Journal of Mathematics and Statistics Studies
ISSN: 2709-4200
DOI: 10.32996/jmss
Journal Homepage: www.al-kindipublisher.com/index.php/jmss

JMSS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Retail Demand Forecasting Using Neural Networks and Macroeconomic Variables

Md Sabbirul Haque
Centennial, CO, The United States of America
Corresponding Author: Md Sabbirul Haque, **E-mail**: sabbir465@gmail.com

| ABSTRACT

With the growing competition among firms in the globalized corporate environment and considering the complexity of demand forecasting approaches, there has been a large literature on retail demand forecasting utilizing various approaches. However, the current literature largely relies on micro variables as inputs, thereby ignoring the influence of macroeconomic conditions on households' demand for retail products. In this study, I incorporate external macroeconomic variables such as Consumer Price Index (CPI), Consumer Sentiment Index (ICS), and unemployment rate along with time series data of retail products' sales to train a Long Short-Term Memory (LSTM) model for predicting future demand. The inclusion of macroeconomic conditions in the predictive model provides greater explanatory power. As anticipated, the developed model, including this external macroeconomic information, outperforms the model developed without this macroeconomic information, thereby demonstrating strong potential for industry application with improved forecasting capability.

| KEYWORDS

Demand forecasting, Neural Networks, Long Short-Term Memory, macroeconomic variable, economic environment, explainable Artificial Intelligence

| ARTICLE INFORMATION

ACCEPTED: 02 July 2023 **PUBLISHED:** 12 July 2023 **DOI:** 10.32996/jmss.2023.4.3.1

1. Introduction

Accurate demand forecasting is crucial for firms as it directly affects firms' financial performance. Future demand forecasting helps firms adjust their operations with market demand and gain a competitive edge in the highly globalized market environment. An increasing number of firms are now moving towards advanced prediction models in an effort to improve supply chain management. The literature, therefore, responds by triggering a large number of studies focusing on forecasting future demand. The literature has leveraged numerous statistical as well as Artificial Intelligence (AI) and Machine Learning (ML) approaches, including various advanced Neural Networks based algorithms. The current literature relies on time series data of customer demand for retail products of interest along with some other relevant variables, such as product price, store characteristics, any special offers or deals, information on any special day or events, etc., in developing forecasting models. However, firms operate in a macroeconomic environment that has a great influence on households' spending. The current literature fails to leverage this influence in forecasting future customer demand. Empirical evidence suggests that economic conditions significantly affect households' shopping behavior and household spending [Scholdra et al., 2022]. Therefore, incorporating those macroeconomic conditions and external environmental factors in addition to capturing historical customer demand data is crucial for customer demand forecasting. In my present study, I address this concern and enrich time series data of customer demand by incorporating macroeconomic variables, such as CPI, ICS, and the unemployment rate, as a reflection of the economic environment. I leverage this enriched data to develop a multi-layer Deep Neural Networks model for forecasting future demand. Specifically, I implement an LSTM model which is capable of capturing unobserved non-linear trends from observed data with the help of hidden layers and by maintaining a memory of past information. In my study, I achieve a reasonably low prediction error with a Root Mean Square Error (RMSE) of 1.60. Furthermore, the model developed including these macroeconomic conditions outperforms the model developed without this macroeconomic information. I further extend my study by providing an interpretation of the developed model with the help of shape values and by providing empirical evidence that the economic environment has a strong influence in forecasting future demand. This study clearly sheds light on the way to further improve firms' forecasting capability of future demand for products which has direct consequences on their financial performance.

2. Literature Review

Because of increasing interest and desire for firms to make better supply chain decisions, a vast literature focuses on improving prediction accuracy with minimal prediction error, relying on various approaches; prediction approaches implemented in the literature include statistical techniques, such as Auto-Regressive (AR), Auto-Regressive Integrated Moving Averages (ARIMA), Seasonal ARIMA (SARIMA), Vector Auto-Regressive (VAR) models, etc. as well as advanced AI algorithms using Artificial Neural Networks. More recent literature tends to rely on various hybrid prediction models that combine several approaches in generating predictions. Each approach has its own limitations and advantages and is suitable for specific applications.

Early studies on demand forecasting are mostly based on statistical approaches, such as MA, ARIMA, SARIMA, VAR, etc. Those statistical approaches require stationarity and rely on the assumption that relations among variables are linear in nature. Fattah et al., 2018 implement various ARIMA models and select an appropriate model using Box-Jenkins time series procedure to forecast demand in a food company using time series data. Results from this study provide managers with some guidelines regarding supply-chain management of the products. Ediger and Akar, 2007 use ARIMA and SARIMA methods to estimate the future primary energy demand of Turkey from 2005 to 2020.

While these statistical approaches are supposed to provide reliable results in the case of linear relationships, most real-world time series data display non-linear relationships, in which case the assumption of linear relationships is violated. Computational intelligence methods, such as Artificial Neural Networks, and different tree-based methods, such as Random Forests, Support Vector Machines (SVM), etc., overcome this limitation and can capture non-linear relations among variables. Several studies compare the performances of statistical approaches with those of computational intelligence methods. Empirical evidence from those studies suggests that models developed using Neural Networks algorithms generally exhibit superior performance. For instance, Wang et al., 2021 implement ARIMA and LSTM to predict future order demand, and they find evidence that LSTM exhibits enhanced performance in the case of short-term forecasting. Prybutok and Mitchel, 1999 compare the performance of ANN with that of Box-Jenkins ARIMA as well as regression models by developing models for forecasting daily maximum ozone levels. Findings from this study also suggest that the Neural Networks model outperforms regression and ARIMA models. Mitrea and Wu, 2009 compare several statistical methods, such as MA and ARIMA, with Neural Networks models using data from an inventory database of Panasonic Refrigeration Devices Company located in Singapore. Findings from their study also support findings from similar studies that Neural Networks-based models provide superior performance compared to other statistical methods.

Neural Networks algorithms come with certain advantages. These models are non-parametric and data-driven, so they can capture any non-linear relationships without prior specification of relationships among variables in the model. This makes these algorithms more suitable for real-world data containing many complex non-linear relationships. Considering this flexibility and superior performance, recent literature mostly relies on Neural Networks-based models for forecasting applications. Tanizaki et al., 2018 use various Machine Learning models to predict restaurant demand. Yue-Fang Gao et al., 2009 develops a demand forecasting model using Holt-Winter's Neural Networks model for the retail industry. Palkar et al., 2020 develop various models for forecasting future demand for liquor products and find evidence of considerably improved accuracy of the Neural Networks-based LSTM model compared to other models. Chen et al. (2021) developed a Neural Networks-based model to forecast future retail demand using data provided by Walmart and offered an interpretation of the developed model using Shapely values. There are also several other studies that develop models in an effort to forecast future demand using various algorithms [Chen et al., 2021, Chung et al., 2023]. A few studies provide systematic reviews of related articles in the literature [Aamer et al., 2021, Lee et al., 2022].

While significant efforts have been made to develop forecasting models with improved performance, the literature fails to utilize the explanatory power of economic conditions in demand forecasting applications. Documented evidence sheds light on the influence of economic conditions on household demands (Scholdra et al., 2022). In my present study, I fill this gap and enrich the historical data of retail demand by including macroeconomic information. I leverage this enriched data to develop an LSTM model to forecast future demand. Specifically, I feed CPI, ICS, and unemployment rate along with time series data of customer demand to train the model. A large number of empirical studies have documented the influence of these selected macroeconomic variables in explaining household demand. For example, Aderamo (2010) shows that CPI, along with other factors, is important in explaining the demand for air transport in Nigeria. Similar evidence has been documented in Obaid (2020) and Icoz and Kozaka, 1998. Findings from Huth (1994) suggest that ICS is useful in predicting future consumer spending. Similarly, Alegre (2013) also documents the explanatory power of unemployment on households' spending on tourism. Results from my study clearly justify the inclusion of information on external economic conditions as a good signal in the model. In my present study, I demonstrate that the RMSE value is smaller when I include and feed those external macroeconomic variables into the model compared to that excluding those variables.

3. Methodology

3.1 Data Preprocessing and Feature Engineering

Data used in this study are historical data of products sold for 3,049 items from ten different stores from three states (CA, TX and WI) over five years and are provided by Walmart, USA and. Each store in the three states has the same set of 3,049 products. In addition to historical data on the product sold, the dataset also includes product id, product price, product category, department, store information, any promotions, day of the week, and any events on the day of the product sold. This information is contained in three different datasets: historical sale data, calendar data, and product price data. The historical sale dataset is a wide dataset that contains information on products sold for 3,049 items over the past 1913 days, with a column for each unique day and a row for each unique product. This dataset also contains product id, product category, and store location. The calendar dataset contains information about special events or holidays, promotion, event type, etc., on each day in each state. The product price dataset contains the price of each product on each day at each store. Macroeconomic variables considered in this study include CPI, ICS, and unemployment rate. Historical CPI data and unemployment data are collected from World Bank's World Development Indicators (WDI) database, and historical ICS data is collected from the University of Michigan's website. First, I join the calendar dataset with macroeconomic data on the "date" column so that I get a combined dataset with a row for each day. This combined dataset is then joined with the product price dataset and sale dataset. This final dataset has product sale, price, and promotion information for each of the 3049 products from each store from three states for each of 1913 days. I restrict the dataset to the most recent 600 days in order to avoid any old trends that may not be valid for current days. Furthermore, I include rolling averages as well as rolling standard deviation of several lag values of sale prices as part of input features.

3.2 Neural Networks Algorithm

In this study, I use the LSTM Neural Networks algorithm to train the forecasting model. Neural Networks, also known as Artificial Neural Networks, try to mimic the human brain and are comprised of a network of nodes with several layers, such as input layers, hidden layers, and output layers. Every node in the input and hidden layers relates to every node in the next layer with a weight. A series of nodes connected to each other are able to recognize complex, non-linear, and unobserved relationships hidden in the data. Neural Networks learn from the training data in a similar way that the human brain learns for the first time.

Each node can be thought of as a linear regression with associated weights. Initial weights are assigned at the beginning of training. After each cycle of forward and backward propagation, these weights are adjusted by minimizing the cost function. The speed of adjustment is determined by a parameter called the learning rate. There are several types of Neural Networks algorithm, and LSTM is one of them. LSTM has a memory that has a dependency on observations over time and is, therefore, appropriate for time-series applications. In my study, I use a multilayer LSTM neural network algorithm to train the model.

4. Results

Before I move on to developing a neural network model, I first examine data in the time domain to understand any temporal dependencies. One of the challenges in this analysis is the fact that most individual time series for the products in this dataset are sparse, and it is difficult to examine any temporal dependency from this data. Therefore, I examine temporal dependency using aggregate data, i.e., the total sale of all products from all ten stores from three states instead of sales of individual products.

Figures 1 and 2 depict autocorrelation and partial autocorrelation, respectively. The autocorrelation estimates (Figure 1) suggest that the current total sale is correlated with its lagged values of up to more than 250 days. The autocorrelation coefficient captures both the direct and indirect effect of lagged values on the current value, whereas partial autocorrelation captures only the direct effect, thereby removing any indirect effects. Partial autocorrelation estimates (Figure 2) also provide strong evidence for serial dependency and demonstrate that current total sale is directly correlated with lagged values with lags of up to 40 days. The evidence of strong temporal dependency with large lagged-values is indicative that the time series in this dataset has explanatory power in forecasting future values. The low RMSE value of the developed LSTM model in this study presented in the next subsection is consistent with this finding of strong serial correlation in the dataset.

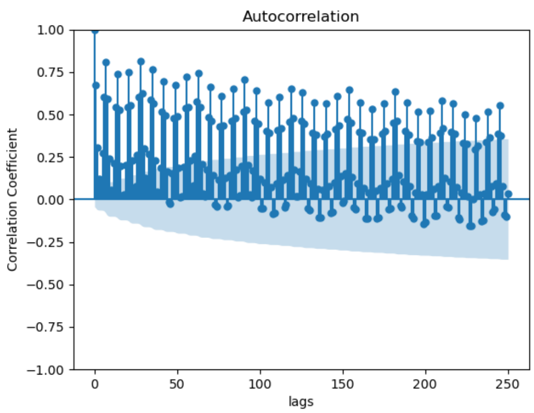

Figure 1: Autocorrelation of total daily sale

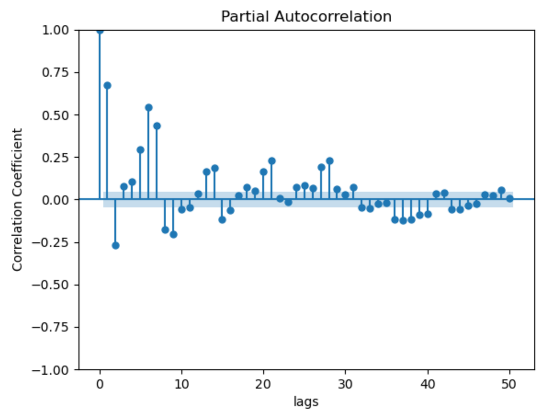

Figure 2: Partial Autocorrelation of daily sale

Table 1: Comparative Model Performances (CA)

Model details	RMSE	
	Without Macro Variables	With Macro variables
LSTM without products' prices	2.010	1.990
LSTM with products' prices	1.964	1.959

Table 2: Comparative Model Performances (CA, TX and WI)

Model details	RMSE	
	Without Macro Variables	With Macro variables
LSTM with products' prices	1.602	1.599

Given the evidence of serial correlation in the time series data, I apply the LSTM model to this data. Results presented from each version of the model are based on the same architecture of the LSTM model, which consists of two hidden layers and is designed for forecasting the demand of 28 future time steps. Initially, I restricted the model to a store CA only, and then I extended the model by including all three states. In order to avoid any past trends that are not present in current data that may result in undesirable results, I feed time series data of the last 600 days to train the model. The basic model does not include product prices, and then I further extend the model by including product price information. To compare the proposed model with that in the present literature, I present model performances, including macroeconomic variables with that without macro variables.

Detailed performances of the initial models are presented in Table 1; In the basic models that do not have product prices, the mol, when including external macroeconomic variables, exhibits superior performance with an RMSE value of 1.990, whereas the model without macro variables has a larger RMSE value of 2.010. A similar conclusion can be reached when I include product prices in the model. The LSTM model with macro variables still outperforms the conventional model without macroeconomic variables. I further extend the model by including data from all three states. Results from this extended model are presented in Table 2. Results from this extended model are similar to those from the basic models. The model with economic information outperforms models without economic environmental information. These findings are consistent with the findings in the literature that provide evidence that the external economic environment can explain household demand. In all cases, the forecasting model developed, including those external variables that can explain household demand, outperforms the model developed without those external variables.

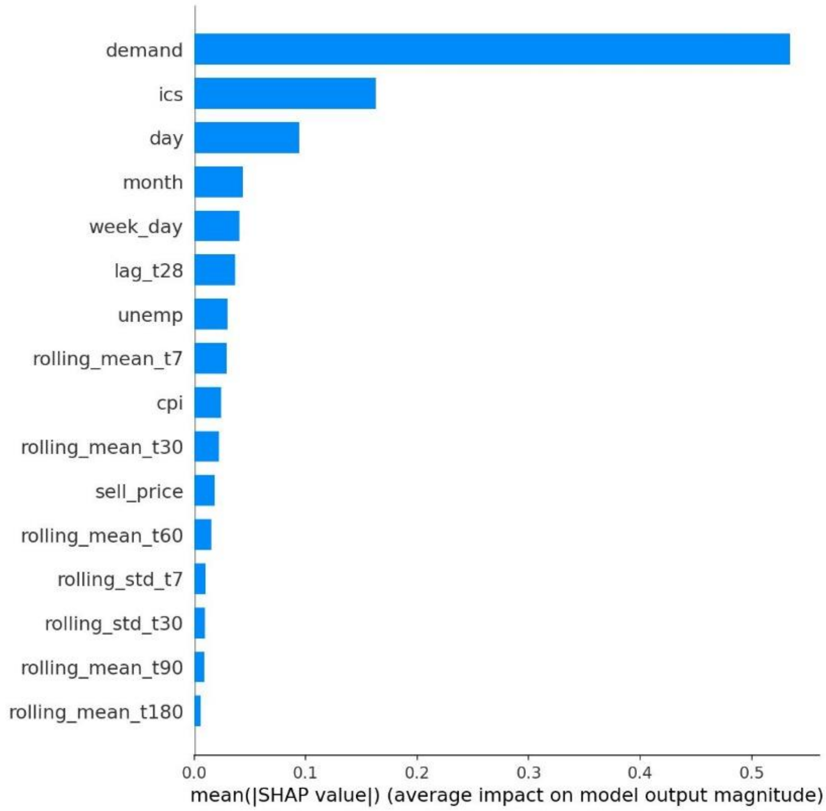

Figure 3: Shapely Value (Feature Importance)

In a time series forecasting problem, stakeholders are not only interested in forecasting future values, but at the same time, they are also interested in identifying the features that are important in forecasting future values. Although, in general, Deep Learning models provide superior prediction performance compared to conventional statistical approaches, these Deep Learning models are considered to be black boxes that do not provide any interpretation of how the predictions are made. I use Shapely values to interpret the LSTM model and examine the contribution of each macroeconomic feature along with other features in forecasting future demand. Shapely values utilize the Collaborative Game theory approach to provide desirable properties and are widely used in the literature for explaining computational intelligence models [Chen et al., 2021]. I explain the importance of the proposed economic and environmental variables compared to other features in the model by using Shapely values. The bar plot of Shapely values (Figure 3) confirms that these macroeconomic variables are, in fact, important and contribute to the model output. Specifically, ICS, unemployment rate, and CPI are the second, seventh, and ninth important features, respectively, out of the sixteen primary features included in the model.

5. Conclusion

In my present study, I developed an LSTM model and included macroeconomic variables along with other demand-related features for forecasting retail demand with improved accuracy. Key findings from this study demonstrate that including macroeconomic variables in the model improves performance. I explore different samples with varying features. Results are robust and provide evidence for superior performance for including proposed features. I also examine the importance of these proposed variables compared to other relevant features and find that these proposed features significantly contribute to the model outcome. My study contributes to the existing literature by providing a solution for an improved forecasting capability, thereby creating opportunities for firms for better supply chain management and improved financial performance. While findings from this study are based on a Neural Networks model, it is still unknown whether these findings are generalizable to other relevant machine learning and statistical approaches. Future studies can focus on examining the applicability of these findings to a broader set of relevant methods and approaches.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Aderamo A. J. (2010). Demand for Air Transport in Nigeria, *J Economics* 1 (1), 23-31

[2] Aamer, A., Eka Yani, L., & Alan Priyatna, I. (2021). Data analytics in the supply chain management: Review of machine learning applications in demand forecasting. Operations and Supply Chain Management: *An International Journal, 14*(1), 1–13

[3] Chen J., Koju W., Xu S., and Liu Z. (2021). Sales Forecasting Using Deep Neural Network and SHAP Techniques. 2021 IEEE 2nd International Conference on Big Data, *Artificial Intelligence and Internet of Things Engineering (ICBAIE),* 135–138

[4] Chung D., Lee C. G., Yang S. (2023). A hybrid machine learning model for demand forecasting: Combination of K-Means, Elastic-Net and Gaussian Process Regression. *International Journal of Intelligent Systems and Applications in Engineering, ISSN* 2147-6799

[5] Dou Z., Sun Y., Zhang Y., Wang T., Wu C., and Fan S. (2021). Regional Manufacturing Industry Demand Forecasting: A Deep Learning Approach. *Applied Sciences, 11*, 6199

[6] Ediger, V. S. and Akar S. (2007). ARIMA forecasting of primary energy demand by fuel in Turkey. Energy Policy, 35 (3), 1701-1708

[7] Fattah J., Ezzine L., Aman Z., Moussami H. E., and Lachhab A. (2018). Forecasting of demand using ARIMA model. *Engineering Business Management*, 10, 1-9

[8] Gao Y-F, Liang Y-S, Liu Y, Zhan S-B and Ou Z-W. (2009). A neural-network-based forecasting algorithm for retail industry. *Proceedings of the eighth international conference on machine learning and Cybernetics, Baoding*

[9] Huth W. L., Eppright D. R., and Taube P. M. (1994). The Indexes of Customer Sentiment and Confidence: Leading or Misleading Guide to Future Buyer Behavior. *Journal of Business Research, 29 (3), 199-206*

[10] Icoz, O., T. Var, and M. Kozaka (1998). Tourism Demand in Turkey. *Annals of Tourism Research, 25,* 236-240.

[11] Kandananond K. A (2012). A comparison of various forecasting methods for autocorrelated time series. *International Journal of Engineering Business Management, 4,* 4

[12] Lee I., Mangalaraj G. (2022). Big data analytics in supply chain management: A systematic literature review and research directions. *Big Data and Cognitive Computing, 6,* 17

[13] Mitrea CA, Lee CKM, and Wu Z. (2009). A comparison between neural networks and traditional forecasting methods: a case study. *Int J Eng Bus Manage*, 1, 11

[14] Obaid, H.M.A. (2020). Factors determining housing demand in Saudi Arabia. *International Journal Economics and Financial Issues, 10,* 150–157

[15] Palkar A., Deshpande M., Kalekar S., and Jaswal S. (2020). Demand Forecasting in Retail Industry for Liquor Consumption using LSTM. Proceedings of the International Conference on Electronics and Sustainable Communication Systems, IEEE Xplore Part Number: CFP20V66-ART

[16] Prybutok VR, Yi J, and Mitchell D. (2000). Comparison of neural network models with ARIMA and regression models for prediction of Houston's daily maximum ozone concentrations. *European Journal of Operational Research, 122*(1),31–40

[17] Scholdra T P., Wichmann J R.K., Eisenbeiss M, and Reinartz W J. (2022). Households Under Economic Change: How Micro-and Macroeconomic Conditions Shape Grocery Shopping Behavior. *Journal of Marketing, 86* (4), 95–117

[18] Tanizaki, T., Hoshino, T., Shimmura, T., and Takenaka, T. (2019). Demand forecasting in restaurants using machine learning and statistical analysis. *Procedia CIRP, 79*, 679–683

[19] Wang, C.-C., Chien, C.-H., Trappey, A. (2021). On the Application of ARIMA and LSTM to Predict Order Demand Based on Short Lead Time and On-Time Delivery Requirements. Processes, 9, 1157