

Applying Machine Learning in Retail Demand Prediction

Subjects: [Operations Research & Management Science](#) | [Computer Science, Artificial Intelligence](#)

Contributor: Mehran Nasser , Taha Falatouri , Patrick Brandtner , Farzaneh Darbanian

In the realm of retail supply chain management, accurate forecasting is paramount for informed decision making, as it directly impacts business operations and profitability.

demand prediction

forecasting

deep learning

1. Introduction

Forecasting remains at the forefront of decision making ^{[1][2]}, especially in the field of supply chain management where accurate predictions of demand and inventory levels can have a significant impact on business operations and profitability ^{[3][4]}. Forecasting methodologies and algorithms to predict more accurate results have been developed throughout the ages ^[5]. In retail, rapid changes to the business environment, shorter planning horizons, lower profit margins, and customer service issues make forecasting more complex ^[6]. In addition, large numbers of product types compel businesses to adopt individual models for specific groups of products, covering linear to complex nonlinear patterns ^[7]. Global uncertainty and the plethora of complexities in the retail industry make forecasting accuracy one of the main priorities in this field.

Recently, researchers have utilized advanced methods of forecasting, such as deep neural networks, like long short-term memory (LSTM) or ensemble learning, to increase accuracy. Ensemble techniques combine different algorithms into an individual method, where each algorithm could be more sensitive under varying conditions ^[8]. Combined forecasts or “ensemble forecasts” have historical roots in aggregating individual forecasts and are not a recent development. However, using them in machine learning models has recently gained increasing popularity in the field of applied artificial intelligence in business. More than 15% of forecasting research in 2021 mentioned combined machine learning models ^[9]. Numerous combination and ensemble techniques have been developed to improve forecasting accuracy. Among the different machine learning models, tree-based methods dominate in both accuracy and uncertainty handling ^[10]. Compared to other machine learning algorithms, tree-based models have relatively low requirements for data preparation tasks, like feature scaling ^[11]. Extra tree regression (ETR), which is a type of tree-based ensemble model, has gained popularity in predictive research due to its ability to learn faster with smaller input dimensions ^[12]. Ensemble forecasting is also utilized in the retail sector to overcome uncertainties in data, parameters, and models and to decrease the risks of relying on a single best model. However, the usage of ensemble models in retail prediction has not been widely investigated in the research ^[7].

2. Applying Machine Learning in Retail Demand Prediction

Traditional machine learning techniques used for forecasting can be categorized into three main groups: (1) time series analysis, (2) regression-based approaches, and (3) supervised and unsupervised methods [13]. Time series analysis methods are the most widely used, encompassing techniques like autoregressive integrated moving average (ARIMA) and Holt Winter Exponential Smoothing (HW). These methods, particularly in the context of retail demand forecasting, are highly regarded for their ability to capture trends and seasonal demand patterns [14][15]. Regression-based methods have the flexibility to consider both independent and dependent variables [5]. And third, supervised and unsupervised models, like artificial neural networks (ANNs) or long short-term memory (LSTM), have been shown to perform better in nonlinear data [16]. The utilization of advanced models in forecasting has experienced growth in various industries, including oil, food and agriculture, public transportation, and retail. The performance of these models has been investigated in these industries, and it was found that ensemble models could outperform other models regarding accuracy [17][18][19].

To gain a deeper understanding of the application of advanced algorithms in demand forecasting, scholars have conducted a search using the keywords “demand forecasting” and “ensemble” in online databases focusing on recent years. The scholars found a significant number of studies that have used advanced machine learning approaches in the field of energy. For instance, Yu et al. (2016) proposed a new method for predicting crude oil prices using an ensemble empirical mode decomposition (EEMD) and extended extreme learning machine (EELM) to forecast electricity load [20]. Ribeiro et al. (2019) have presented a framework for short-term load forecasting using the Wavenet ensemble [21]. The framework involves transforming the data, determining an optimal time window, and selecting features. The proposed framework outperforms existing similar forecasting techniques, like multilayer perceptron neural networks. In electricity price forecasting, Zhang et al. (2022) introduced a hybrid deep neural network approach, which utilizes the Catboost algorithm for feature selection and a bidirectional long short-term memory neural network (BDLSTM) as the main forecasting engine [22]. In a recent study, Da Silva et al. (2021) proposed a new method for short-term prediction in microgrids called the Ensemble Prediction Network (EPN). The EPN comprises an ensemble of nine linear predictive nodes and is designed to provide an optimal estimate of predicted demand through least-squares optimization under certain constraints [23]. To overcome uncertainty, Yang et al. (2017) used combination approaches in a HAR model that considers lags of realized volatility and other potential predictors [24]. In the tourism industry, Cankurt (2016) developed and implemented ensemble learners for tourism demand forecasting based on M5P and M5-Rule model trees and random forest algorithms. The learners were evaluated using bagging, boosting, randomization, stacking, and voting techniques for forecasting tourism demand in Turkey [17]. Ensemble models have also been applied in public transportation, where Dai et al. (2018) presented a data-driven framework for short-term metro passenger flow prediction that utilizes both spatial and temporal information. The passenger flow information is obtained from smart-card data, and passenger flow patterns are explored. The proposed framework consists of two basic prediction models and a probabilistic model selection method (random forest classification) to combine the outputs for better prediction [25]. In an agricultural application field, Ribeiro and dos Santos Coelho (2020) investigated the accuracy of forecasting agricultural commodity prices through regression ensembles. The aim of their study was to compare the performance of ensembles (bagging, boosting, and stacking) with reference models such as support vector machine (SVR),

multilayer perceptron (MLP), and K-nearest neighbor (KNN) in forecasting prices one month ahead. Their study used monthly time series data for the price of soybean and wheat in the state of Parana, Brazil [26]. In the steel industry, Raju et al. (2022) compared the performance of different machine learning models for demand forecasting in the steel industry. Their study found that the best results came from a combination of models called STACK1 (extreme learning machine + gradient boosting + XGBR-SVR [27].

In the retail industry, a new heuristic approach was applied in a Turkish retail chain (SOK Market) with 4000 stores and 1500 SKUs. The results led to a reduction in stock outs, increased revenue by 30%, a 10% decrease in stock days, and a 34% reduction in waste for perishable products [8]. Das Adhikari et al. (2017) introduced a new ensemble technique using an averaging method that prioritizes algorithms with good accuracy and reduces deviation from actual sales. The method gives importance to algorithms that perform well based on historical data and penalizes those that deviate from actual sales [28]. Wang et al. (2018) applied ensemble empirical mode decomposition (EEMD) for global food price volatility and decomposed the original food price series into intrinsic mode functions and a residual. In their study, they mentioned that the low-frequency component contributes more to food price volatility, which is caused by notable events and policies. High-frequency components are mainly influenced by small events and market adjustments in a time series analysis. In the long term, food prices are determined by an intrinsic trend from global economic development. The findings reveal that food price volatility is a complex issue with multiple factors affecting both low and high frequencies [29]. In another study, daily optimal ordering quantities of fresh products using six methodologies (LSTM, SVR, RFR, GBR, XGBoost, and ARIMA) were analyzed. The paper compared the performance of conventional statistics, like ARIMA, and various machine learning algorithms, including RNNs (LSTM), SVRs, decision trees/ensemble methods (RFRs), and boosted trees (GBR, XGBoost). It was found that the LSTM and SVR machine learning algorithms outperformed the other demand forecasting models for the dataset [30].

Arora et al. (2020) focused on forecasting sales demand using historical data from a wholesale alcoholic beverage distributor. They employed an ensemble approach by combining traditional statistical models, multivariate models, and deep learning models. The study showed a reduction in the sale forecasting error by almost 50% and 33.5% for the most sold and highest revenue-grossing products respectively, compared to a naive model. The authors concluded that each product needs a unique model for accurate demand forecasting [31]. Sharma and Omair Shafiq (2020) used historical retail purchase data to predict the probability of item purchases. An ensemble learning model was built using random forests (RFs), Convolutional Neural Networks (CNNs), Extreme Gradient Boosting (XGBoost), and a voting mechanism. The model was evaluated using metrics such as accuracy, precision, F1 score, sensitivity, and specificity, and they experienced better performance using ensemble models than existing solutions [32].

Zhang et al. (2022) aimed to forecast weekly retail sales using Walmart's retail data from over five years. The forecast subject was divided into twenty-one time series based on different departments and states. Four machine learning models (naïve, moving average, prophet, ETS) were used to train the data, and stacking was used as the ensemble technique. The results showed that while the ensemble model using linear regression performed the best in the validation stage, the weighted average method supported by random forests was the best in the testing

stage. Linear regression was found to be overfitting. The research concluded that ensemble learning, especially weighted average, was a recommended method for forecasting [33]. Another previous study examined data mining's role in predicting retail sales for Walmart's outlets using supervised machine learning techniques. By analyzing factors like past sales, promotions, holidays, and economic indicators, the research helps businesses optimize sales forecasts and marketing strategies. The results suggest that simple regression techniques might not be optimal for short-term sales prediction with limited historical data. Ensemble learning techniques, involving the averaging of results from multiple decision trees, show better accuracy. Thus, for such scenarios, business owners are advised to opt for ensemble learning models [34]. Seyedan et al. (2022) proposed a demand forecasting methodology for the sports retail industry using ensemble learning. The methodology includes a cluster-based demand prediction using the time-series forecasting methods LSTM and prophet and majority voting and BMA as ensemble learning techniques. The aim was to improve the accuracy of future daily demand forecasting by combining different models and assigning higher weights to better-performing models. The results show that the clustered–ensembled approach improves prediction accuracy compared to using single models, leading to minimum values of MAPE, MAE, and RMSE. Their proposed framework had a considerable increase in prediction accuracy in various seasonal and monthly cases [18]. Ma et al. (2022) introduced a Spatial–Temporal Graph Attentional LSTM (STGA-LSTM) neural network for predicting short-term bike sharing demand utilizing various data sources. This model outperforms baseline approaches, leveraging deep learning to capture spatiotemporal patterns in bike sharing systems [35].

These models, showcased through studies ranging from predicting oil prices and electricity load to optimizing passenger flow and commodity prices, consistently demonstrate superior accuracy and performance compared to traditional methods. This trend extends to retail, where ensemble approaches have led to reduced stock outs, increased revenue, and enhanced sales forecasting precision. The remarkable versatility and success of these models across various sectors highlight their potential to reshape and refine demand forecasting practices, ultimately leading to more informed and effective decision-making processes.

References

1. Petropoulos, F.; Apiletti, D.; Assimakopoulos, V.; Babai, M.Z.; Barrow, D.K.; Ben Taieb, S.; Bergmeir, C.; Bessa, R.J.; Bijak, J.; Boylan, J.E.; et al. Forecasting: Theory and practice. *Int. J. Forecast.* 2022, 38, 705–871.
2. Brandtner, P.; Udokwu, C.; Darbanian, F.; Falatouri, T. Applications of Big Data Analytics in Supply Chain Management: Findings from Expert Interviews. In *Proceedings of the ICCMB 2021: 2021 the 4th International Conference on Computers in Management and Business*, Singapore, 30 January–1 February 2021; ACM: New York, NY, USA, 2021; pp. 77–82.
3. Brandtner, P. Predictive Analytics and Intelligent Decision Support Systems in Supply Chain Risk Management—Research Directions for Future Studies. In *Proceedings of the Seventh*

- International Congress on Information and Communication Technology, London, UK, 21–24 February 2022; Yang, X.-S., Sherratt, S., Dey, N., Joshi, A., Eds.; Lecture Notes in Networks and Systems; Springer Nature: Singapore, 2023; Volume 464, pp. 549–558.
4. Brandtner, P. Requirements for Value Network Foresight-Supply Chain Uncertainty Reduction. In ISPIIM Conference Proceedings; LUT Scientific and Expertise Publications: Lappeenranta, Finland, 2020; pp. 1–12.
 5. Falatouri, T.; Darbanian, F.; Brandtner, P.; Udokwu, C. Predictive Analytics for Demand Forecasting—A Comparison of SARIMA and LSTM in Retail SCM. *Procedia Comput. Sci.* 2022, 200, 993–1003.
 6. Fildes, R.; Ma, S.; Kolassa, S. Retail forecasting: Research and practice. *Int. J. Forecast.* 2022, 38, 1283–1318.
 7. Ma, S.; Fildes, R. Retail sales forecasting with meta-learning. *Eur. J. Oper. Res.* 2021, 288, 111–128.
 8. Akyuz, A.O.; Bulbul, B.A.; Uysal, M.O. Ensemble approach for time series analysis in demand forecasting: Ensemble learning. In Proceedings of the 2017 IEEE International Conference on Innovations in Intelligent Systems and Applications (INISTA), Gdynia, Poland, 3–5 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 7–12.
 9. Wang, X.; Hyndman, R.J.; Li, F.; Kang, Y. Forecast combinations: An over 50-year review. *Int. J. Forecast.* 2022, 39, 1518–1547.
 10. Januschowski, T.; Wang, Y.; Torkkola, K.; Erkkilä, T.; Hasson, H.; Gasthaus, J. Forecasting with trees. *Int. J. Forecast.* 2022, 38, 1473–1481.
 11. Kotu, V.; Deshpande, B. Classification. In *Data Science*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 65–163.
 12. Djarum, D.H.; Ahmad, Z.; Zhang, J. River Water Quality Prediction in Malaysia Based on Extra Tree Regression Model Coupled with Linear Discriminant Analysis (LDA). In Proceedings of the 31st European Symposium on Computer Aided Process Engineering, Computer Aided Chemical Engineering, Istanbul, Turkey, 6–9 June 2021; Elsevier: Amsterdam, The Netherlands, 2021; Volume 50, pp. 1491–1496.
 13. Sumaiya Farzana, G.; Prakash, N. (Eds.) Machine Learning in Demand Forecasting—A Review. In Proceedings of the 2nd International Conference on IoT, Social, Mobile, Analytics & Cloud in Computational Vision & Bio-Engineering, Thoothukudi, India, 29–30 October 2020.
 14. Arunraj, N.S.; Ahrens, D.; Fernandes, M. Application of SARIMAX model to forecast daily sales in food retail industry. *Int. J. Oper. Res. Inf. Syst. (IJORIS)* 2016, 7, 1–21.

15. Da Marques, F.; Alexandre, R. A Comparison on Statistical Methods and Long Short Term Memory Network Forecasting the Demand of Fresh Fish Products. Master's Thesis, Faculty of Engineering of the University of Porto, Porto, Portugal, 2020.
16. Alon, I.; Qi, M.; Sadowski, R.J. Forecasting aggregate retail sales: A comparison of artificial neural networks and traditional methods. *J. Retail. Consum. Serv.* 2001, 8, 147–156.
17. Cankurt, S. Tourism demand forecasting using ensembles of regression trees. In *Proceedings of the 2016 IEEE 8th International Conference on Intelligent Systems (IS)*, Sofia, Bulgaria, 4–6 September 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 702–708.
18. Seyedan, M.; Mafakheri, F.; Wang, C. Cluster-based demand forecasting using Bayesian model averaging: An ensemble learning approach. *Decis. Anal. J.* 2022, 3, 100033.
19. Liu, Z.; Jiang, P.; Wang, J.; Zhang, L. Ensemble forecasting system for short-term wind speed forecasting based on optimal sub-model selection and multi-objective version of mayfly optimization algorithm. *Expert Syst. Appl.* 2021, 177, 114974.
20. Yu, L.; Dai, W.; Tang, L. A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting. *Eng. Appl. Artif. Intell.* 2016, 47, 110–121.
21. Ribeiro, G.T.; Mariani, V.C.; Coelho, L.d.S. Enhanced ensemble structures using wavelet neural networks applied to short-term load forecasting. *Eng. Appl. Artif. Intell.* 2019, 82, 272–281.
22. Zhang, F.; Fleyeh, H.; Bales, C. A hybrid model based on bidirectional long short-term memory neural network and Catboost for short-term electricity spot price forecasting. *J. Oper. Res. Soc.* 2022, 73, 301–325.
23. da Silva, R.G.; Ribeiro, M.H.D.M.; Moreno, S.R.; Mariani, V.C.; dos Santos Coelho, L. A novel decomposition-ensemble learning framework for multi-step ahead wind energy forecasting. *Energy* 2021, 216, 119174.
24. Yang, K.; Tian, F.; Chen, L.; Li, S. Realized volatility forecast of agricultural futures using the HAR models with bagging and combination approaches. *Int. Rev. Econ. Finance* 2017, 49, 276–291.
25. Dai, X.; Sun, L.; Xu, Y. Short-Term Origin-Destination Based Metro Flow Prediction with Probabilistic Model Selection Approach. *J. Adv. Transp.* 2018, 2018, 5942763.
26. Ribeiro, M.H.D.M.; dos Santos Coelho, L. Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series. *Appl. Soft Comput.* 2020, 86, 105837.
27. Raju, S.M.T.U.; Sarker, A.; Das, A.; Islam, M.; Al-Rakhami, M.S.; Al-Amri, A.M.; Mohiuddin, T.; Albogamy, F.R. An Approach for Demand Forecasting in Steel Industries Using Ensemble Learning. *Complexity* 2022, 2022, 9928836.

28. Das Adhikari, N.C.; Garg, R.; Datt, S.; Das, L.; Deshpande, S.; Misra, A. Ensemble methodology for demand forecasting. In Proceedings of the 2017 International Conference on Intelligent Sustainable Systems (ICISS), Palladam, India, 7–8 December 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 846–851.
29. Wang, L.; Duan, W.; Qu, D.; Wang, S. What matters for global food price volatility? *Empir. Econ.* 2018, 54, 1549–1572.
30. Priyadarshi, R.; Panigrahi, A.; Routroy, S.; Garg, G.K. Demand forecasting at retail stage for selected vegetables: A performance analysis. *J. Model. Manag.* 2019, 14, 1042–1063.
31. Arora, T.; Chandna, R.; Conant, S.; Sadler, B.; Slater, R. Demand Forecasting In Wholesale Alcohol Distribution: An Ensemble Approach. *SMU Data Sci. Rev.* 2020, 3, 7.
32. Sharma, A.; Shafiq, M.O. Predicting purchase probability of retail items using an ensemble learning approach and historical data. In Proceedings of the 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA), Miami, FL, USA, 14–17 December 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 723–728.
33. Zhang, Y.; Zhu, H.; Wang, Y.; Li, T. Demand Forecasting: From Machine Learning to Ensemble Learning. In Proceedings of the 2022 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS), Dalian, China, 11–12 December 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 461–466.
34. Raizada, S.; Saini, J.R. Comparative Analysis of Supervised Machine Learning Techniques for Sales Forecasting. *Int. J. Adv. Comput. Sci. Appl.* 2021, 12, 102–110.
35. Ma, X.; Yin, Y.; Jin, Y.; He, M.; Zhu, M. Short-Term Prediction of Bike-Sharing Demand Using Multi-Source Data: A Spatial-Temporal Graph Attentional LSTM Approach. *Appl. Sci.* 2022, 12, 1161.

Retrieved from <https://encyclopedia.pub/entry/history/show/114123>