Survey on During Pandemic scenario: Domestic Retail Sales Forecasting of Passenger Vehicles in India using Time Delay Neural Network

[1]Dr. Vijay Bhardwaj, [2]Dr. Sunny Arora
[1]University College of Engg & Technology
[1]Guru Kashi University, Talwandi Sabo

ABSTRACT
Accurate sales forecast is crucial to plan the production process, especially in automotive industry, where a large number of factors –macro-economic, micro-economic and others dynamically influence the demand.Forecasting in such a dynamic industry, especially with the onset of COVID, has become extremely important with the preference for personal transportation on the increase. In this paper, a neural network model to forecast the domestic retail sales of passenger vehicles in India is formulated. A Time Delay Feed Forward Neural Network with Back Propagation (TDNN)is used and the forecasting accuracies determined by comparing the values of Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Deviation (MAD) and Mean Absolute Percentage Error (MAPE). The macro-economic indicators identified were Unemployment rate, GDP growth rate, Private Final Consumption Expenditure (PFCE), Index of Industrial Production (IIP), Bank Lending Rate (BLR), Inflation Rate and Crude oil prices. The indicators were taken as the input variables and Passenger Vehicle Retail Sales was selected as the target variable to formulate the TDNN model. It was observed that the TDNN model was able to forecast the output with great accuracy.
Keywords: Demand Forecasting, Artificial Neural Network, Delayed Neural Networks, Passenger Vehicles, Retail Sales, GDP growth rate, Inflation rate, IIP, Crude oil price.

1. Introduction

Automobile industry has been in the process of rapid transformation in recent years. It is one of the fast-changing industry influenced by changes in technologies, customer preferences, demand patterns, economic and econometric factors. India is the fifth largest automobile market with an average sales increase of 8.3% year -on-year and is expected to reach the third position globally in terms of volume by 2026. It holds a strong market presence inboth domestic demand and exports. The sector contributes 7.1% to the country's GDP, 22% to the manufacturing GDP and has generated 35 million employment opportunities [1]. The automotive mission plan 2020-26 (AMP 2026), a collective vision of Govt of India and the Indian AutomotiveIndustry, has set the objective of increasing the contribution of the Indian Automotive sector in excess of 12% to the country's GDP and more than 40% of the manufacturing GDP over the next decade [2].Passenger vehicles account for 13% market share of the total vehicles sold in financial year 19. Urbanization, rising incomes and larger young population and furthermore, the onset of COVID, has increased the demand for personal transportation by passenger vehicles, especially cars.

Fierce market competitions in both domestic and export segments call for accurate sales forecasting. Manufacturers need to consider all the macroeconomic, microeconomic and other factors which influence the customer demand and hence the sales. The prediction accuracy affects the firms in their judgement of demand patterns, competitive strengths and strategies for product development. To effectively manage resources and to maximize the revenue, it is imperative that product portfolios need to be designed and positioned suitably (Sa-ngasoongsong et al., 2012). Automobile production involves long production cycles typically from 6-24 months. Hence effective production planning requires accurate long-term forecasting of sales and demand. Long term forecasting serves as an input to process planning and business planning strategies such as expansion or contraction of existing production units. Inaccuracies in demand forecasting has led to increased costs, losses in revenue and high sub-optimal levels of production workforce (Monks 1987). It can lead to disruption of supply chain due to shortage of materials if supply is outsourced. Sales forecasting (Chu and Zhang, 2003; Danese and Kalchschmidt, 2011; Luxhøj et al., 1996) is therefore considered to be a realistic assessment of expected future demand (Mentzer and Moon, 2005). Sales forecasting of automobiles has received quite a lot of attention in the recent years. Many models have been proposed by various researchers to forecast the demand and sales of automobiles. Statistical time series models such as the ARIMA model, Regression Analysis, Advanced exponential smoothing models have all been proposed with a fairly desirable level of forecasting accuracy. However, with the advent of Artificial Intelligence and Machine learning techniques, the forecasting capabilities have taken a new high. Use of Artificial Neural Networks (ANNs) in forecasting (Sharda, 1994) have resulted in better prediction capabilities with the advantage of capturing underlying relationships among nonlinear variables which otherwise were difficult to be determined. ANN can be treated as a nonlinear, non-parametric statistical method (White, 1989; Ripley, 1993; Cheng and Titterington, 1994) [3]. Evolution of ANNs in their structure and algorithms combined with use of statistical methods in combination has contributed in further enhancement of the prediction level. Multilayer Feed Forward is the most common architecture in neural networks which finds applications particularly in forecasting. Back propagation algorithm is the commonly used algorithm in multilayer networks. The presence of seasonality and non-linearity characteristics in automobile sales data makes the use of an ensemble model (A combination of two models: statistical model and MLFF-BPNN in this case), ideal for the purpose.

2. Literature Survey

2.1 Forecasting Models

Chow (1957) and Nerlove (1957) used an aggregate stock adjustment modelling framework which used a system of aggregate demand models to predict a desired level of automobile holdings. Suits (1958) used causal aggregate models to model and forecast automobile demand.

The application of time series models to forecasts of the registrations of new vehicles was originally established by Lewandowski (1970). The LR 799 model proposed by Tanner (1979), tried to predict the demand of cars in Britain using GDP per capita and operating costs as the variables. The model proved to be complex and required large quantity of data. Berkovec (1985) modelled the aggregate automobile market demand as the sum of individual consumer demands using household attributes and vehicle ownership levels[4]. Mannering and Winston (1987) developed a disaggregate model which used an elaborate algorithm to establish equilibrium vehicle prices for all makes, models and vintages of automobiles. Hamed and Abdulwahab (1996) used disaggregate models to forecast future automobile demand based on certain policy scenarios.

Garcia-Ferrer et al., (1997) used a variation of ARIMA model to forecast automobile sales and to compare the forecasting performance with other methods.ARIMA method suffered by its univariate nature and hence could not fully account for the effects of economic or social factors.Romilly et al. (1998) used linear regression model with differenced and lagged values of variables to estimate the forecast [4]. Greenspan and Cohen (1999) predicted aggregate sales of new motor vehicles based on quantifying vehicle stock and scrappage rates. Franses (1994b) extended the deterministic Gompertz model (Franses, 1994a; Meade, 1984) to predict Dutch new car sales. Armstrong et al. (2000) developed intentions-based forecasting methods to predict French and US automobile sales.Dudenh¨offer and Borscheid (2004) published a very important application of time series methods to the German automobile market.Kunhui et al. (2007) predicted Chinese automobile sales using support vector regression (SVR). Br¨uhl et al. (2009) proposed a sales forecast model for the German Automobile market based on time series analysis and data mining methods [5]. Hulsmann et al. (2012) applied data mining algorithms to model German and US automobile markets.Conparison of different models of time series forecasting can be observed involving traditional models and ANNs. Many research publications report the forecasting quality displayed by ANNs over traditional models like ARIMA and regression (Ansuj et al. 1996, Ho et al. 2002, Karbasi et al. 2009, Chattopadhyay 2010). (Goyal 2006, Díaz-Robles et al. 2008,Koutroumanidis et al. 2009, Khashei et al. 2012, Liu 2012, Pektaş et al.2013) proposed hybrid methods to perform better than the individual models. Zhang (2003) proposed a hybrid approach to combine statistical and ANN models to obtain better results. His model was modified and used by many researchers (Aburto 2007, Zou et al. 2007, Pham et al. 2010).Asilkan and Irmak (2009) predicted the future prices of automobiles by using ANNs. An automotive pricing model was formulated by İşeri and Karlık (2009) based on technical and physical characteristics. Hosoz and Ertunç used ANN to predict automobile performances. Karaatlı et al. (2012) estimated new cars sales in Turkey using ANN.

2.2 Macro-economic indicators

It is seen that macro-economic factors contribute widely to the automobile sales especially in the domestic segment. The indicators such as Unemployment rate, GDP growth rate, Private Final Consumption Expenditure (PFCE), Index of Industrial Production (IIP), Bank Lending

Rate (BLR), Consumer Price Index (CPI) which represents Inflation and Crude oil prices cannot be controlled by the companies, but directly affect the sales of automobiles. According to research publications in this area, several economic factors have been identified to influencing automobile sales which include fluctuation in fuel prices, interest on loans, unemployment rates etc

Ludvigson (1998) tested the impact of bank loans for car purchase. The increase of basic interest rate caused a significant negative impact on car sales. Dargay (2001) used Family Expenditure Survey from 1970 to 1995 to propose that the statistics of vehicle ownership recorded a positive upward trend with income increase. However, there is a negative correlation when there is an income reduction [6]. Dynaquest (2002) proposed that there is exists a strong relationship between new car sales and the nominal GDP. McManus (2007) proposed a link between vehicle sales and gasoline prices. Kongsberg Automotive (2008) analyzed the relationship between global car sales and global GDP from 1998 until 2008. The results show that there is a high correlation between global car sales and global GDP [7]. Smith and Chen (2009) established a historical correlation between annualized GDP and vehicle sales growth in US which indicated that positive vehicle sales growth depends on three percent or higher GDP growth [7]. Shahabudin (2009) in his research discovered that the variables like income level, interest rate, financial aggregate and unemployment rate significantly influenced car sales. His model however displayed heteroscedasticity.Mian and Sufi (2010) published the similiarity in the evolution patterns of automobile sales, unemployment rate and new housing building permits [4].Babatsou and Zervas (2011) found a good linear correlation between GDP and passenger car sales in the European Union countries.Muhammad, et al. (2012) analyzed the impact of economic variables on automobile sales in five ASEAN countries. They proposed that gross domestic product, inflation, unemployment rate and loan rate had significant long-term correlation with automobile sales in these ASEAN countries [6].Karaatlı et al. (2012) used gross domestic product, real sector confidence index, investment expenditures, consumer expenditures, consumer confidence index, dollar exchange rate and time as independent variables [8]. Chifurira, et al. (2014) examined the impact of inflation on the automobile sales in South Africa over the sample period of 1969 to 2013 [9]. The finding of the study indicated that there is a unidirectional causal effect (one-way causality) from inflation to new vehicle sales.

A large number of literatures were available on ANNs, hybrid models and on the macroeconomic indicators which influenced automobile sales, minimal works have been published which addresses the application of these models to forecast automobile sales using macroeconomic indicators. Shahabuddin (2009) modeled vehicle sales using indicators such as durable personal consumption and durable industrial demand [10]. Bruhlet al. ‘s (2009) data-driven model for the German automobile market relates various economic indicators such as Gross Domestic Product (GDP), Consumer Price Index (CPI), interest rate, unemployment rate and gas prices with automobile sales. The results using the empirical models suggest that these indicators have a significant effect on automobile sales.Wang et al. (2011) used an adaptive network-based fuzzy inference system to estimate new automobile sales in Taiwan with economic indicators.

3 Data

The time series output quantity considered in this paper is domestic passenger vehicle sales (which includes utility vehicles and vans) in India from Jan 2016 to November 2020 [11] (see Fig. 1). The input parameters in the forecast model are the macroeconomic indicators (ref Table.1): Gross Domestic Product (represented as growth percentage in year – on – year basis) [12], Unemployment rate (represented as growth percent in year – on – year) [13], Consumer Price Index which represents the inflation [14], Private Final Consumption Expenditure(expressed in growth percentage year – on – year), Index of Industrial Production [15], Prime Lending Rate (expressed in growth percentage year – on – year)and Crude Oil Prices (Rs/barrel) [16]. The indicators which affect sales quantity exhibit dynamic influence on demand behavior and are able to represent national economy and changes in the economic cycle. Based on the survey of literatures and expert opinion, the economic indicators selected for the study (see Fig. 2) are: Unemployment rate, GDP growth rate, Private Final Consumption Expenditure (PFCE) growth rate, Index of Industrial Production (IIP), Bank Lending Rate (BLR), Consumer Price Index (CPI) and Crude oil prices [17,18,19,20].It is seen that the domestic retail sales of passenger vehicles (Fig. 1.) exhibits a seasonal behavioral pattern with ups in the months of March and September and dips in June and Decemberfor the years 2016, 2017, 2018 and 2019. However, during April 2020, the lowest dip in sales was observed due to lockdown imposed in India. The sales of passenger vehicles duringthe month was recorded as zero.No significant trend has been noticed in the dataset. GDP growth rate (Fig. 2a) shows a slow decline in the trend over the period of study and recorded negative growth during the first half of 2020. Consumer Price Index (Fig 2b.) which is a true representation of inflation shows an upward trend until the last quarter of 2019. Revisions in monetary policies and strict measures adopted by the Reserve Bank of India in December 2019 has resulted in a slight decrease of inflation during January 2020. Unemployment rate (Fig 2c.) experienced a slowdown from first quarter of 2016 to the second quarter of 2017 and has been increasing hence forth. Private Final Consumption Expenditure which represents the customer spending on goods and services (Fig 2d.) has shown a downward trend indicating the selective spending and reluctance to spend money on purchases of non-essential goods. Index of Industrial Production . It has however started to decline from the first quarter of 2020 due to changes in regulations by OPEC and the growing impact of the pandemic COVID-19. The data set for the input and output parameters have been collected from various reliable secondary sources published by organizations under the Govt of India (ref Table. 1) from the period January 2016 to November 2020. The data has been collected to show the influence of BS IV emissionstandard regulations put forward during 2017 and BS VI regulations from January 2020 onwards.

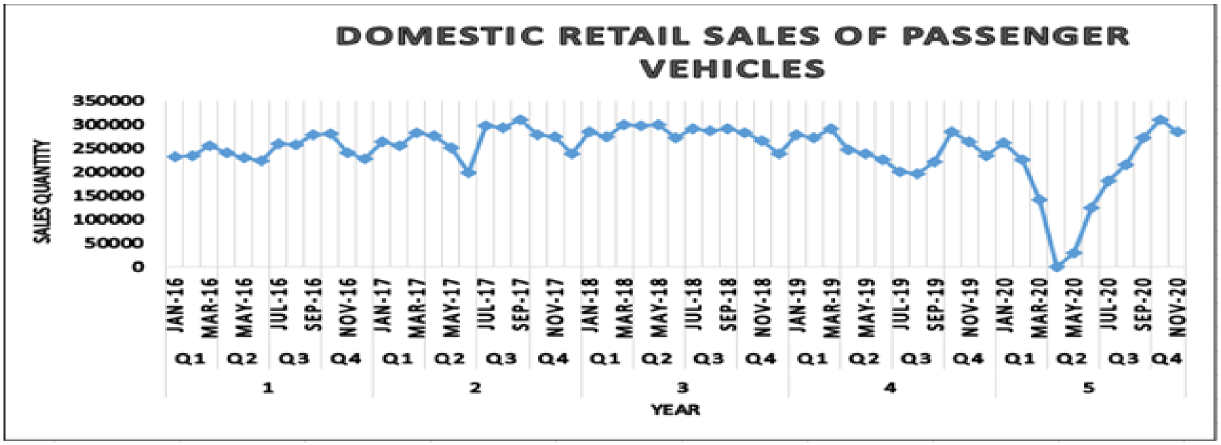
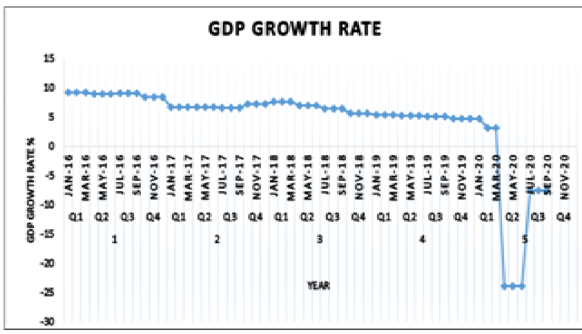


Fig. 1. Domestic retail sales of Passenger Vehicles



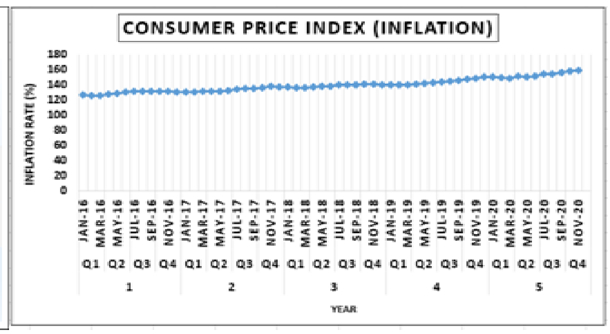
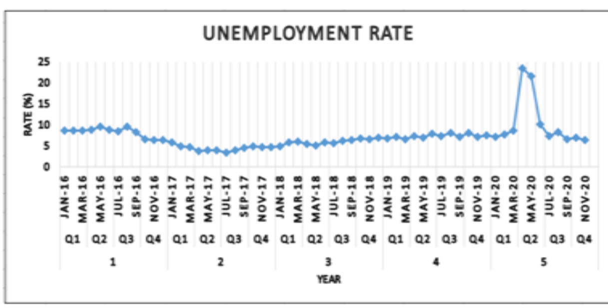


Fig. 2a.GDP Growth Rate **Fig. 2b.**Consumer Price Index (Inflation)



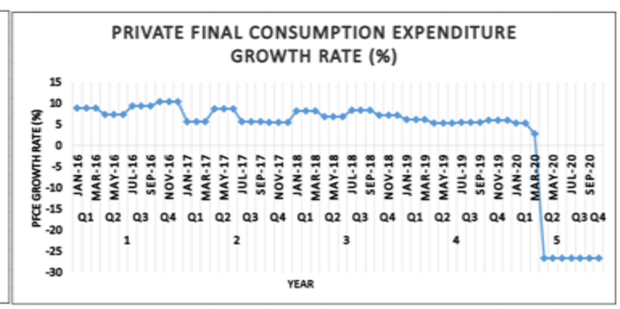


Fig. 2c. Unemployment Rate **Fig. 2d.**Private Final Consumption Expenditure

4 Forecasting Model

Only very few forecasting models have been applied to predict automobile sales which makes use of macroeconomic indicators. Time Delay Artificial Neural Network with 7 input parameters has been usedin this paper to estimate the forecasted value for sales and hence the measures of forecasting accuracy – Mean Squared Error, Root Mean Square Error, Mean Absolute Deviation and Mean Absolute Percentage Error.

4.1 Time Delay Feed Forward Neural Networks with Back Propagation

Artificial Neural Networks have been used in a wide range of applications where nonlinear data is involved. It provides a flexible computation framework for nonlinear modeling. ANN modeling does not require any apriori assumptions. The feed forward architecture with back propagation algorithms are the most widely used network structure of Artificial Neural Networks particularly in forecasting. There are multiple layers of interconnected neurons, input, hidden and output layers. Each neuron in one layer is connected with all neurons of the immediate next layer. The links between neurons are known as arcs which are characterized by arc weights (w_{ij}). The threshold coefficient of $i$thneuron (D.Svozil, 1997) is characterized by the equation:

$x_i = f(\xi_i),$ (4.1)

$\xi_i = \vartheta_i + \sum_{j\epsilon\Gamma^{-1}} w_{ij} x_j$(4.2)

Where,Γ is the mapping function which assigns to each neuron i a subset $\Gamma_i \subseteq V$ which consists of all ancestors of the given specified neuron. All predecessors of the given neuron i is represented by the subset $\Gamma_i^{-1} \subseteq V$. ξ_i is the potential of the ith neuron and $f(\xi_i)$is the transfer function and x_j is the bias.

The transfer function$f(\xi_i)$= $\frac{1}{1+\exp(-\xi)}$ (4.3)

The generalized equation for the neural network is shown:

$$\hat{y}_2 = w_0 + \sum_{j=1}^{H} w_j f(\xi_i)\left(x_j + \sum_{i=1}^{N} w_{ij} y_{t-i}\right) + e_t \qquad (4.4)$$

Where $\hat{y}_2$ is the estimated value of Passenger Vehicle Retail Sales using the ANN model, w_{ij}, w_j are the model weights, H and N are the number of hidden and input nodes respectively, e_t is the noise or the error term and $f(\xi_i)$is the transfer function of the hidden layer neurons.

The threshold coefficient and weight coefficients are varied to minimize the mean squared error, root mean squared error, mean absolute deviation and mean absolute percentage error values between the computed and required output values.

Matlab R2017a has been used to investigate the neural network model in the present study after training, testing and validation on the dataset. The following set of procedures have been used for the purpose:

1 The arc weights are assigned randomly.
2 The minimum test error is initialized to the maximum real value
3 Training dataset is passed to the network iteratively
4 Back propagation is performed using mean squared error as the stop criteria for learning
5 Testing is performed using the test dataset and measures of performance of learning and testing dataset are computed

6 Validation is performed where the predicted values are printed to the output file which contains actual prediction and corresponding test error
7 If the error is less than desirable value, the values of arc weights are saved. Else the network is retrained for new values of arc weights
8 Network evaluation is performed to calculate RMSE, MAD and MAPE.

5. Results and Discussion

5.1Time Delay Feed Forward Neural Networks with back propagation (TDNN)

The neural network model for the study is as shown below:

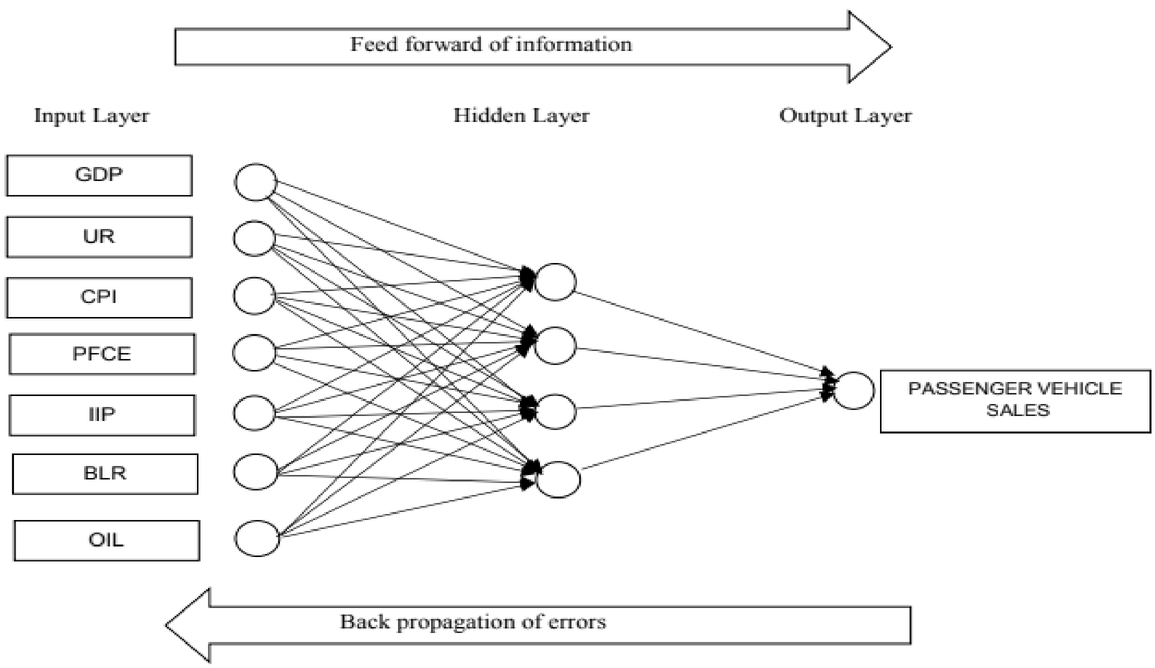
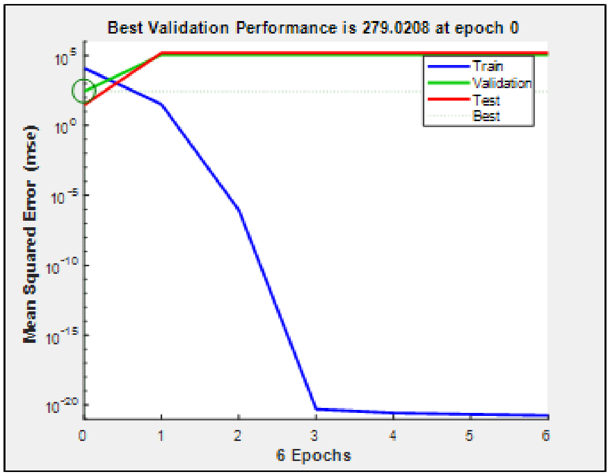


Fig 3. General framework of Feed forward Back Propagation Neural Network

The neural network is modeled with the 7 input parameters and yielded the following results(Table 5). All the models were trained using TRAINLM(Levenberg-Marquardt algorithm) as the training function, LEARNGDM as the learning function, TANSIG (sigmoid function) as the Transfer function for hidden layer, PURELIN (linear function) as the transfer function for output layer and MSE as the performance measure.Since a concrete rule was not available in selecting the number of hidden neurons, a trial and error method was adopted to range the number of hidden neurons.



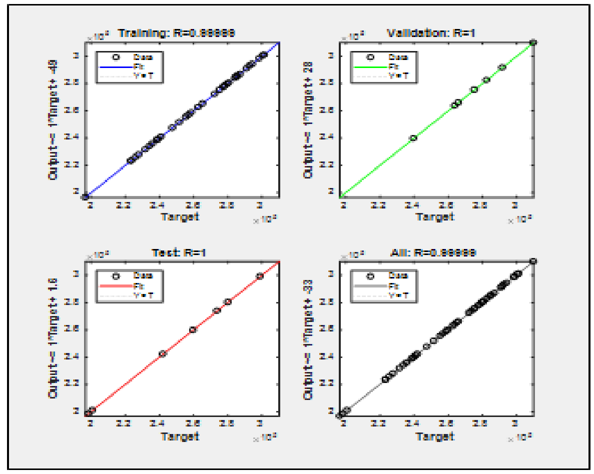


Fig 4. Validation performance of network 7-13-1 **Fig 4.** Validation performance of network 7-13-1

The values for accuracy measures obtained for the TDNN is shown in TABLE2.

TABLE 2
Accuracy measures for Time Delay Neural Network

	TDNN	
	All factors	Significant factors
MSE	19319.35	145629.0
RMSE	138.99	381.61
MAD	27.35	159.98
MAPE	0.0119	0.0625

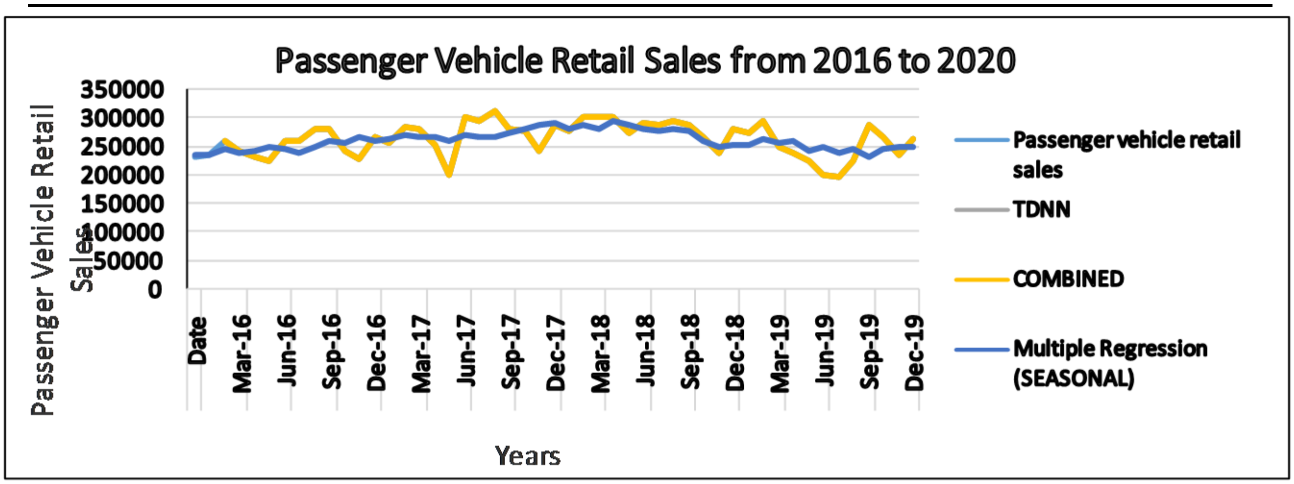


Fig 6. Passenger Vehicle Retail Sales modeled using TDNN

It is observed that the TDNN with 7 input parameters, 1 hidden layer, 14 hidden neurons and 1 output neuron (7-14-1) was able to model the output quantity with 99% accuracy.

6. Conclusion

The automobile industry is a major contributor to India's GDP, of which the major share is occupied by passenger vehicles. It is therefore critical to analyze the business plans of the industry carefully and to foresee the upturns and downturns in the economy which might have a direct effect on the country economic growth. The business plans thus developed carefully should be effective, realistic and reliable. Thus, it becomes essential to make use of a forecasting method which can predict the business conditions with a very high level of accuracy. Retail chains, being very close to the customer need to respond quickly with variations in demand level.

This study attempts to investigate the effects of macroeconomic indicators such as GDP, UR, CPI, PFCE, IIP, PLR and OIL on the domestic sales of passenger vehicles in India. The empirical results obtained indicate a long-term effect of macroeconomic indicators on the sales over the sample period from 2016 to 2020. Use of new generation computational techniques like ANN in prediction has greatly improved the accuracy of results by capturing underlying relationships among nonlinear variables which otherwise are difficult to be determined. A time delay neural network (TDNN) with feedforward, back propagation was used in the paper for modeling purposes. The number of input neurons were fixed at 7 (for 7 macroeconomic indicators). Since no concrete rule is available in fixing the number of hidden neurons of the hidden layer, the optimal number was found by an extensive iterative process.Results indicated that TDNN was able to model the parameters effectively.

The general model thus formulated can be effectively applied to forecast the sales of various segments in the automotive industry. Future research directions in this area can be oriented towards incorporating more efficient algorithms such as particle swarm optimization (PSO) to decrease the effect of local minima and use of microeconomic and internal factors to the organization to capture all the underlying relationships between these variables for modeling the system completely.

REFERENCES

[1] Malini Goyal, 'With India's economy growing at about 7%, why the auto industry is hurting so badly?', Economic Times, 2019

[2] Jayabharat Maruti Limited, Annual report 2017-18

[3] Sa-ngasoongsong et.al, Multi-step sales forecasting in automotive industry based on structural relationship identification, International Journal of Production Economics, 140(2) (2012) 875-887

[4] Monks, J.G. Operations management: Theory and problems, 2ed.McGraw-Hill, New York, USA (1987)

[5] Chu, C.-W., Zhang G.P., A comparative study of linear and nonlinear models for aggregate retail sales forecasting. International Journal of Production Economics 86 (2003) 217–231.
[6] Danese P., Kalchschmidt M., The role of the forecasting process in improving forecast accuracy and operational performance, International Journal of Production Economics 131 (2011) 204–214.
[7] Luxhøj J.T., Riis J.O., Stensballe B., A hybrid econometric - neural network modeling approach for sales forecasting. International Journal of Production Economics 43 (1996) 175–192.
[8] Mentzer, J.T., Moon M.A., Sales forecasting management: A demand management approach, 2ed.SagePublication, Inc, Thousand Oaks, California (2005).
[9] Wilson, R.L. and Sharda R, Bankruptcy Prediction Using Neural Networks. Decision Support Systems 11 (1994) 545-557.
[10] White, H, Learning in artificial neural networks: A statistical perspective. Neural Computation 1(4) (1989) 425–464.
[11] Ripley, Neural Networks: A Review from Statistical Perspective, Statistical Science 9 (1) (1994) 45-48.
[12] Cheng, Titterington, Neural Networks: A Review from a Statistical Perspective 9(1) (1994) 2-30.
[13] Yang, Li, The Combination Forecasting Model of Auto Sales Based on Seasonal Index and RBF Neural Network, International Journal of Database Theory and Application 9(1) (2016) 67-76.
[14] Chow. G, Demand for Automobiles in the United States. Amsterdam: North-Holland Publishing Company (1957).
[15] Nerlove, A Note on Long-Run Automobile Demand, The Journal of Marketing 22(1) (1957) 57-64.
[16] Suits, D. B., The demand for new automobiles in the United States 1926-1956, Review of Economics and Statistics 40(3) (1958) 273-280.
[17] Lewandowski. R, Prognose- und Informationssysteme und ihreAnwendungen. de Gruyter, Berlin – New York (1974)
[18] Tanner J. C, Car Ownership Trends and Forecasts, Report LR 799, Transport and Road Research Laboratory (1977)
[19] Berkovec. J, Forecasting automobile demand using disaggregate choice models, Trans, Res 19B (1984)
[20] Mannering F. and Winston C, U.S. automobile market demand, Chapter 3 in Clifford Winston and Associates. Blind Intersection. Washington D.C.: The Brookings Institution (1987).
[21] Garciâ -ferrer et.al, Univariate Forecasting Comparisons: The Case of the Spanish Automobile Industry, Journal of Forecasting 16 (1997) 1-17.

[22] Peter Romilly, Haiyan Song and Xiaming Liu, Journal of Transport Economics and Policy 32 (2) (198) 165-185.