Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 8, August : 2024

Optimizing Food Forecasts: A Study on ML and DL Approaches for Accurate Demand Prediction

[1]**Konda Sravani**, [2] **P. Subhan basha**
[1]M. Tech student, [2] Associate Professor
[1,2]Department of CSE
[1,2]Sri Sai Institute of Technology and Science, Masapeta, Rayachoty (M), Annamayya Dt, Andhra Pradesh, India.
subhan.mahammad@gmail.com, kondasravani789@gmail.com

ABSTRACT

Accurate demand forecasting is crucial in the food industry due to the short shelf life of many products, making proper inventory management essential to minimize waste and loss. Recent advancements in machine learning and deep learning have significantly improved handling time-dependent data. This paper uses the 'Food Demand Forecasting' dataset from Genpact to analyze various factors affecting demand and proposes a comparative study of seven regressors for order prediction. The study employs Random Forest, Gradient Boosting, LightGBM, XGBoost, CatBoost, LSTM, and Bidirectional LSTM. Results indicate the superior performance of LSTM models. Additionally, the CNN2D algorithm was utilized to optimize dataset features, achieving more accurate forecasting compared to other algorithms.

Index Terms: Food, LSTM

INTRODUCTION:

In today's dynamic market, companies face the dual challenge of meeting consumer demands and staying ahead in an increasingly competitive environment. To tackle this, many businesses are turning to demand forecasting as a key strategy for effective demand-supply chain management. Accurate demand forecasting is essential because it directly influences a company's profitability. If demand is overestimated, it can lead to excess inventory, resulting in high storage costs and waste. Conversely, underestimating demand can cause stockouts, driving customers to competitors. Therefore, reliable demand forecasting is crucial for maintaining a balance between supply and demand, optimizing resources, and maximizing profits.

Demand forecasting plays a pivotal role across various departments within a company. The financial department relies on forecasts to estimate costs, profits, and capital requirements. Accurate predictions help in budgeting and financial planning, ensuring that the company allocates resources efficiently. The marketing department uses demand forecasts to shape its strategies, assessing the potential impact of different marketing campaigns on sales volumes. This allows for more targeted and effective marketing efforts, enhancing customer engagement and boosting sales.

LITERATURE SURVEY:

N. Bibi ***et al***

Efficient electricity price modeling and forecasting are vital in today's competitive markets, yet challenging due to the complex nature of price series. This study evaluates an ensemble-based technique for short-term electricity spot price forecasting in the Italian market (IPEX). The approach splits the price time series into deterministic components, estimated using semi-parametric techniques, and stochastic components, modeled with time series and machine learning algorithms. Results show that the ensemble-based model excels based on standard accuracy measures, with random forest and ARMA also proving highly competitive.

C.-Y. Chen ***et al***

Time series forecasting has long been a significant research area, employing statistical, machine learning, and hybrid methods to tackle its challenges. However, assumptions of data normality are often unaddressed. This research extends forecasting models to account for heavy-tailed noise distributions, focusing on normal and Student's t distributions. We compare the SARIMAX model, using maximum likelihood, with the regression tree-based random forest method. Our study includes not only forecasts but also prediction intervals, which are crucial for practical applications. While centered on these models, our findings have broader implications for other systems.

PROBLEM STATEMENT:

Accurate forecasting is become necessity in food industry to meet demand supply requirements. Many food products has shelf life and if demand forecast is not accurate and then either short life products will be wasted or in some scenarios it goes for shortage. Many deep learning or machine learning algorithms was introduced for accurate forecasting but they lack support of Time series or LAG data

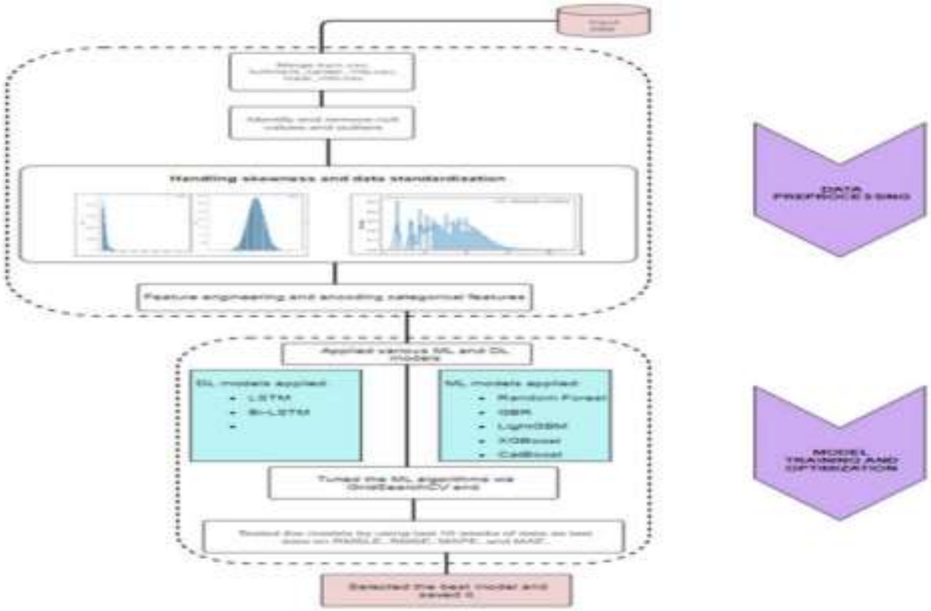
PROPOSED METHOD:

To address this problem, the author extracts lag data from the dataset and assigns weights to the target variable using an alpha value of 0.5. Products that are recent or in high demand receive higher weights. Time series data is gathered for 10-week periods, and forecasting is conducted for the subsequent 10 weeks. The extracted lag data is trained using various machine learning and deep learning algorithms, including Random Forest, Gradient Boosting, XGBOOST, CATBOOST, LIGHT GBM, LSTM, and BI-LSTM. Each algorithm's performance is evaluated based on RMSE, MAE, MAPE, and RMSLE, which measure the difference between actual and predicted values. The lower the difference, the better the algorithm performs. LSTM showed the lowest MAE and RMSE error rates among all algorithms.

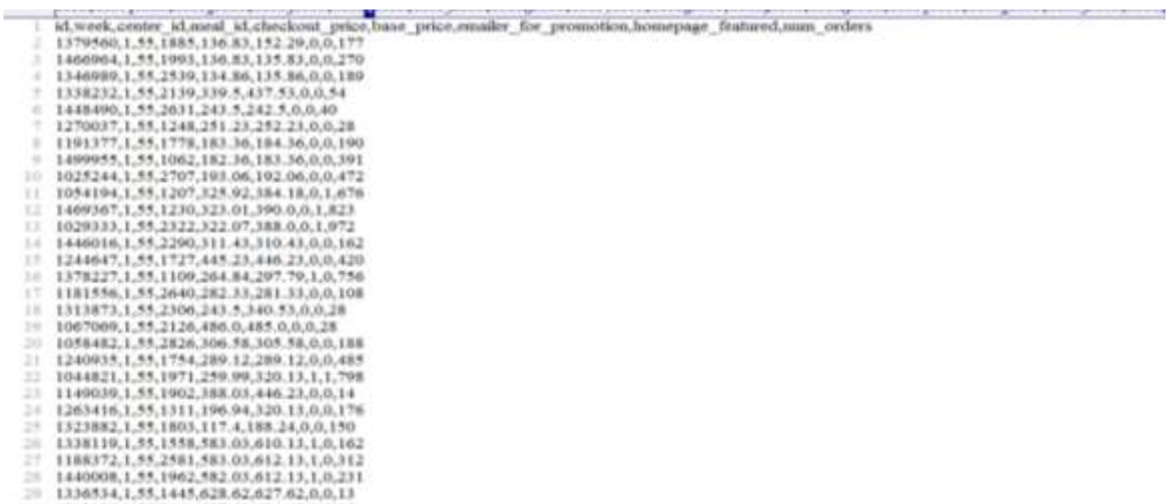
Industrial Engineering Journal

ISSN: 0970-2555

Volume : 53, Issue 8, August : 2024



ARCHITECTURE



```
id,week,center_id,meal_id,checkout_price,base_price,emailer_for_promotion,homepage_featured,num_orders
1379560,1,55,1885,136.83,152.29,0,0,177
1466964,1,55,1993,136.83,135.83,0,0,270
1346989,1,55,2539,134.86,135.86,0,0,189
1338232,1,55,2139,339.5,437.53,0,0,54
1448490,1,55,2631,243.5,242.5,0,0,40
1270037,1,55,1248,251.23,252.23,0,0,28
1191377,1,55,1778,183.36,184.36,0,0,190
1499955,1,55,1062,182.36,183.36,0,0,391
1025244,1,55,2707,193.06,192.06,0,0,472
1054194,1,55,1207,325.92,384.18,0,1,676
1469367,1,55,1230,323.01,390.0,0,1,823
1029333,1,55,2322,322.07,388.0,0,1,972
1446016,1,55,2290,311.43,310.43,0,0,162
1244647,1,55,1727,445.23,446.23,0,0,420
1378227,1,55,1109,264.84,297.79,1,0,756
1181556,1,55,2640,282.33,281.33,0,0,108
1313873,1,55,2306,243.5,340.53,0,0,28
1067009,1,55,2126,486.0,485.0,0,0,28
1058482,1,55,2826,306.58,305.58,0,0,188
1240935,1,55,1754,289.12,289.12,0,0,485
1044821,1,55,1971,259.99,320.13,1,1,798
1149039,1,55,1902,388.03,446.23,0,0,14
1263416,1,55,1311,196.94,320.13,0,0,176
1323882,1,55,1803,117.4,188.24,0,0,150
1338119,1,55,1558,583.03,610.13,1,0,162
1188372,1,55,2581,583.03,612.13,1,0,312
1440008,1,55,1962,582.03,612.13,1,0,231
1336534,1,55,1445,628.62,627.62,0,0,13
```

TIME SERIES DATASET

The first row contains dataset column names, while the remaining rows hold dataset values. The last column, representing orders as sales, helps calculate the target variable using the EWMA (Exponentially Weighted Moving Average) formula. This dataset will be used to train and test the performance of each algorithm.

METHODOLOGY:

1. Data Collection

Load the Datasets

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 8, August : 2024

Two key datasets are crucial for this analysis: the meal sales dataset and the fulfillment center dataset. The meal sales dataset typically contains information on sales quantities, dates, and possibly meal types. The fulfillment center dataset provides details about various centers, including location, type, and operational metrics.

Merge the Datasets

Combining these datasets involves merging them on common attributes, such as fulfillment center IDs or dates. This integration creates a unified dataset that consolidates all relevant information, providing a comprehensive view of the data and enabling more robust analysis.

2. Exploratory Data Analysis (EDA)

Visualize the Distribution of Features

To gain insights into the data, it is crucial to visualize the distribution of different features. Histograms and box plots are commonly used to illustrate the spread and central tendencies of numerical variables. Histograms show the frequency distribution of values, while box plots provide a visual summary of the data's central tendency, dispersion, and presence of outliers.

Analyze Relationships Between Features

Correlation matrices and heatmaps are useful for exploring relationships between different features. A correlation matrix quantifies the strength and direction of linear relationships between variables, helping identify which features are strongly related. Heatmaps visually represent these correlations, making it easier to detect patterns and anomalies.

Investigate Trends and Patterns

Examining meal sales across various dimensions, such as different regions, fulfillment center types, and time periods (e.g., weeks), helps uncover trends and patterns. Analyzing these aspects can reveal seasonal variations, regional preferences, and operational impacts on sales, providing valuable insights for forecasting and strategic planning.

3. Feature Engineering and Preprocessing

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 8, August : 2024

Perform Feature Engineering

Feature engineering involves creating new features or modifying existing ones to improve the model's performance. This process might include generating features like sales trends, promotional periods, or interactions between different variables. These engineered features can provide additional context and improve predictive accuracy.

Handle Missing Values and Encode Categorical Variables

Data often contains missing values, which must be addressed through imputation or removal. Categorical variables, such as fulfillment center types, need to be encoded into numerical formats for model compatibility. Techniques like Label Encoding convert categories into numeric labels, ensuring that machine learning algorithms can process these features effectively.

Normalize the Features

Normalization is essential to ensure that features with different scales contribute equally to the model. Min-Max scaling is a common technique that scales features to a range between 0 and 1. This step standardizes the input data, improving the convergence and performance of machine learning models.

Prepare Lagged Data

For time-series forecasting, incorporating lagged data (previous sales figures) can provide valuable information about temporal patterns. By including these lagged features, models can learn from past trends to make more accurate predictions about future sales.

4. Model Selection and Training

Traditional Machine Learning Models

Several traditional machine learning models are trained to predict meal sales:

- **RandomForestRegressor**: An ensemble method that combines multiple decision trees to improve predictive performance and handle non-linear relationships.

- **GradientBoostingRegressor**: Another ensemble technique that builds models sequentially to correct errors made by previous models.

Hyperparameter tuning using techniques like GridSearchCV optimizes model parameters to enhance performance. Evaluation metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Logarithmic Error (RMSLE) are used to assess the accuracy and reliability of these models.

Deep Learning Models

Deep learning models offer advanced capabilities for capturing complex patterns in the data:

- **Long Short-Term Memory (LSTM) Networks**: Effective for sequential data, LSTMs can capture long-term dependencies and trends in time-series data.
- **Bidirectional LSTM (Bi-LSTM)**: Enhances LSTM performance by processing data in both forward and backward directions.
- **Convolutional Neural Networks (CNN)**: Typically used for image data but can be adapted for time-series forecasting to capture local patterns through convolutional layers.

These models are compiled and trained using appropriate loss functions and optimizers. Techniques like dropout are implemented to prevent overfitting, ensuring that the models generalize well to unseen data. The best-performing models are saved using tools like ModelCheckpoint to avoid the need for retraining.

5. Model Evaluation and Performance Metrics

Calculate Performance Metrics

To evaluate model performance, metrics such as MAE, RMSE, MAPE, and RMSLE are calculated. MAE measures the average magnitude of errors in predictions, while RMSE penalizes larger errors more significantly. MAPE expresses errors as a percentage of actual values, and RMSLE focuses on relative errors, especially useful when dealing with exponential growth.

Visualize Predictions

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 8, August : 2024

To understand model performance, predicted sales are compared against actual sales. Visualization techniques, such as scatter plots or line charts, help assess how well the models' predictions align with real data, highlighting areas where the models may be overestimating or underestimating sales.

6. Results Interpretation and Conclusion

Analyze Model Performance

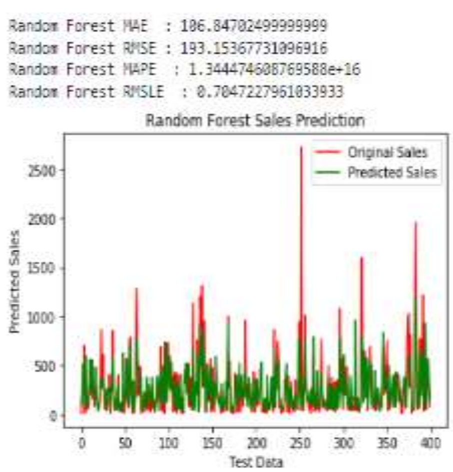
The performance of each model is analyzed based on the evaluation metrics. This analysis includes comparing the effectiveness of different models in predicting meal sales and identifying which models provide the most accurate and reliable forecasts.

Discuss Insights and Improvements

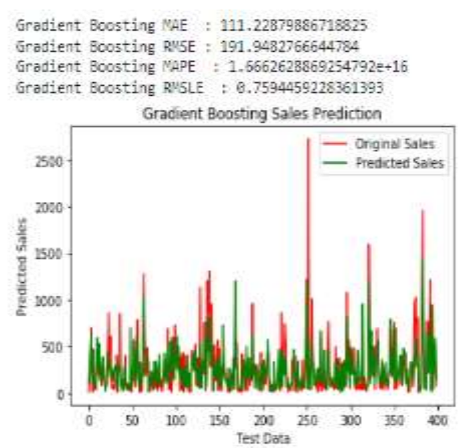
The insights gained from the analysis are discussed, including any trends, patterns, or anomalies identified. Potential areas for improvement are also explored, such as refining feature engineering, adjusting model parameters, or incorporating additional data sources.

Analysis of meal sales using a combination of traditional and deep learning models, along with thorough data preprocessing and feature engineering, provides a comprehensive understanding of sales patterns and forecasting accuracy.

RESULTS:



Training Random Forest with tuning parameters on train dataset and then testing on test data to calculate RMSE values



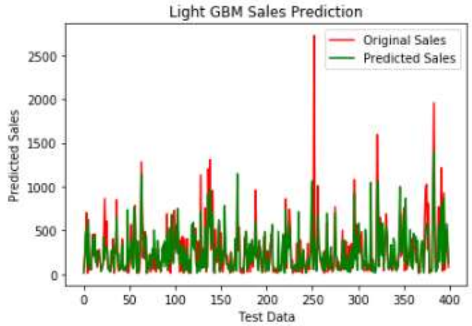
Industrial Engineering Journal

ISSN: 0970-2555

Volume : 53, Issue 8, August : 2024

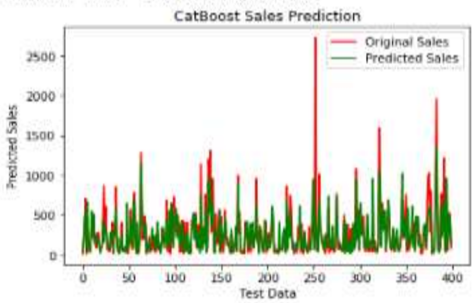
Training gradient boosting and its MAE values is 111

```
Light GBM MAE  : 107.33588861995717
Light GBM RMSE : 190.58221989704842
Light GBM MAPE  : 1.4588182313031324e+16
Light GBM RMSLE  : 0.6972932697969428
```



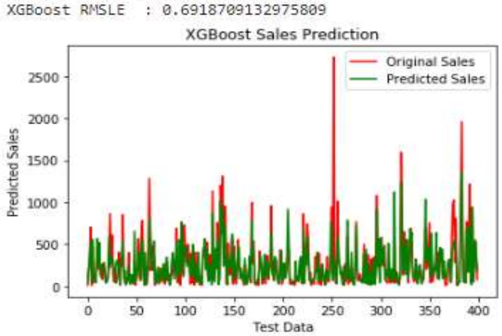
Training LIGHTGBM got 107 as MAE

```
CatBoost MAE  : 98.27724852062401
CatBoost RMSE : 176.9991813036784
CatBoost MAPE  : 9841604899730246.0
CatBoost RMSLE  : 0.6563690268193625
```



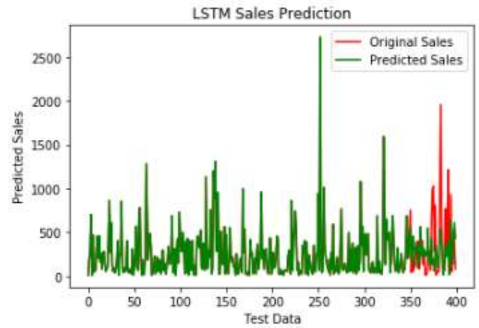
Training CATBOOST 98 as MAE

```
XGBoost MAE  : 101.92435542106628
XGBoost RMSE : 187.53665026021528
XGBoost MAPE  : 7687138193070486.0
XGBoost RMSLE  : 0.6918709132975809
```



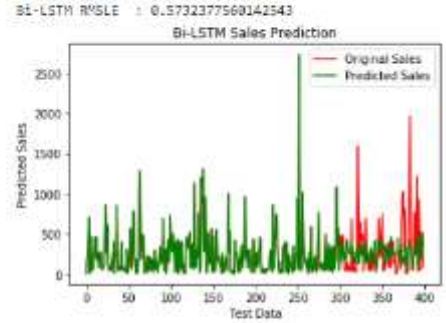
Training XGBOOST got 101 as MAE

```
LSTM MAE  : 29.706652455329895
LSTM RMSE : 127.53145787605564
LSTM MAPE  : 1876499844774064.8
LSTM RMSLE  : 0.4352678722296506
```



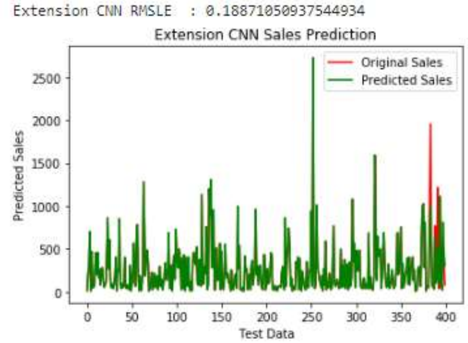
Training LSTM got 29

```
Bi-LSTM MAE  : 53.219437787532804
Bi-LSTM RMSE : 162.82636632633066
Bi-LSTM MAPE  : 1876499844774064.8
Bi-LSTM RMSLE  : 0.5732377560142543
```

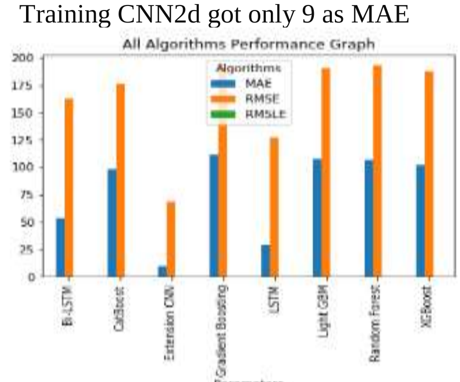


Training BI-LSTM got 53% MAE

```
Extension CNN MAE  : 9.761380290985112
Extension CNN RMSE : 68.73260903573932
Extension CNN MAPE  : 1876499844774064.8
Extension CNN RMSLE  : 0.18871050937544934
```



Training CNN2d got only 9 as MAE



Industrial Engineering Journal

ISSN: 0970-2555

Volume : 53, Issue 8, August : 2024

In all algorithms LSTM and extension CNN2d got less MSE and RMSE error rates

Prediction:

```
Test Data : [1151666 1 89 2640 281.33 280.33 0 0 703 56 'TYPE_A' 4.8] Predicted Sales ===> 205.91173
Test Data : [1048572 1 89 1878 282.33 280.33 0 0 703 56 'TYPE_A' 4.8] Predicted Sales ===> 289.71167
Test Data : [1379525 1 89 2306 243.5 242.5 0 1 703 56 'TYPE_A' 4.8] Predicted Sales ===> 556.96893
Test Data : [1152138 1 89 1216 456.93 454.93 0 1 703 56 'TYPE_A' 4.8] Predicted Sales ===> 412.3625
Test Data : [1478586 1 89 2126 487.0 485.0 0 0 703 56 'TYPE_A' 4.8] Predicted Sales ===> 68.18612
Test Data : [1092935 1 89 2826 341.44 342.44 0 0 703 56 'TYPE_A' 4.8] Predicted Sales ===> 129.0887
Test Data : [1090744 1 89 1754 284.27 283.27 0 0 703 56 'TYPE_A' 4.8] Predicted Sales ===> 300.8519
```

We can see predicted sales for that week

CONCLUSION

Our research focused on predicting meal sales with high accuracy using various machine learning and deep learning models. By merging datasets and analyzing features, we examined sales patterns across different centers, regions, and weeks. We pre-processed data and fine-tuned models, evaluating Random Forest, Gradient Boosting, LSTM, and more. Notably, the CNN2D algorithm excelled, achieving a Mean Absolute Error (MAE) of just 9, outperforming traditional models. This underscores the power of convolutional neural networks in sales forecasting, highlighting the significance of advanced deep learning methods for precise predictions, optimal resource allocation, and enhanced decision-making in the food industry.

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,'' Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.

[7] S. Yadav, T.-M. Choi, S. Luthra, A. Kumar, and D. Garg, ''Using Internet of Things (IoT) in agri-food supply chains: A research framework for social good with network clustering analysis,'' IEEE Trans. Eng. Manag., vol. 70, no. 3, pp. 1215–1224, Mar. 2023, doi: 10.1109/TEM.2022.3177188.

[8] 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.

[9] 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.

[10] Y. Zhang, L. Wang, X. Chen, Y. Liu, S. Wang, and L. Wang, ''Prediction of 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] I. Shah, H. Bibi, S. Ali, L. Wang, and Z. Yue, ‘‘Forecasting one-dayahead electricity prices for Italian electricity market using parametric and nonparametric approaches,’’ IEEE Access, vol. 8, pp. 123104–123113, 2020.
[16] 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.
[17] E. S. Gardner Jr., ‘‘Exponential smoothing: The state of the art,’’ J. Forecasting, vol. 4, no. 1, pp. 1–28, 1985.
[18] C. C. Holt, ‘‘Forecasting seasonals and trends by exponentially weighted moving averages,’’ Int. J. Forecasting, vol. 20, no. 1, pp. 5–10, 2004.
[19] G. E. P. Box and G. M. Jenkins, Time Series Analysis: Forecasting and Control. San Francisco, CA, USA: Holden Day, 1970.
[20] 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.