International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Improving Food Demand Forecasting through Stacked Regression and Hybrid Machine Learning Methods

Naresh Kumar*[1], Dr. Chitresh Banerjee[2], Dr. Minu Bala[3]

Submitted:14/03/2024 **Revised**: 29/04/2024 **Accepted**: 06/05/2024

Abstract: Food demand forecasting is crucial in modern disaster management, particularly amid challenges magnified by the COVID-19 pandemic. The pandemic emphasized the need for accurate forecasting, especially with increased food deliveries and demand surges. Proactive measures are essential to mitigate potential food shortages and prevent losses. Integrating blockchain and Machine Learning (ML) algorithms enhance forecasting accuracy. This study evaluates the efficacy of different ML algorithms - in food demand forecasting, utilizing a monthly ration consumption dataset of 20 districts of Jammu & Kashmir UT. Data pre-processing techniques, including handling missing values, outlier removal, and data normalization, were employed. Additionally, feature selection approaches identified the most relevant predictors for demand forecasting. Our aim is to create a predictive model that achieves high accuracy and efficiency through the integration of innovative machine learning methods and hybridizing current methodologies. Our tests result show that Stacking Regression, which combines three different algorithms (Random Forest, Support Vector Regression, and Ridge), performs better than other methods at predicting outcomes with a lower Root Mean Squared Error (RMSE) score, which indicates how close our predictions are to the actual values. After Hyperparameter tuning, the best RMSE score obtained was 0.11252 using the Ridge model. However, the Stacking Regression model resulted the best RMSE score of 0.10945. This shows that Stacking Regression emerges as the optimal algorithm for demand forecasting within JkBFMs: Blockchain food supply chain management systems, offering simplicity, interpretability, and efficiency in decision-making.

Keywords: *Machine Learning, Food Demand Prediction, Stacking Regression, Root Mean Squared Error, Mean Absolute Error*

1. Introduction

Food demand forecasting, integral to the Blockchain food supply chain management system, has evolved alongside advancements in related fields[1]. These prediction methods empower authorities to proactively address and mitigate the impacts of food-related issues on communities and infrastructure. In modern society, food prediction models have emerged as indispensable assets in disaster management, owing to their ability to anticipate and prepare for potential food-related crises. Advancements in technology and data science have propelled the refinement of these models, enhancing their sophistication and accuracy in evaluation of blockchain food supply chain management system [2]. The development of food prediction models is deeply linked with understanding and managing multifaceted factors within the food supply chain system. Through comprehensive analysis and modeling, food prediction models provide accurate forecasts of food events, aiding authorities in making informed decisions and implementing interventions. This study explores machine learning techniques for forecasting food, utilizing data from January 2021 to December 2023. Insights gained will highlight the effectiveness of machine learning in improving food prediction, thereby advancing disaster preparedness and response strategies.

The pandemic has highlighted the importance of accurate demand forecasting, especially amidst the surge in demand and the need for efficient deliveryof food grains to doorsteps [3]. Addressing the detrimental effects of potential food shortages in specific areas necessitates proactive measures to minimize damages and prevent loss of food. In this dynamic landscape of food supply management, the integration of blockchain and Machine Learning (ML) algorithms holds significant promise for enhancing demand forecasting accuracy [4]. This study evaluates the efficacy of different Machine Learning algorithms (Linear Regression (LR), Decision Tree (DT), Support Vector Regression (SVR), Random Forest (RF), Ridge and Lasso) in food demand forecasting. The aim of paper is to introduce a prediction model of Stacking Regression to predict consumption of monthly ration.

1.1. Motivation and Contributions

1.1.1. Motivation:

Accurate food demand forecasting is crucial for the efficient functioning of blockchain-based food supply chain management systems and plays a vital role in disaster management by anticipating food-related crises.

[1] *Amity Institute of Information Technology, Amity University, Jaipur, Rajasthan, India*
ORCID ID:0009-0005-3735-7438
[2] *Amity Institute of Information Technology, Amity University, Jaipur, Rajasthan, India*
ORCID ID:0000-0002-2303-9894
[3] *SPPND Govt. Degree College,Samba, Jammu & Kashmir, India*
ORCID ID:0000-0002-3359-6715
* *Corresponding Author Email: nareshfeb1990@gmail.com*

Developing robust food prediction models involves comprehensive analysis and management of various supply chain factors. This study examines the effectiveness of machine learning techniques, using data from January 2021 to December 2023, in enhancing food prediction accuracy. Addressing potential food shortages requires proactive strategies to minimize damage and prevent food loss. Integrating blockchain technology with machine learning algorithms holds significant potential for improving demand forecasting accuracy. The study evaluates various machine learning algorithms, including Linear Regression, Decision Tree, Support Vector Regression, Random Forest, Ridge, and Lasso, and introduces a Stacking Regression model for predicting monthly ration consumption. To address limitations and challenges of the Public Distribution System (PDS), especially highlighted by the COVID-19 pandemic [5].

1.1.2. Contributions:

- The study involves training machine learning models using six algorithms: Linear Regression (LR), Decision Tree (DT), Support Vector Regression (SVR), Random Forest (RF), Ridge, and Lasso.
- The dataset comprises over 12,320 records from 20 districts in Jammu & Kashmir, categorized by months, districts, food items, schemes, ration card counts, entitled quantity, opening and closing balances, and monthly consumptions.
- Models were assessed using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), with Ridge achieving the lowest RMSE.
- A Stacking Regression approach was tested, combining different algorithms. The combination of RF, SVR, and Ridge yielded the best RMSE score (0.10945).
- The research demonstrated that stacking regression outperforms standard algorithms in food demand forecasting, improving the RMSE score and thereby enhancing the accuracy of food supply chain management.

1.2. Structure of Paper

The paper is structured as follows:

- The Section I introduces Blockchain Supply Chain Management systems and Machine Learningbackground information.
- In Section II, we review the literature on demand forecasting.
- Section III describes and analyzes the dataset and explains how we calculate the features used for prediction. It also covers the data pre-processing techniques.
- Section IV the forecasting techniques and the evaluation metrics are presented.
- Section V compares the performance of the proposed models.
- Finally, in Section VI, we conclude with the study's results, highlighting that the Stacking Regression model achieved the best performance with an RMSE score of 0.10945 in food demand forecasting.

2. Related Work

The existing literature has explored a wide range of machine learning (ML) algorithms, as indicated by various surveys. However, most of these algorithms have been developed and tested primarily in industrialized nations, leading to limited application in addressing challenges within developing countries. A key area of focus has been food demand prediction, where researchers have applied different ML techniques. Several studies have utilized models such as Linear Regression, Lasso, Ridge, Decision Tree, Random Forest, SVM, and Hybrid Regression. Hybrid Regression, which combines various algorithms into a single model, has gained popularity by sectioning the dataset for each algorithm and then merging the outcomes to create the final model [6].

S. Yadav et al. emphasized the importance of accurate predictions to prevent waste, especially in Agri-Food Supply Chains where IoT technologies are used; without precise forecasts, these technologies may not perform effectively [7]. J. Zheng et al. designed an iterated greedy algorithm to address food preparation time in online food delivery [8], while Y. Zhang et al. used ensemble learning to predict wheat production at a national level[9]. Li Zhiyu et al. analyzed multidimensional time series data to predict fresh produce sales [10]. S. K. Panda Tarallo et al. highlighted the advantages of ML models over traditional forecasting methods [11], and Krishna et al. found that boosting algorithms outperformed others in their comparative study [12]. Hewamalage et al. conducted an empirical study on Recurrent Neural Networks (RNNs) for forecasting, noting that RNNs can directly model seasonality if patterns are uniform; otherwise, and deseasonalization is needed. The LSTM network, a type of RNN, is effective for modeling long-term connections in data [13]. K. Lutoslawski used a hybrid model combining nonlinear autoregressive exogenous with neural network (NARXNN) for time series prediction [14]. J. Kim and N. Moon demonstrated that Bi-LSTM's performances on multivariate time series showed neural network-based methods are superior to statistical methods for data with nonlinear trends [15]. H. D. Nguyen et al. discussed two key problems in supply chain management: forecasting sales and detecting anomalies. They suggested using LSTM-

based methods for multivariate time series forecasting and combining LSTM Auto encoder with SVM for anomaly detection [16].

Researchers aim to enhance accuracy by combining these approaches in a hybrid manner. Some studies have used a combination of descriptive and predictive analytics, such as Saadat et al., who predicted computer manufacturing test failure patterns to reduce costs and speed up manufacturing operations [17].

Table 1. A summary of the above discussion is tabulated

Authors	Model Study	Methods Applied
S. Yadav et al.	IoT Based	Data clustering
J. Zhang et al.	Iterated greed algorithm	sampling method
Y. Zhang et al.	Ensemble model	Regression Matrix
Li Zhiyuet al.	ARIMA-LSTM	Regression Matrix
S. K. Panda	Machine Learning	Regression Matrix
Krishna et al.	Regression model and Boosting Techniques	Regression Matrix
Hewamalage et al.	Recurrent Neural Networks (RNN)	Regression Matrix
K. Lutoslawski	Nonlinear Autoregressive Exogenous Neural Network (NARXNN)	Regression Matrix
J. Kim and N. Moon	Bi-LSTM Model	Regression Matrix
H. D. Nguyen et al	LSTM	Classification Matrix
Saadat et al.	Hybrid (SVM and Random Forest)	Classification Matrix

3. Data Analysis and Pre-Processing

In this section, we detailed the dataset features, data collection process, and applied pre-processing techniques.

3.1. Data Set

The dataset containing 20 District of Jammu & Kashmir in which more than 12320 records both in the training dataset and testing dataset, each characterized by 10 parameters. It includes data categorized by Months, Districts, Fooditems, Schemes, Ration Card, RC Counts (Ration Card Count), Entitled Quantity, Opening Balance, Closing Balance and Monthly Consumptions.

3.2. Data Collection

In this study, data for independent features like Months, Districts, Fooditems, Schemes, RC Counts, Entitled Quantity, Opening Balance, Closing Balance and Monthly Consumptions were collected from different sources in a ".csv" file. This data is used to predict the dependent feature, Monthly Consumptions.

3.3. Data Pre-Processing

Data Cleaning & Transformation: Data cleaning is a crucial step in pre-processing, involving the removal of undesired data and rectifying missing or NA values. Additionally, inaccurate and outlier data points that may affect prediction models are eliminated.

The chosen dataset underwent pre-processing to handle null values. Imputation was used to address missing data by estimating the missing values. Mean/Mode imputation replaces missing data with the mean or mode of known values, suitable for numerical and categorical variables respectively, assuming Missing Completely at Random (MCAR). Z-score identifies and removes outliers by quantifying the deviation of data points from the mean [18].

$$Z = \frac{x-\mu}{\sigma} \quad (1)$$

Where:

- xis the original data point or variable.
- μ represents the mean of the dataset
- σis the standard deviation of the population or sample.

When x exceeds μ the Z-score becomes positive, suggesting that the data point surpasses the mean. If x is lesser than μ, the Z-score turns negative, indicating that the data point falls below the mean. When x equals μ, the Z-score amounts to zero, signifying that the data point aligns precisely with the mean.

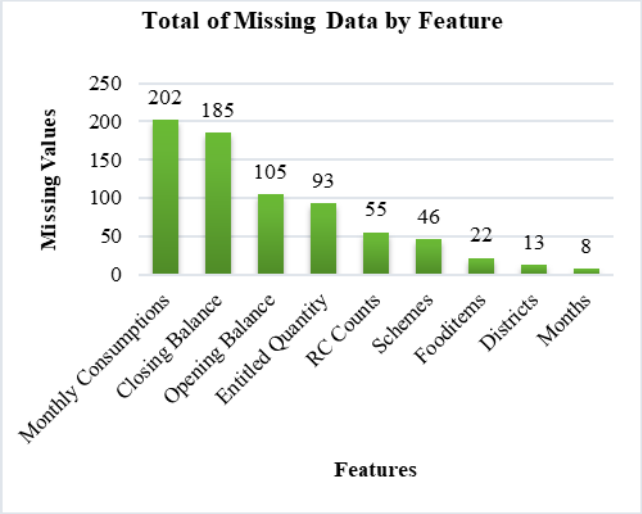

Fig. 1. Missing Values on Data.

3.4. Feature Selection

The dataset containing 20 District of Jammu & Kashmir in which more than 12320 records both in the training dataset and testing dataset, each characterized by 10 parameters. It includes data categorized

Table 2. KBest Results

Test Case	Number of Features	RMSE Score
1	3 Feature	0.15621
2	4 Feature	0.14322
3	5 Feature	0.13221
4	7 Feature	0.13990
5	9 Feature (Without Filtering)	0.12336

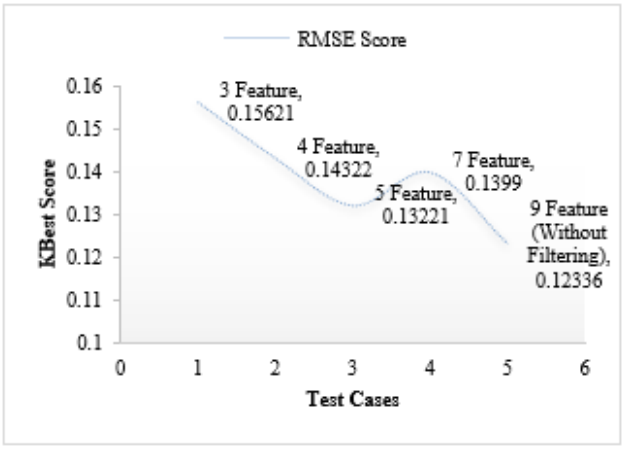

Fig. 2.KBest RMSE score based on features.

In the analysis, every feature demonstrates its importance in predicting the target feature, implying that keeping all features intact is more favourable than discarding any. The results from the KBest score test, as shown in Fig. 2, indicate that the RMSE score performs better when no feature filtration is conducted as opposed to when certain features are removed.

3.5. Data Splitting

We split the dataset into a training set and a test set, maintaining a ratio of 75:25, resulting in 9240 samples for training and 3080 samples for testing. Additionally, we employed KFold cross-validation in the modeling process. KFold divides the dataset into k subsets and iteratively uses one subset for validation while the rest are used for training. In this instance, we set k to 5, resulting in the dataset being split into 5 subsets, with 4 subsets used for training and 1 subset for testing in each iteration. This process is repeated k times with different combinations of training and test sets. The choice of the number of folds is not standardized and may vary depending on the specific requirements of the analysis.

$$PERFORMANCE = \frac{1}{K}\sum_{k=1}^{N} Metric_k \quad (2)$$

Where:

- K is the number of folds (e.g., 2-fold, 20-fold).
- $Metric_k$is the performance metric (e.g., accuracy, error) obtained for the kth fold during validation.

Table 3. Best Fold Number

Number of Fold	RMSE Score
2	0.12641
5	0.12622
10	0.12521
15	0.12362
20	0.12311
25	0.12324

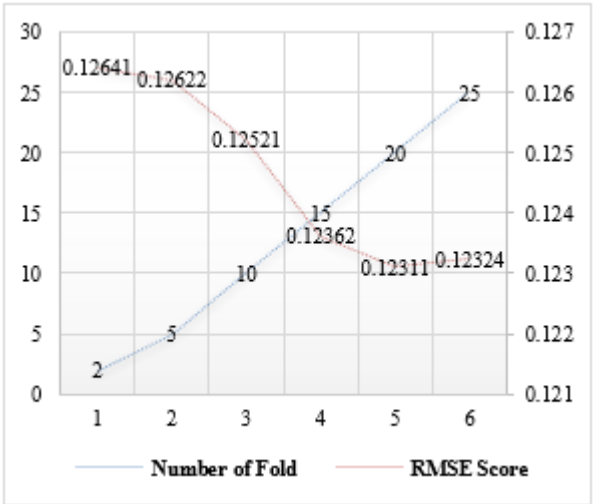

Fig. 3. Best Fold Numbers RMSE Score.

As the results of the Fig. 3 results indicate that when we look at the best folds based on their RMSE score, both using 20 folds and 25 folds yield similar scores. However, using 20 folds actually gives us a slightly better score than using 25 folds. We stopped the test at 25 folds because it takes too much time to process when we test with more than 25 folds.

4. Methodology and Evaluation Metrics

To create an effective predictive model, we can choose from a range of Machine Learning methods available today [19][20]. We considered the requirements of the main goal of the paper and related research, and decided to use the most commonly used regression model i.e. LR, DT, SVR, RF, Ridge and Lasso.The proposed workflow is shown in Fig.4.

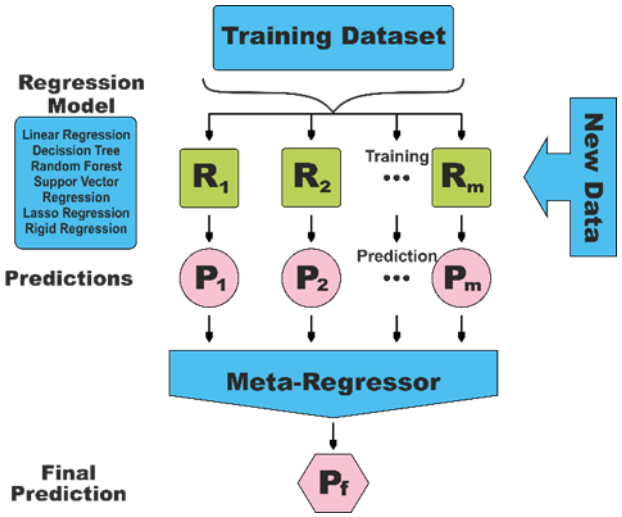

Fig. 4. Proposed Workflow of Stacking Regression.

4.1. Evaluation Matrix

There are several evaluation metrics available to ascertain the accuracy of a model's predictions up to a certain performance threshold [21].

Mean Absolute Error (MAE) represents the average of all the differences between predicted and actual values, without considering the direction of those differences:

$$MAE = \frac{1}{N}\sum_{j=1}^{N}\left|\mathcal{Y}actual_j - \mathcal{Y}pred_j\right| \quad (3)$$

Root Mean Squared Error (RMSE) represents the average of the differences between predicted and actual values, then finding the square root of that average:

$$RMSE = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(\mathcal{Y}actual_j - \mathcal{Y}pred_j\right)^2} \quad (4)$$

- N represents the number of instances or observations.
- $\mathcal{Y}\text{actual}_j$represents the actual value of the dependent variable for the jth data point.
- $\mathcal{Y}\text{pred}_j$represents the predicted value of the dependent variable for the jth data point.

5. Experimental Results

Table 4. Basic Algorithm Modeling Score

Algorithms	MAE Score	RMSE Score
LR	0.12311	0.12652
DT	0.12271	0.12886
SVR	0.13225	0.15345
RF	0.12629	0.13526
Ridge	**0.11252**	**0.11851**
Lasso	0.12324	0.14552

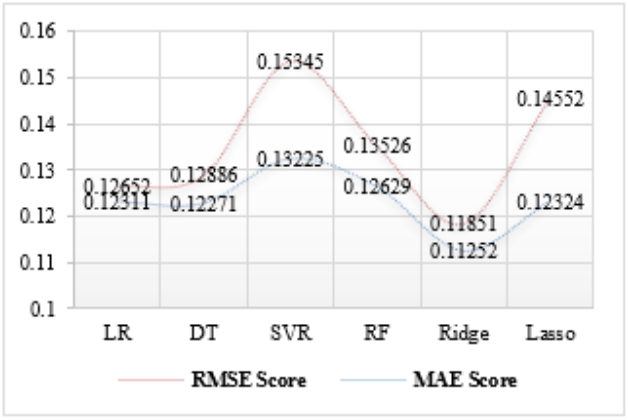

Fig. 5. Basic Modeling RMSE and MAE Score

After trying out different methods to analyze data, we found that the Ridge algorithm performed the best overall, with the lowest RMSE and the best MAE scores. Interestingly, LR, DT and RF algorithms all had similar scores to Ridge. An alternative method involves GridSearchCV, which helps find the best parameter values by exploring a range of choices. It evaluates parameter options to pinpoint the optimal values within the specified range.

$$\text{Hyperparameters} = \arg\max_{\theta}\frac{1}{K}\sum_{k=1}^{K}\text{Metric}_k(\theta) \quad (4)$$

Where:

Hyperparameters are the hyperparameters that maximize the average performance metric across all folds.

- θ represents the hyperparameter combination being evaluated.
- K is the number of folds in cross-validation.
- $\text{Metric}_k(\theta)$is the performance metric (e.g., accuracy, loss) obtained for the kth fold during validation, given the hyperparameter

combination θ.

Table 5. Grid Search CV Hyperparameter and Result

Algorithms	Parameters
LR	• Fit intercept = True • normalize: True • copy_X: True • Splitter = random
DT	• Max_Depth = 10 • Min_Sample_Leaf = 4
SVR	• Max_Depth = 20 • C = 500 • Kernel = rbf
RF	• Estimators = 200 • Min_Sample_Split = 4 • Min_Sample_Leaf = 4 • Max_Depth = 20
Ridge	• Alpha = 11
Lasso	• Alpha = 0.01

After testing various parameter values, the optimal parameters, as indicated in Table 5, were determined. Following this, a hyper-parameter test will be conducted to refine the model algorithms and assess the differences pre and post tuning.

Table 6. Basic Algorithm Compared After Hyperparameter

Number of Fold	RMSE Score Before	RMSE Score After Hyper parameter
LR	0.12652	0.12311
DT	0.12886	0.12271
SVR	0.15345	0.13225
RF	0.13526	0.12629
Ridge	**0.11851**	**0.11252**
Lasso	0.14552	0.12324

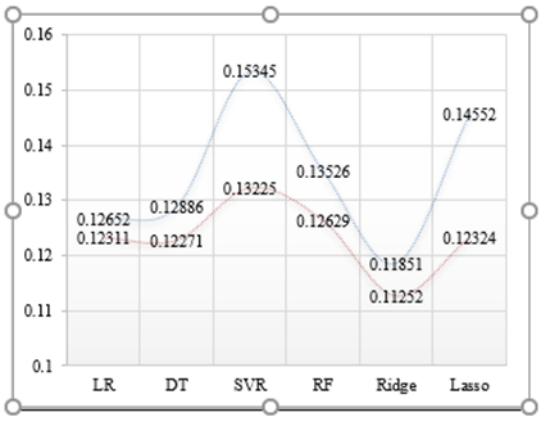

Fig. 6. Comparison of Algorithm Before and After Hyperparameter Testing

After hyper-parameter testing for our model, we found that the Ridge algorithm consistently gives us the lowest RMSE, which is a measure of how well our model predicts outcomes. LR and DTs also perform well, but not as well as Ridge. Fig. 6 summarizes our results and found that Ridge had the best scores across multiple tests. Now, we're moving on to a technique called stacking regression. The initial move in stacking is to find the best secondary model to combine with our main model. Then, we'll figure out which algorithms work best when combined in this stacking process.

Table 7. Finding Meta Regressor

Algorithm	RMSE Score
LR	0.12241
DT	0.12221
SVR	0.11662
RF	0.12322
Ridge	0.11310
Lasso	0.12324

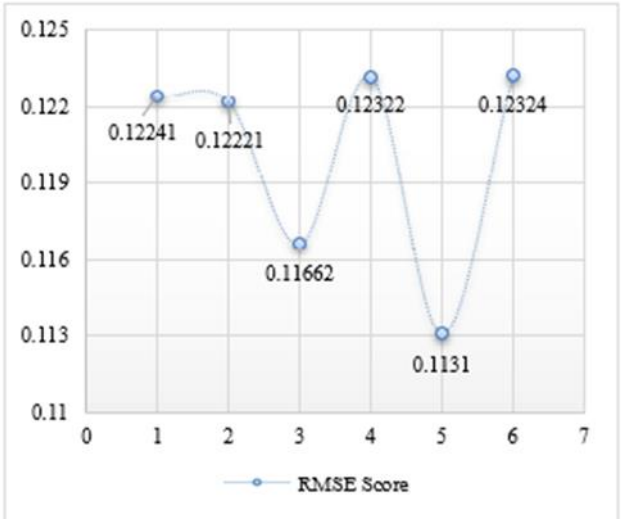

Fig. 7. RMSE Score of Meta Regressor

As shown in Table 7 shows that the ridge algorithm outperforms others in terms of RMSE for basic stacking. The algorithms considered in this research are LR, DT, SVR, RF, Ridge and Lasso. For the upcoming stacking test cases, the ridge algorithm will serve as the meta regressor. To determine the optimal combination of algorithms for stacking regression, several methods will be explored. Initially, algorithms will be categorized based on their method: Linear-Based methods like LR, Lasso, Tree-Based methods such as DT and RF, and Ridge, and Distance-Based methods exclusively including SVR.

Table 8. Stacking Regressor Test Case 1

Combination Algorithm	RMSE Score
Linear-Based	0.11556
Tree-Based	0.11225
Linear-Based + Tree-Based + Distance-Based	0.11325

In the test case 1 of stacking regression experimentation, the results showed in table 8 that linear-based algorithms as input led to a better Root Mean Square Error (RMSE). In the test case 2 of stacking regression, we attempted various combinations by initially incorporating all available algorithms and subsequently eliminating them one by one to identify the optimal combination.

Table 9. Stacking Regressor Test Case 2

Combination Algorithm	RMSE Score
LR + SVR + Ridge + Lasso	0.11312
RF + SVR + Lasso	0.11212
DT + Ridge + LR + SVR	0.11225
Lasso + SVR + RF +Ridge	0.11328
RF + Ridge + SVR + LR	0.10994
SVR +RF + Ridge	########
SVR + Ridge	0.11844

Table 9 showing the outcome, obtained as the best result of 0.10954, was achieved through a combination of algorithms, including RF, SVR, and Ridge, during the test case 2 of stacking regression.

6. Conclusion and Future Direction

6.1. Conclusion

In our research, we found that stacking regression performed better than standard algorithms when measuring accuracy with RMSE Score. After Hyperparameter tuning, the best RMSE score we got with the Ridge algorithm was 0.11252. However, when we used stacking regression, which involves combining different algorithms, we tested various combinations and achieved a better RMSE score of 0.10945, the best in food demand forecasting.

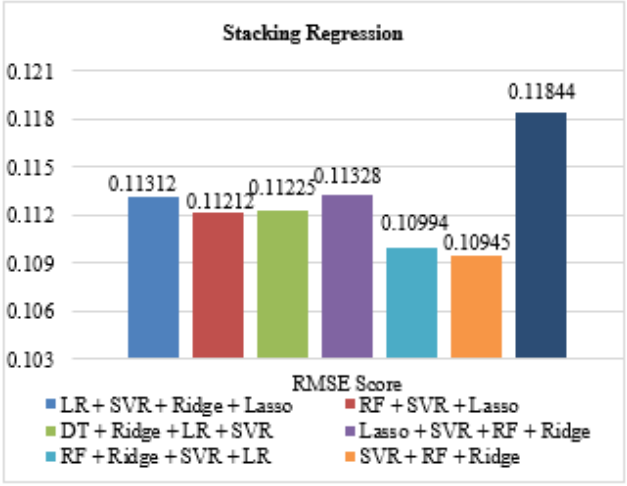

Fig. 8. Stacking Regression Results on Data

Our result shows in Fig. 8. that the Stacking Regression model, which uses SVR, RF, and Ridge as input algorithms, performed better than other models, resulting in a lower RMSE score. Compared to what other studies have found, our research successfully improved the RMSE score by using this stacking approach. This method combines the strengths of different algorithms, leading to better overall performance of food supply chain management system for food demand forecasting.

6.2. Future Directions

Future scope for this research will focus on exploring adaptability of the suggested algorithm through experimentation with various components. While our existing algorithm draws from traditional machine learning methodologies, there exists an opportunity to integrate neural networks into the stacking framework as it holds promise for offering high flexibility and productivity of the model.

Acknowledgment

I would like to express my heartfelt gratitude to my research supervisor Dr. Chitresh Banerjee, Associate Professor, Amity Institute of Information Technology, Amity University, Jaipur, Rajasthan and Dr. Minu Bala, Associate Professor, SPPND Govt. Degree College, Samba, Jammu & Kashmir for their guidance and unwavering support during my research.

Author contributions

Naresh Kumar: Conceptualization, Methodology, Software, Field study, Writing-Original draft preparation**Dr. Chitresh Banerjee:** Data curation, Software, Validation, Visualization, Investigation **Dr. Minu Bala:** Visualization, Investigation, Writing-Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Naresh Kumar, C. Banerjee, Minu Bala, "JkBFMs: blockchain food supply chain management system", *In: Proc. Of AdvIntellApplInnov Approach,* 10(1063/5):0148937 (2023).

[2] Kumar, N., Banerjee, C. & Bala, M. "Performance evaluation of blockchain food supply chain management system: JkBFMs", *Innovations SystSoftwEng,* (2024). https://doi.org/10.1007/s11334-024-00560-7

[3] M. H. D. M. Ribeiro, R. G. da Silva, V. C. Mariani, and L. dos Santos Coelho, "Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil," *Chaos, Solitons& Fractals, vol. 135, p. 109853, 2020*.

[4] Feizabadi, J.,"Machine learning demand forecasting and supply chain performance."*International Journal of Logistics Research and Application(2022).*

[5] F. Rustam et al., "COVID-19 future forecasting using supervised machine learning models," *IEEE access, vol. 8, pp. 101489-101499, 2020.*

[6] Behrooz Keshtegar, Aliakbar Gholampour, DucKien Thai, Osman Taylan , "Hybrid regression and machine learning model for predicting ultimate condition of FRP-confined",*concrete: Composite Structures(2021).*

[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] 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*

[10] Li Zhiyu1, Yu Wei1, Zhu Wenwei1, Wan Haojie1, PengJingjing, Wang Hui, "Research on Fresh Produce Sales Prediction Algorithm for Store Based on Multidimensional Time Series Data Analysis",*EAI Endorsed Transactionson Scalable Information Systems (2023)*

[11] S. K. Panda and S. N. Mohanty, "Time Series Forecasting and Modeling of Food Demand Supply Chain Based on Regressors Analysis", *IEEE Access, vol. 11, pp. 42679-42700, 2023, doi: 10.1109/ACCESS.2023.3266275*

[12] A. Krishna, A. V. A. Aich, and C. Hegde, ''Sales-forecasting of retail stores using machine learning techniques'',*In Proc. 3rd Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solutions (CSITSS), Dec. 2018, pp. 160–166.*

[13] H. Hewamalage, C. Bergmeir, and K. Bandara, ''Recurrent neural networks for time series forecasting: Current status and future directions,'' *Int. J. Forecasting, vol. 37, no. 1, pp. 388–427, 2021*.

[14] K. Lutoslawski, M. Hernes, J. Radomska, M. Hajdas, E. Walaszczyk and A. Kozina, "Food Demand Prediction Using the Nonlinear Autoregressive Exogenous Neural Network", *IEEE Access, vol. 9, pp. 146123-146136, 2021, doi: 10.1109/ACCESS.2021.3123255*

[15] J. Kim and N. Moon, ''BiLSTM model based on multivariate time series data in multiple fields for forecasting trading area'',*J. Ambient Intell. Hum. Comput., pp. 1–10, Jul. 2019.*

[16] H. D. Nguyen, K. P. Tran, S. Thomassey, and M. Hamad, ''Forecasting and anomaly detection approaches using LSTM and LSTM autoencoder techniques with the applications in supply chain management'',*Int. J. Inf. Manage., vol. 57, Apr. 2021, Art. no. 102282*

[17] Saadat R, Syed-Mohamad SM, Azmi A, et al, "Enhancing manufacturing process by predicting component failures using machine learning", *Neural Computing and Applications 34(20):18155–18169*

[18] VaibhavAggarwa,Agarwal, V. Gupta, V. Singh, P. Sharma, K., & Sharma, N., "Detection of spatial outlier by using improved Z-score test," *Proc. of the International Conference on Trends in Electronics and Informatic*s,(2019)

[19] Vithitsoontorn, C. and Chongstitvatana, "Demand forecasting in production planning for dairy products using machine learning and statistical method", *International Electrical Engineering Congress (2022).*

[20] E. Tarallo, G. K. Akabane, C. I. Shimabukuro, J. Mello and D. Amancio, "Machine learning in predicting demand for fast-moving consumer goods: Exploratory research", IFAC-PapersOnLine, vol. 52, (2019).

[21] Davide Chicco, Matthijs, J. Warren,"The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation", *Peerj Computer Science(2021)*