JBMS

AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Product Demand Forecasting with Neural Networks and Macroeconomic Indicators: A Comparative Study among Product Categories

Tuan Ngoc Nguyen[1], ✉ Mahfuz Haider[2], Afjal Hossain Jisan[3], Md Azad Hossain Raju[4], Touhid Imam[5], Md Munsur Khan[6] and Abdullah Evna Jafar[7]

[1]VNDirect Securities, 97 Lo Duc, Hai Ba Trung, Hanoi, Vietnam

[2]Department of Clinical operations, University of Virginia Physicians Group, USA

[3]Department of Supply Chain & Information Systems, The Pennsylvania State University, University Park, Pennsylvania, USA

[45]Department of Computer Science, University of South Dakota, Vermillion, South Dakota, USA

[6]College of Graduate and Professional Studies, Trine University, Angola, IN, USA

[7]Department of Economics & Decision Sciences, University of South Dakota, Vermillion, SD, USA

Corresponding Author: Tuan Ngoc Nguyen, **E-mail**: tuan.nguyenngoc@vndirect.com.vn

| **ABSTRACT**

In the fiercely competitive global corporate arena, the intricacies of demand forecasting in the retail sector have become a focal point. While previous research has delved into various methodologies, it consistently overlooks the distinct performances of forecasting models within different retail product categories. Understanding these variations in prediction performances is pivotal, enabling firms to fine-tune forecasting models for each category. This study bridges this gap by scrutinizing the prediction performances of models tailored to different product categories. Building on recent research, we incorporate external macroeconomic indicators like the Consumer Price Index, Consumer Sentiment Index, and unemployment rate, alongside time series data of retail sales spanning various categories. This amalgamated dataset is employed to train a Long Short Term Memory model, projecting future demand across product categories. We further extend the analysis by identifying features that contribute most towards explaining product demand and quantifying their strength. The fitted models yield comprehensive insights into their performances and pinpoint the product categories warranting more focused model development.

1. Introduction

In a highly competitive market, accurate demand forecasting holds paramount importance for firms, directly influencing their financial performance. It enables firms to align their operations with market demand, granting them a competitive edge in the highly globalized market landscape. An increasing number of firms are turning to advanced prediction models to enhance their supply chain management. This trend has spurred a surge in studies dedicated to forecasting future demand. The literature has explored a wide array of statistical techniques as well as advanced Artificial Intelligence (AI) and Machine Learning (ML) approaches, including sophisticated Neural Networks-based algorithms. The literature predominantly relies on time series data of customer demand for specific retail products, coupled with relevant variables such as product price, store attributes, promotions, and information about special events or occasions, in developing forecasting models.

With the current development of forecasting capabilities, firms typically develop a general predictive model that is common for all types of product categories. However, it is important to recognize that different product categories have different patterns of

product demand. Therefore, developing a single forecasting model for all of the product categories together is not optimal for capturing all the underlying patterns across product categories. This necessitates building separate forecasting models for individual product categories.

In this study, we address this crucial concern by developing a separate forecasting model for each product category. We also enrich time series data of customer demand through the inclusion of macroeconomic variables, namely CPI, ICS, and the unemployment rate, as reflective indicators of the economic environment. This enriched dataset is employed to develop a multi-layer Deep Neural Networks model for forecasting future demand for each product category. Specifically, we implement LSTM models, which are capable of capturing latent non-linear trends from observed data through hidden layers and retaining a memory of past information. We further extend the analysis by examining the contribution of features using dominance analysis and quantifying the strength of features in explaining retail product demand. Our proposed approach yields reasonably low prediction errors for all models developed.

2. Literature Review

Due to the escalating need for companies to enhance supply chain decisions, a substantial body of literature has emerged, focusing on refining prediction accuracy through various methodologies. These encompass statistical techniques like Auto-Regressive (AR), Auto-Regressive Integrated Moving Averages (ARIMA), Seasonal ARIMA (SARIMA), and Vector Auto-Regressive (VAR) models alongside advanced AI algorithms utilizing Artificial Neural Networks. Recent studies have leaned towards hybrid prediction models, amalgamating diverse approaches to generate forecasts. Each approach has its own set of strengths and limitations, tailored for specific applications.

Initially, early studies in demand forecasting predominantly relied on statistical techniques such as MA, ARIMA, SARIMA, and VAR, which necessitate stationarity and assume linear relationships between variables. For instance, Fattah et al. (2018) applied various ARIMA models, utilizing the Box-Jenkins time series procedure to forecast demand in a food company. Ediger and Akar (2007) employed ARIMA and SARIMA methods to estimate Turkey's future primary energy demand from 2005 to 2020. However, these statistical approaches, designed for linear relationships, may fall short in real-world scenarios characterized by non-linear associations in time series data.

To address this, computational intelligence methods like Artificial Neural Networks, as well as tree-based techniques such as Random Forests and Support Vector Machines (SVM), emerged to overcome the limitations of linear assumptions. Empirical findings consistently highlight the superior performance of models developed using Neural Networks algorithms. For example, Wang et al. (2021) compared ARIMA and LSTM for predicting future order demand, with LSTM demonstrating enhanced performance in short-term forecasting. Prybutok and Mitchel (1999) compared the performance of Neural Networks with Box-Jenkins ARIMA and regression models for forecasting daily maximum ozone levels, favoring the Neural Networks model. Similarly, Mitrea and Wu (2009) compared various statistical methods with Neural Networks models, corroborating the superiority of the latter.

Neural Network algorithms possess distinct advantages; being non-parametric and data-driven, they can capture non-linear relationships without the need for prior specification. This makes them well-suited for real-world data characterized by complex non-linear associations. Given this adaptability and superior performance, recent literature predominantly gravitates towards Neural Networks-based models for forecasting applications. Studies by Tanizaki et al. (2018), Yue-Fang Gao et al. (2009), Palkar et al. (2020), and Chen et al. (2021) exemplify the widespread adoption of Machine Learning models, particularly Neural Networks-based approaches, in diverse forecasting contexts. Haque (2023) utilizes Neural Networks and takes into account macroeconomic variables to forecast retail demands.

Despite the strides made in developing models with improved forecasting accuracy, the literature has mostly ignored the importance of examining model performances individually across product categories. Furthermore, existing studies in the literature have failed to highlight feature dominance. In this study, we contribute to the literature by providing a comprehensive evaluation of the comparative performances of forecasting models for different categories, along with a detailed analysis of the comparative strengths of the features included in the models. Following Haque (2023) and Haque et al. (2023), we enrich historical retail demand data with macroeconomic information. This augmented dataset is leveraged to construct an LSTM model for projecting future demand for each category.

3. Methodology

3.1 Data

Data employed in this study is generously provided by Walmart. The original historical sales dataset is expansive, encapsulating information on the sale of 3049 items over a duration of 1,913 days. It is structured with a column for each unique day and a corresponding row for each unique product. This dataset also encompasses product IDs, product categories, and store locations. The calendar dataset furnishes insights into special events, holidays, promotions, event types, and related information for each day

across the respective states. Furthermore, the product pricing dataset contains records of the price of each individual product each day at every store.

3.2 Data Preprocessing

We restrict our analysis to 300 product items due to compute power limitations. Data employed in this study comprises the sales history of 300 distinct items of three different product categories across ten different stores situated in three states (CA, TX, and WI): 100 food items, 100 hobby items and 100 household items. Notably, each store within these three states stocks an identical set of 300 products. In addition to the records of products sold, the dataset encompasses essential attributes, including product ID, pricing details, product category, department, store specifications, promotional activities, day of the week, and any pertinent events coinciding with the sale of a product. This comprehensive information is distributed across three distinct datasets: historical sales data, calendar data, and product pricing data. We finally split this dataset into three different datasets, where each dataset represents a group of 100 items within a product category. Subsequently, we train a model for each product category comprising 100 items.

Incorporated into this study are key macroeconomic variables, namely the Consumer Price Index (CPI), Consumer Sentiment Index (ICS), and the unemployment rate. The historical CPI and unemployment data were sourced from the World Bank's World Development Indicators (WDI) database, while the historical ICS data was obtained from the University of Michigan's website.

To ensure comprehensive data integration, we merged the calendar dataset with the macroeconomic data based on the "date" column, culminating in a unified dataset, which now features a dedicated row for each day. This combined dataset was subsequently merged with the product pricing dataset and the sales dataset. This final data set encapsulates critical information regarding product sales, pricing, and promotional activities for each of the 300 products across each store in the three states. To maintain relevance and avoid incorporating outdated trends, we restricted the dataset to encompass the most recent 600 days. Additionally, we introduced rolling averages and rolling standard deviation calculations for several lag values of sale prices as integral components of the input features.

3.3 Neural Networks

In this investigation, we employ the LSTM Neural Networks algorithm to train a predictive model for each product category. Neural Networks also referred to as Artificial Neural Networks, seek to emulate the workings of the human brain. They consist of interconnected nodes arranged in layers, including input, hidden, and output layers. Each node in the input and hidden layers is linked to every node in the subsequent layer with a specific weight. This interconnected network of nodes has the capacity to discern intricate, non-linear, and concealed relationships within the data. Neural Networks learn from training data in a manner akin to how the human brain learns initially.

Conceptually, each node can be likened to a linear regression with associated weights. These initial weights are set at the outset of the training process. Following each cycle of forward and backward propagation, these weights are fine-tuned to minimize the cost function. The rate at which this adjustment occurs is dictated by a parameter known as the learning rate. Various types of Neural network algorithms exist, with LSTM being one of them. LSTM possesses a memory component that relies on past observations, rendering it well-suited for applications involving time-series data. In this study, we opt for a multi-layer LSTM neural network algorithm to facilitate the model training process.

4. Results and Discussion

As discussed in the previous sections, we implemented the LSTM model on three different datasets to train three different models. Furthermore, we compare the outcomes derived from a dataset containing macroeconomic indicators with those from a dataset devoid of such variables. The outcomes from each model are based on identical architecture, featuring several hidden layers designed to forecast demand for 28 future days. To ensure that any prior trends not present in current data do not lead to undesirable outcomes, we utilize time series data from the most recent 600 days for training.

Table 1: Comparative Model Performances

Model details	Without Macroeconomic Indicators		With Macroeconomic Indicators	
	RMSE (1)	MAE (2)	RMSE (3)	MAE (4)
LSTM for Hobby Items	1.61	2.03	1.51	2.01
LSTM for Household Items	1.57	1.66	1.39	1.62
LSTM for Food Items	1.25	1.09	1.25	1.08
LSTM for all Items together	2.59	2.61	2.57	2.60

Comprehensive results for model performances for three types of items are outlined in Table 1. Columns 1 and 2 represent performances obtained from the dataset that is devoid of macroeconomic indicators, whereas columns 3 and 4 are from a dataset that includes such variables. For the dataset that excludes such variables, the LSTM model trained on Hobby items has an RMSE of 1.61 and MAE of 2.03. The RMSE and MAE of the LSTM model trained on household items are 1.57 and 1.66, respectively. Similarly, the model trained on food items data has an RMSE of 1.25 and MAE of 1.09. Finally, we have also trained a legacy model that has been trained on items comprising all three categories together. The RMSE and MAE of the model trained on data from all three categories are 2.59 and 2.61, respectively. Moreover, concerning the dataset incorporating external variables, results align closely with those of the dataset omitting macroeconomic factors, albeit with slight indications suggesting potential enhancement in model performance.

In time series forecasting, stakeholders seek not only to predict future values but also to discern the contribution of each feature in explaining the outcome. While Deep Learning models typically outperform conventional statistical methods in prediction accuracy, they are often perceived as black boxes, lacking interpretability. We continue our examination by identifying the relevant features for the model building and quantifying relative strengths by using dominance analysis. Dominance analysis compares pairs of features across all possible subsets of the features in a model to determine the additional contribution that each feature makes to the model. A candidate feature is considered dominant when it makes a larger contribution in terms of R-square value to every possible subset of features than any of the other features. We use the Python Dominance Analysis package to implement it and find the strength of dominance. Results for the most dominant features are presented in Table 2. These results show that product price, 7 day rolling mean of sale, ics, cpi, day of week and unemployment rates are dominant predictors; results also report their strength of importance.

Table 2: Dominance Analysis Results

Dominant Feature Name	Percent Strength of Dominance
product price	38.02%
Previous 7 day rolling mean of sale	29.62%
ics	15.08%
cpi	9.59%
day of week	4.09%
unemployment rate	2.05%

A comparative picture of the performances of the models trained on different datasets highlights the importance of training separate models for individual categories of products. We also observe that the average RMSE and MAE are smaller when we train individual models than the RMSE and MAE of the model trained on the data set for all three product categories. Furthermore, the error is lowest for the model trained on food items. These results are intuitive as food items are less elastic and are less affected by external factors, and more predictable, whereas luxury items are more sensitive to market conditions and are less predictive.

5. Conclusion

In our current research, we advocate for developing separate tailored forecasting models for different product categories. In our study, we identify the most dominant features and their relative strengths and develop individual models for different product categories. In general, performance is superior for individual models when compared with a model trained using all product categories. The analysis also reveals that necessity items are more predictive and have lower prediction errors, whereas luxury items have higher prediction errors. Our study makes a noteworthy contribution to the existing body of literature by empirically demonstrating ways to improve model performances. The insights gleaned from this study present valuable opportunities for firms to refine their supply chain management practices and enhance overall financial performance. These results are based on an openly available data set available in Kaggle and examine three different product categories. Future studies can look into a larger set of product categories to make sure that the results are generalizable.

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]   Aamer, A., Eka Y., L., & Alan P., 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

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

[3]   Bristi, U., Rajib, M. I., & Shoukat, S. (2020). Implementation of corporate social responsibility in the ready-made garment industry of Bangladesh in order to achieve sustainable development goals: a study on Bangladeshi garment manufacturers. *Journal of Textile Science and Technology*, 6(4), 232-246.

[4]   Chung D., Lee C. G. and 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]   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

[6]   Chowdhury, O. S., & Baksh, A. A. (2017). IMPACT OF OIL SPILLAGE ON AGRICULTURAL PRODUCTION. *Journal of Nature Science & Sustainable Technology*, *11*(2).

[7]   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

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

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

[10] 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

[11] Haque, M. S. (2023). Retail Demand Forecasting Using Neural Networks and Macroeconomic Variables. *Journal of Mathematics and Statistics Studies,* 4(3), 01–06. https://doi.org/10.32996/jmss.2023.4.3.1

[12] Haque, M. S., Amin, M. S., Miah, J., Cao, D. M., Sayed, M. A., & Ahmed, S. (2023). Retail Demand Forecasting: A Comparative Study for Multivariate Time Series. *Journal of Mathematics and Statistics Studies*, 4(4), 40–46. https://doi.org/10.32996/jmss.2023.4.4.4

[13] 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

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

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

[16] Chowdhury, M. S., Nabi, N., Rana, M. N. U., Shaima, M., Esa, H., Mitra, A., ... & Naznin, R. (2024). Deep Learning Models for Stock Market Forecasting: A Comprehensive Comparative Analysis. *Journal of Business and Management Studies*, 6(2), 95-99.

[17] Linkon, A. A., Shaima, M., Sarker, M. S. U., Nabi, N., Rana, M. N. U., Ghosh, S. K., ... & Chowdhury, F. R. (2024). Advancements and Applications of Generative Artificial Intelligence and Large Language Models on Business Management: A Comprehensive Review. *Journal of Computer Science and Technology Studies*, 6(1), 225-232.

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

[19] Shaima, M., Nabi, N., Rana, M. N. U., Linkon, A. A., Sarker, M. S. U., Anjum, N., & Esa, H. (2024). Machine Learning Models for Predicting Corticosteroid Therapy Necessity in COVID-19 Patients: A Comparative Study. *Journal of Computer Science and Technology Studies*, 6(1), 217-224.

[20] Ray, R. K., Linkon, A. A., Bhuiyan, M. S., Jewel, R. M., Anjum, N., Ghosh, B. P., ... & Shaima, M. (2024). Transforming Breast Cancer Identification: An In-Depth Examination of Advanced Machine Learning Models Applied to Histopathological Images. *Journal of Computer Science and Technology Studies*, 6(1), 155-161.

[21] 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

[22] Nath, F., Asish, S., Debi, H. R., Chowdhury, M. O. S., Zamora, Z. J., & Muñoz, S. (2023, August). Predicting hydrocarbon production behavior in heterogeneous reservoir utilizing deep learning models. In *Unconventional Resources Technology Conference, 13–15 June 2023* (pp. 506-521). Unconventional Resources Technology Conference (URTeC).

[23] Nath, F., Chowdhury, M. O. S., & Rhaman, M. M. (2023). Navigating Produced Water Sustainability in the Oil and Gas Sector: A Critical Review of Reuse Challenges, Treatment Technologies, and Prospects Ahead. *Water*, *15*(23), 4088.

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

[25] 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

[26] 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

[27] 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

[28] Shoukat, S. (2016). Now and then: The neckline history of women. *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, *26*(2), 33-52.

[29] Shoukat, S., & Rabby, S. (2017). A Total Journey of Footwear with Material Analysis. *Journal of Scientificand Engineering Research*, *4*(6), 187-191.

[30] Shoukat, S., & Islam, M. A. (2016). Development of Wedding Costume from Bengal to Bangladesh.

[31] Shoukat, S., & Kwon, W. S. (2024, January). Effects of AI Agent Anthropomorphism on Consumers' Affective, Cognitive, and Social Shopping Experiences. In *International Textile and Apparel Association Annual Conference Proceedings* (80, 1). Iowa State University Digital Press.

[32] 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

[33] Wang, C.-C., Chien, C.-H. and 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

[34] Jewel, R. M., Linkon, A. A., Shaima, M., Sarker, M. S. U., Shahid, R., Nabi, N., ... & Hossain, M. J. (2024). Comparative Analysis of Machine Learning Models for Accurate Retail Sales Demand Forecasting. *Journal of Computer Science and Technology Studies*, *6*(1), 204-210.

[35] Nabi, N., Pabel, M. A. H., Rahman, M. A., Mozumder, M. A. S., Al-Imran, M., Sweet, M. M. R., ... & Sharif, M. K. (2024). Unleashing Deep Learning: Transforming E-commerce Profit Prediction with CNNs. *Journal of Business and Management Studies*, *6*(2), 126-131.