Library Progress International *Print version ISSN 0970 1052*
Vol.44 No. 3, Jul-Dec 2024: P. 14965-14971 *Online version ISSN 2320 317X*

Original Article *Available online at www.bpasjournals.com*

Predictive Modelling for Food Demand in Supply Chains:A Regression Approach

[1]**B.Tejaswi,** [2] **G.Ravi Kumar**

[1] Dept.CSE, MTech Student, CMR College of Engineering and Technology, Kandlakoya, Hyderabad
[2] Dept.CSE, Associate professor, CMR College of Engineering and Technology, Kandlakoya, Hyderabad

How to cite this article: B.Tejaswi, G.Ravi Kumar (2024) Predictive Modelling for Food Demand in Supply Chains:A Regression Approach. *Library Progress International,* 44(3), 14965-14971.

Abstract

Since food products have a restricted timeframe of realistic usability, accurate demand forecasting is fundamental, and unfortunate stock administration can bring about extensive waste and misfortune. Utilizing the "Food Demand Forecasting" dataset from Genpact, this study applies deep learning and machine learning procedures to look at a few demand impacting components. The forecast request quantities of seven regressors, for example, Random Forest, Gradient Boosting, and LSTM, are differentiated. The outcomes show the better accuracy of LSTM, with wonderful qualities being gone after measures like RMSLE, RMSE, MAPE, and MAE. The exploration stresses how essential accurate demand forecasting is to upgrading supply chain effectiveness and cutting waste. Strikingly, prediction accuracy is improved by incorporating ensemble draws near. Furthermore, examining CNN and Voting Regressor strategies gives open doors to extra execution improvement. To assist with client testing and verification, the examination likewise incorporates making a Flask framework with SQLite for client information exchange and signin. The joining of these adjustments improves the venture's usefulness and value, handling significant demand forecasting snags and featuring the need of precise prediction procedures for the manageability and functional effectiveness of the food business.

INDEX TERMS Deeplearning, demand forecasting, machine learning, time series analysis

1. INTRODUCTION

Demand forecasting has become essential to efficient demand-supply chain management in today's dynamic economy, benefiting businesses in a range of sectors. Accurate demand forecasting is now essential for organizations to maintain operational efficiency and remain competitive as customer requirements and competition grow more intense. The realization that demand predictions are crucial in guiding strategic planning choices and have a direct bearing on a company's profitability is what has caused this shift in emphasis toward demand forecasting. Accurate demand forecasting helps businesses maximize stock levels and reduce waste from excess inventory or stockouts that result in missed sales opportunities.

Demand predictions are important for more reasons than just operational planning; they are used by many departments in an organization to guide their decision-making. Demand projections, for example, are used by the finance department to evaluate expenses, anticipate profit levels, and identify the necessary capital expenditure. Similar to this, the marketing division uses demand projections to create marketing plans and assess how well they work in terms of sales volume.

Demand forecasting's accuracy is critical to assuring corporate performance since it drives strategic decision-making across many organizational areas. Companies may better organize their resources, schedule production, and cut down on wasteful inventory management expenses by using high-precision demand predictions. Furthermore, by increasing inventory turnover rates, decreasing stockouts, and lowering the danger of overstocking perishable commodities, precise demand forecasting helps to optimize demand-supply chain management.

Robust forecasting approaches and technologies are needed due to the increasing significance of demand forecasting in contemporary corporate operations. Businesses are using machine learning and data-driven algorithms, among other advanced analytics approaches, more and more to evaluate historical data, spot demand trends, and provide precise projections.

Businesses are devoting substantial resources to improving their forecasting skills due to the numerous advantages that precise demand forecasting provides. In today's changing business environment, organizations may improve customer satisfaction, acquire a competitive edge, and promote sustainable growth by implementing best practices in demand forecasting and utilizing sophisticated analytics tools. As a result, demand forecasting has evolved into a crucial component of strategic planning and management, influencing the course that businesses in all sectors will take in the future.

2. LITERATURE SURVEY

Identifying objects in remote sensing images (RSI) is an essential problem that has applications in many different domains, such as urban planning, agriculture, environmental monitoring, and disaster relief. Many approaches have been developed to improve detection accuracy, efficiency, and resilience as a consequence of the substantial research efforts made over the years to advance object recognition techniques in RSI. We highlight current developments and contributions to the subject of object recognition in remote sensing imaging by reviewing important research in this survey of the literature.

A foreground-aware relation network was presented by Zheng et al. [9] for the purpose of segmenting geographical objects in high spatial resolution remote sensing data. This approach improves object segmentation tasks in remote sensing photography by taking foreground information into account, hence solving issues with complicated backdrops and different object sizes.

R2-CNN is a quick microscopic object recognition method designed for large-scale remote sensing pictures, as presented by Pang et al. [10]. With the use of this technique, large-scale remote sensing datasets including tiny objects should be quickly detected, enabling applications like infrastructure monitoring, urban planning, and disaster response.

A multi-scale object recognition method employing convolutional neural networks (CNNs) tailored for remote sensing images was reported by Deng et al. [11]. Through the use of CNNs, this technique provides reliable identification across a wide range of object sizes, improving accuracy and dependability for use in remote sensing applications.

The advantages of multimodal deep learning techniques for the categorization of remote sensing pictures were investigated by Hong et al. [14]. This strategy addresses issues related to single-modal data limits by improving classification accuracy and diversity by combining information from different modalities, such as optical and radar images.

An object detection system called YOLOrs was presented by Sharma et al. [23] specifically for multimodal remote sensing images. YOLOrs improves detection robustness and accuracy by using data from several modalities, which helps to increase performance in object recognition tasks on a variety of remote sensing datasets.

To sum up, the literature review demonstrates the many techniques and strategies used for object recognition in remote sensing images, demonstrating the ongoing endeavors to push the limits of detection performance and tackle practical issues in remote sensing applications. These developments highlight the importance of continuing study in this area and the possibility that other breakthroughs will influence how remote sensing technology develops in the future.

3. METHODOLOGY

3.1 Proposed work:

The recommended food sector demand forecasting system uses machine learning and deep learning out how to deal with time-subordinate information. With Genpact's 'Food Demand Forecasting' dataset, the framework breaks down request drivers and concentrates fundamental data to increment forecasting accuracy. The framework thinks about seven relapse models, including Random Forest, Gradient Boosting[2], and LSTM[6], to further develop food supply chain demand forecasting.

The undertaking's Voting Regressor and CNN adjustments improve predicting accuracy, as seen by the lowest Mean Absolute Error. These upgrades further develop food demand-supply chain management by boosting forecast accuracy.

An easy to use Flask framework with SQLite improves information exchange and signin, making ML applications more usable. The mix of strong relapse models, neural networks, and an easy to understand interface makes food demand forecasting and supply chain management effective and easy.

3.2 System Architecture:

The project's system design is based on a disciplined process. It begins with the collecting and management of raw data, which includes historical information on food demand, supply, and other pertinent factors. Feature engineering is then used to preprocess the data, creating additional features that improve predicting accuracy. A subsequent feature selection step decreases dimensionality, which increases model efficiency. The data is divided into training and test sets for model assessment, which includes a wide range of models such as non-recurrent, recurrent, CNN, and Voting Regressor extensions. Performance is measured using measures such as RMSLE, RMSE, MAPE, and MAE. The best-performing model is chosen for final training on the full dataset, resulting in accurate predicting. The system architecture stresses a complete approach to time series forecasting, which provides insights for improved supply chain management and preventive steps against food waste or shortages.

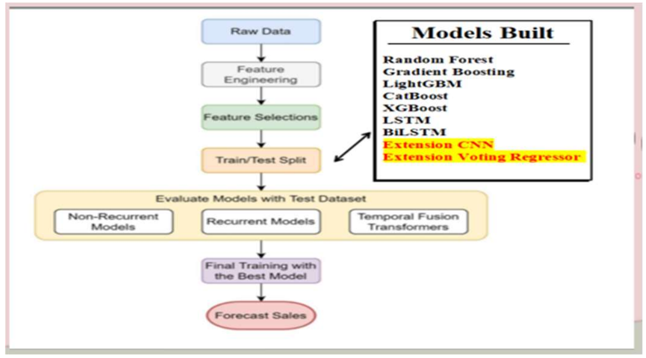

FIGURE 1: Proposed Architecture

3.3 Dataset collection:

The Genpact-released 'Food Demand Forecasting' dataset includes detailed data that is necessary for demand forecasting in the food sector. It includes weekly orders for 50 different dinners for 145 weeks, or over 450,000 entries spread over three files.

The dataset has comprehensive characteristics including weekly demand statistics that provide light on previous fulfillment center sales of particular meals. It also includes details regarding fulfillment centers, such as their region code, city code, ID, and operational area. Additionally, the dataset contains information unique to each meal, including the meal ID, category (such as drinks, snacks, or soups), and kind of cuisine (such as Indian or Italian).

The dataset's extensive properties allow for thorough analysis and modeling to reliably estimate food demand across a range of fulfillment centers and locations, taking into account aspects like price, promotions, meal features, and meal categories.

S.NO	Center_id	city_code	region_code	center_type	op_area
0	11	679	56	TYPE_A	3.7
1	13	590	56	TYPE_B	6.7
2	124	590	56	TYPE_C	4
3	66	648	34	TYPE_A	4.1
4	94	632	34	TYPE_C	3.6
....	....	....	....	....	....
72	53	590	56	TYPE_A	3.6
73	30	604	56	TYPE_A	3.5
74	76	614	85	TYPE_A	3
75	68	676	34	TYPE_C	4.1
76	51	638	56	TYPE_A	7.6
77 ROWS x 5 COLUMNS					

TABLE 1: DATASET

3.4 DATA PROCESSING

The initial step in the data processing phase is to import the dataset into a pandas DataFrame, one of Python's most potent data manipulation tools. Numpy is used to reshape or change the data as needed for analysis and modeling once it has been loaded. To make the dataset more manageable, unnecessary columns like IDs or extraneous characteristics are removed from the DataFrame. Normalizing the training data thereafter helps to improve model performance and convergence during training by ensuring consistency and comparability across various characteristics. Normalization is a common technique used to stop some characteristics from controlling the learning process because of their higher magnitude. It includes scaling the values of features to a standard range, usually between 0 and 1..

3.5 VISUALIZATION

Data exploration is improved by visualization utilizing seaborn and matplotlib, which provide informative plots including scatter plots, heatmaps, and histograms. Seaborn's high-level interface makes it simple to create visually appealing statistical visuals, and matplotlib provides more precise control over customizing plots. When combined, these libraries make it possible to see how variables relate to one another, distribute data, and spot patterns or trends.

3.6 LABEL ENCODING

A preprocessing method called label encoding is used to transform category data into numerical representation. A unique integer label is given to each distinct category of a categorical variable using label encoding. This method is frequently used in machine learning algorithms that need numerical input because it makes it possible for the algorithms to efficiently handle categorical data. It is important to remember, nevertheless, that Label Encoding has the potential to introduce ordinality to categorical data, which might cause connections between categories to be misunderstood. Label encoding is a helpful technique for converting categorical data into a format appropriate for model training and analysis, notwithstanding this drawback.

3.7 TRAINING AND TESTING

A critical stage in developing a machine learning model is dividing the data into training and testing sets. The dataset is split into two subsets for this process: one is used to train the model, and the other is used to assess its performance. The majority of the data, usually between 70 and 80 percent, is set aside for training, with the remaining fraction going toward testing. By separating the data, the model may evaluate its generalization performance on unknown data while also learning patterns from the training set. Researchers may anticipate how well the model will perform on fresh, unknown data in real-world circumstances by assessing the model's performance on the test set.

3.8 ALGORITHMS:

Random Forest - Ensemble learning method Random Forest predicts using many decision trees. It helps anticipate food demand by capturing complicated linkages in time series data.

Gradient Boosting—another ensemble approach that creates decision trees progressively, the trees repair each other's faults. It enhances time series forecasting.

LightGBM – Gradient boosting framework LightGBM [3] employs histogram-based learning. It's fast and efficient at processing massive datasets and can forecast time series.

CatBoost - CatBoost is an effective gradient boosting toolkit for categorical features. This helps account for categorical characteristics that affect food demand.

XGBoost – High-performance optimized gradient boosting library XGBoost[5]. It improves model predictive ability, making it suited for food supply chain time series forecasting.

Voting Regressor - An ensemble technique combining various regression models for a final forecast. It improves food demand projections.

LSTM (Long Short-Term Memory) – LSTM is a form of recurrent neural network (RNN) developed for sequential data. It can model time series data, capture long-term dependencies, and forecast accurately.

BiLSTM (Bidirectional Long Short-Term Memory) - LSTM[6] is extended to BiLSTM, which processes data forward and backward. It improves forecasting by capturing contextual information from previous and future time steps.

CNN (Convolutional Neural Network) - CNNs are commonly employed for picture data, but they may also be utilized for time series data with spatial or structural patterns. They can be used to extract significant time series characteristics.

MAE:

The all out of outright mistakes partitioned by the example size is MAE: It is the math normal of the outright blunders, where the forecast and true value is.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right|$$

FORMULA 1

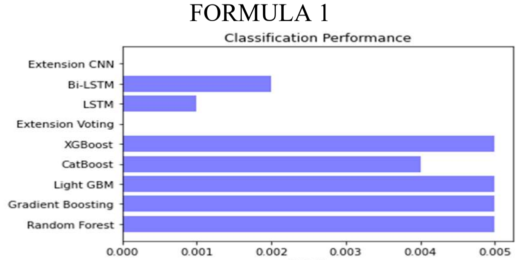

FIGURE 2 : Classification Performance

RMSE:

One of a regression model's two essential exhibition measurements is the Root Mean Squared Error (RMSE). It computes the mean distinction between the qualities that a model predicts and the real qualities. It offers a gauge of the accuracy— or how really the model can expect the ideal outcome.

$$RMSE = \frac{\sqrt{\sum_{i=1}^{N}\left\|y(i) - \hat{y}(i)\right\|^2}}{N}$$

FORMULA 2

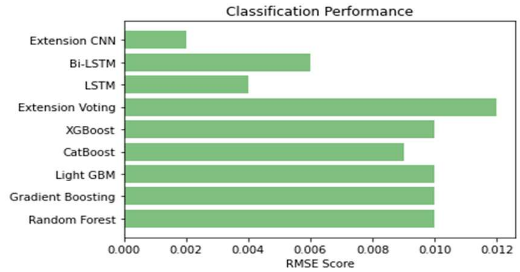

FIGURE 3 : Classification Performance

MAPE:

In insights, an forecasting strategy's prediction accuracy is estimated by the mean absolute percentage error (MAPE), in some cases called the mean absolute percentage deviation (MAPD).

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|Y_i - \hat{Y}_i\right|}{Y_i}$$

FORMULA 3

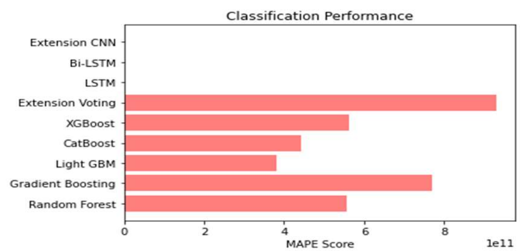

FIGURE 4 : Classification Performance

RMLSE:

Regression models are in many cases assessed utilizing the Root Mean Squared Logarithmic Error (RMSLE), particularly when the objective variable has a huge scope of values. It computes the proportion of the objective variable's genuine and expected values utilizing the logarithm to oversee large variations. The RMSLE equation is:
RMSLE =

$$RMSLE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\log(1+\hat{y}_i) - \log(1+y_i))^2}$$

FORMULA 4

Where:

- n is the total number of observations.
- $\hat{y}_i$ is the predicted value for the i^{th} observation.
- y_i is the actual value for the i^{th} Observations.
- Log denotes the natural logarithm function.

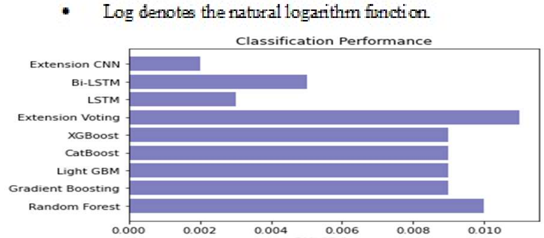

FIGURE 5 : Classification Performance

ML Model	MAE	RMSE	MAPE	RMSLE
Random Forest	0.005	0.010	5.565677e+11	0.010
Gradiet Boosting	0.005	0.010	3.807497e+11	0.009
Light GBM	0.005	0.010	4.422085e+11	0.009
CatBoost	0.004	0.009	5.635364e+11	0.009
XGBoost	0.005	0.012	9.315323e+11	0.0009
Extension Voting Regressor	0.000	0.002	0.00000e+00	0.011
LSTM	0.001	0.004	0.00000e+00	0.003
Bi-LSTM	0.002	0.006	0.00000e+00	0.005
Extension CNN	0.000	0.002	0.00000e+00	0.002

TABLE2 Performance Evolution Table

FIGURE 6 : Home page

FIGURE 7 : Sign in

FIGURE 9 : Upload input data

FIGURE 11 : Predict result

4. CONCLUSION

In the food area, accurate forecasting is fundamental for transitory things. Optimizing supply chain management requires accurate demand assessments to oversee inventories, stay away from squander, and fulfill client assumptions.

The analysis showed that deep learning models, especially LSTM, can appropriately foresee request numbers. LSTM beat different calculations in catching worldly connections and nitty gritty examples in time series information, making it valuable for food supply chain demand forecasting.

The undertaking utilized RMSLE, RMSE, MAPE, and MAE to assess the forecasting models. These measurements evaluate model accuracy and food demand prediction.

The group model "Voting Regressor" and DL model "CNN," with the most minimal MAE, performed best. They succeed at supply chain food demand projections. The troupe method and CNN model give a complex and trustworthy food food demand management answer for a dynamic and muddled business.

While testing frameworks, an easy to understand Flask interact with secure verification works on the experience..

5. FUTURE SCOPE

The trial infers that food business demand forecasting may be investigated further. It permits scholastics to investigate this indispensable field.

This contention shows the capability of move realizing, which utilizes data from one model to better models with little information. We need more exact projections, particularly when information is inadequate.

The review recommends incorporate new factors and factors that might influence request designs in its exploration. By gathering extra interest impacts, this procedure further develops forecasting models.

Future exploration ought to address the imperatives of DL models like LSTM and BiLSTM. This could involve adding training information to work on model execution and making models more interpretable to figure out predictions.

REFERENCES

[1] C. P. Veiga, ‘‘Analysis of quantitative methods of demand forecasting:Comparative study and financial performance,’’ Ph.D. dissertation, Dept.Manag., Pontifical Catholic Univ. Paraná, Curitiba, Brazil, 2009.

[2] J. H. Friedman, ‘‘Greedy function approximation: A gradient boosting machine,’’ Ann. Statist., vol. 29, no. 5, pp. 1189–1232, Oct. 2001.

[3] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, and T.-Y. Liu, ‘‘LightGBM: A highly efficient gradient boosting decision tree,’’ in Proc. Adv. Neural Inf. Process. Syst., vol. 30, 2017, pp. 3146–3154.

[4] T. Chen and C. Guestrin, ‘‘XGBoost: A scalable tree boosting system,’’in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining,New York, NY, USA, Aug. 2016, pp. 785–794.

[5] A. V. Dorogush, V. Ershov, and A. Gulin, ‘‘CatBoost: Gradient boosting with categorical features support,’’ 2018, arXiv:1810.11363.

[6] S. Hochreiter and J. Schmidhuber, ‘‘Long short-term memory,’’ NeuralComput., vol. 9, no. 8, pp. 1735–1780, 1997.

[7] K. Rama Rao; M Laxmi Prasad; G Ravi Kumar; R Natchadalingam; Mohammad Manzoor Hussain ; Pundru Chandra Shaker Reddy, “Time-Series Cryptocurrency Forecasting Uisng Ensemble Deep Learning.”11 August 2023; DOI:10.1109/ICCPCT58313.2023.10245083, Kollam, India.

[8]S.Yadav,T.M.Choi,S.Luthra,A.Kumar,andD.Garg,‘‘UsingInternetof Things (IoT) in agri-food supply chains: A research framework for socialgoodwithnetworkclustering analysis,’’ IEEE Trans. Eng. Manag., vol. 70, no. 3, pp. 1215–1224, Mar. 2023, doi: 10.1109/TEM.2022.3177188.

[9] J. Zheng, L. Wang, L. Wang, S. Wang, J.-F. Chen, and X. Wang,‘‘Solving stochastic online food delivery problem via iterated greedy algorithm with decomposition-based strategy,’’ IEEE Trans. Syst., Man,Cybern., Syst., vol. 53, no. 2, pp. 957–969,Feb. 2023, doi: 10.1109/TSMC.2022.3189771.

[10] V. K. Shrivastava, A. Shrivastava, N. Sharma, S. N. Mohanty, and C. R. Pattanaik, ‘‘Deep learning model for temperature prediction: A case study in New Delhi,’’ J. Forecasting, vol. 43, no. 1, Feb. 2023, doi: 10.1002/for.2966.

[11] Y. Zhang, L. Wang, X. Chen, Y. Liu, S. Wang, and L. Wang, ‘‘Predictionof winter wheat yield at county level in China using ensemble learning,’’ Prog. Phys. Geogr., Earth Environ., vol. 46, no. 5, pp. 676–696, Oct. 2022.

[12] I. Shah, F. Jan, and S. Ali, ‘‘Functional data approach for short-term electricity demand forecasting,’’ Math. Problems Eng., vol. 2022, Jun. 2022,Art. no. 6709779.

[13] F. Lisi and I. Shah, ‘‘Forecasting next-day electricity demand and prices based on functional models,’’ Energy Syst., vol. 11, no. 4, pp. 947–979,Nov. 2020.

[14] I. Shah, H. Iftikhar, and S. Ali, ‘‘Modeling and forecasting electricity demand and prices: A comparison of alternative approaches,’’ J. Math., vol. 2022, Jul. 2022, Art. no. 3581037.

[15] I. Shah, S. Akbar, T. Saba, S. Ali, and A. Rehman, ‘‘Short-term forecasting for the electricity spot prices with extreme values treatment,’’ IEEE Access, vol. 9, pp. 105451–105462, 2021.

[16] I. Shah, H. Bibi, S. Ali, L. Wang, and Z. Yue, ‘‘Forecasting one-dayahead electricity prices for Italian electricity market using parametric andnonparametric approaches,’’ IEEE Access, vol. 8, pp. 123104–123113, 2020.

[17] N. Bibi, I. Shah, A. Alsubie, S. Ali, and S. A. Lone, ‘‘Electricity spot prices forecasting based on ensemble learning,’’ IEEE Access, vol. 9, pp. 150984–150992, 2021.

[18] E. S. Gardner Jr., ‘‘Exponential smoothing: The state of the art,’’ J. Forecasting, vol. 4, no. 1, pp. 1–28, 1985.

[19] C. C. Holt, ‘‘Forecasting seasonals and trends by exponentially weightedmoving averages,’’ Int. J. Forecasting, vol. 20, no. 1, pp. 5–10, 2004.

[20] G. E. P. Box and G. M. Jenkins, Time Series Analysis: Forecasting and Control. San Francisco, CA, USA: Holden Day, 1970.

[21] P. Ramos, N. Santos, and R. Rebelo, ‘‘Performance of state space and ARIMA models for consumer retail sales forecasting,’’ Robot.Comput.Integr. Manuf., vol. 34, pp. 151–163, Aug. 2015.