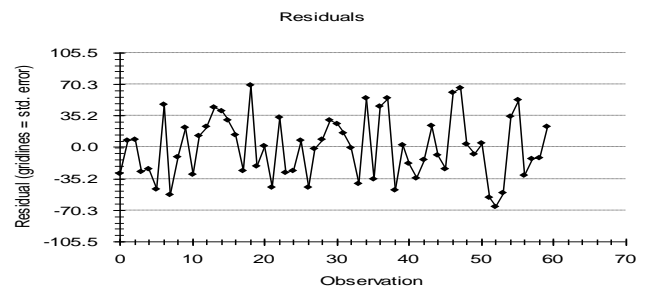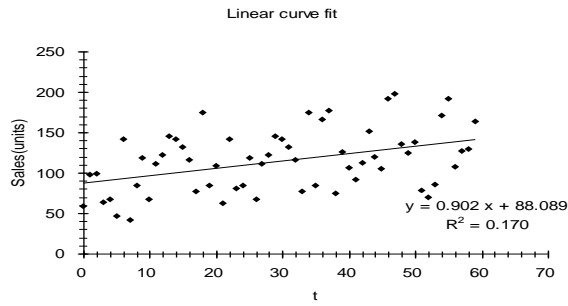

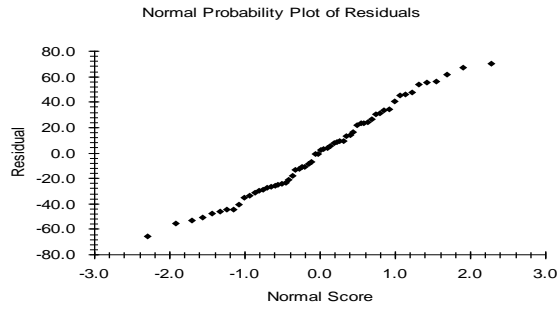**Figure 10:- training error for sales forecasting model**

5. RESULTS OBTAINED FROM LINEAR REGRESSION AND ANN
5.1 Regression Results:-

Regression output
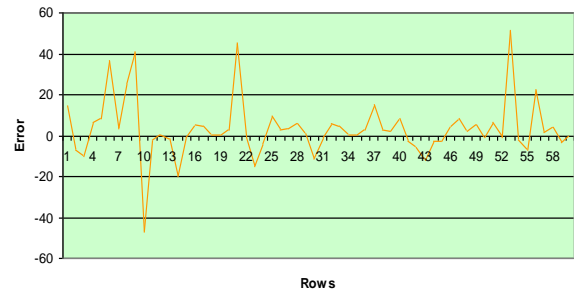
variables	coefficients	std. error	t (df=58)	p-value
Intercept	88.0885	8.9667	9.824	5.97E-14
t	0.9021	0.2621	3.441	.0011

Normal Probability Plot of Residuals

Deviation Graph

Linear curve fit

$y = 0.902 x + 88.089$
$R^2 = 0.170$

Actual vs. Forecast (Scatter Plot)
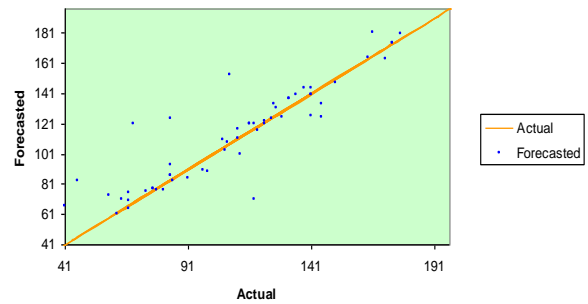
5.2 ANN Results:-

Actual vs. Forecast
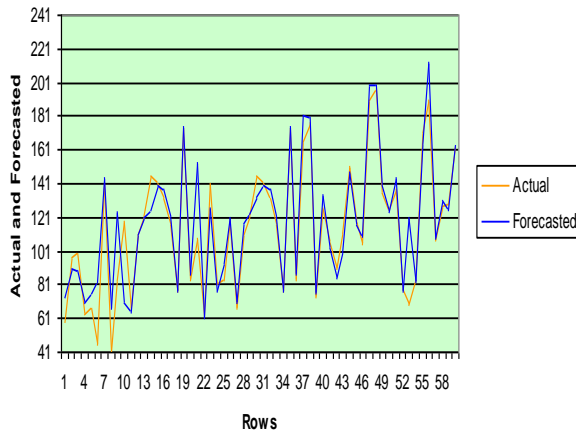
	Training set	**Test set**
# of rows:	50	10
CCR:	n/a	n/a
Average AE:	7.094532448	17.204825
Average MSE:	176.2832379	454.14623
Tolerance type:	Relative	Relative
Tolerance:	10%	30%
# of Good forecasts:	42 (84%)	7 (70%)
# of Bad forecasts:	8 (16%)	3 (30%)

Scatter Plot of Monthly Forecasted Sales Data (LR, ANN, ANFIS):-

Monthly Forecasted Sales by LR

y = 0.902 x + 87.186
R^2 = 1.000

2008-Forecast by ANFIS

y = 1.143 x + 75.795
R^2 = 0.021

Monthly Forecasted Sales by ANN

y = 0.901 x + 90.637
R^2 = 0.175
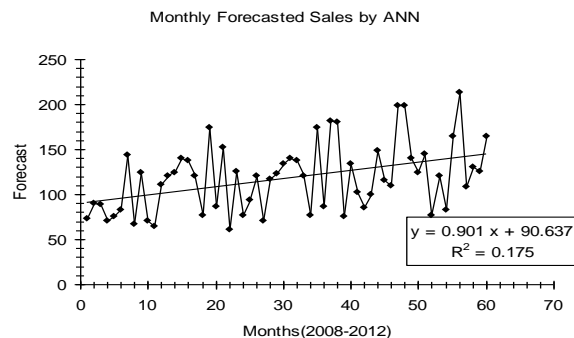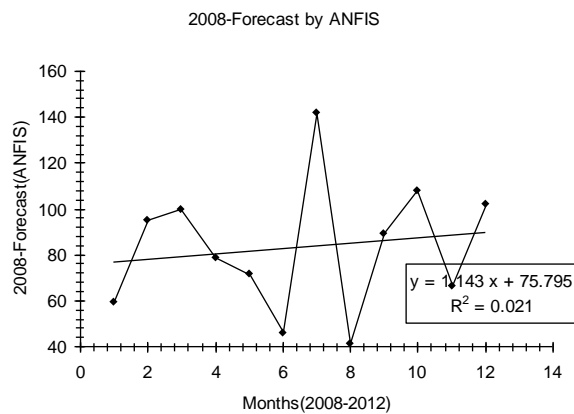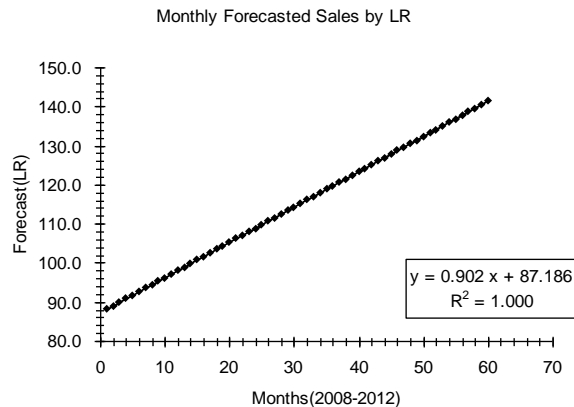
6. CONCLUSION

Forecasting the sales data is an important task in today business to predict the actual future sales demand of product in this competitive market. In this paper, we use ANFIS method to predict the sale of automobile industry in India. In this paper, we consider MA & ES forecasted sales of actual sales data as input to ANFIS system. We forecast automobile sale with ANFIS, ANN and linear regression. By considering RMSE as a compared scale, it is shown that ANFIS is better than other two.

The study shows that the performance of sales prediction can be significantly enhanced by using ANFIS. There is much scope in extending the present work. For example, considering quality factors such as appearance that a significant effect on demand in the forecasting model. Compare the proposed model with other forecasting techniques. Finally train and test the data with other functions of ANFIS or high-tech technology.

7. REFERENCES

[1] Chang, P., C., Liu, C.H., and Fan, C.Y., Data clustering and fuzzy neural network for sales forecasting: A case study in printed circuit board industry, Knowledge-Based Systems, Volume 22, Issue 5, Pages 344355,2009.

[2] Abu-Eisheh, S. A., & Mannering, F. (2002). Forecasting automobile demand for economies in transition, a dynamic simultaneous-equation system approach. Transportation Planning and Technology, 25, 311–331.

[3] Kuo, R.J., Wu, P., and Wang, C.P., An intelligent sales forecasting system through integration of artificial neural networks and fuzzy neural networks with fuzzy weight elimination, Neural Netw. Sep;15(7):909-25., 2002.

[4] Fiordaliso, A. (1998). A nonlinear forecasts combination method based on Takagi-Sugeno fuzzy systems. International journal of forecasting, 14:367-379.

[5] Makridakis, S. and Wheelwright, S. C., *In Forecasting Methods for Management*, 5th Ed., Wiley, Chichester (1989).

[6] Winklhofer, H. and Diamantopoulos, A., "A model of export sales forecasting behavior and performance: development and testing," International Journal of Forecasting, Vol. 19, pp. 271-285 (2003).

[7] Mentzer, J. T. and Bienstock, C. C., Sales Forecasting Management: Understanding the Techniques, Systems and Management of the Sales Forecasting Process, Sage publications, Thousand Oaks, CA (1998).

[8] Diamantopoulos, A. and Winklhofer, H., "Export sales forecasting by UK firms technique utilization and impact on forecast accuracy," *Journal of Business Research*, Vol. 56, pp. 45-54 (2003).

[9] C.C. Holt, Forecasting seasonals and trends by exponentially weighted moving averages. International Journal of Forecasting 20(1), pp.5-10, 2004.

[10] Wood, D. and Dasgupta, B., "Classifying trend movements in the MSCI U.S.A. capital market index-a comparison of regression, aroma and neural network methods," *Computers & Operations Research*, Vol. 23,No. 6, pp. 611-622 (1996).

[11] Shtub, A., Versano, R., Estimating the cost of steel pipe bending, a comparison between neural networks and regression analysis, International Journal of Production Economics, vol.62, no.3, pp.201-207, 1999.

[12] Jang, J. S. R., & Gulley, N. (1995). The fuzzy logic toolbox for use with MATLAB. Natick, MA:The Math WorksInc..

[13] Firat, M., & Güngör, M. (2008). Hydrological time series modeling using an adaptive neuro-fuzzy inference system. Hydrological Processes, 22, 2122–2132.

[14] K.Y.Chen, Combining linear and nonlinear model in fore-casting tourism demand, Expert Systems with Applications 38(8), pp.10368-10376, 2011.

[15] Martino, J.P., Technological forecasting for decision mak-ing (3rd ed.). New York: McGraw-Hill, 1993.

[16] Boyacioglu, M.A., Avci, D., An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the prediction of stock market return: The case of the Istanbul Stock Exchange, Expert Systems with Applications: An International Journal , Volume 37 Issue 12, December, 2010.

[17] S.H.Lim, Forecasting models of additional use of mobile digital contents: A comparison of artificial neural networks and logistic regression analysis, International Journal of Computer Science and Network Security 6(6), 146-149,2006.