I.J. Information Technology and Computer Science, **2025, 3, 97-105**
Published Online on June 8, 2025 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijitcs.2025.03.07

Forecasting Agriculture Commodity Price Trend using Novel Competitive Ensemble Regression Model

R. Ragunath
Department of Computer Science, Periyar University, Salem-636 011, India
E-mail: rrcspu@gmail.com
ORCID iD: https://orcid.org/0000-0003-4992-5947

R. Rathipriya*
Department of Computer Science, Periyar University, Salem-636 011, India
E-mail: rathi_priyar@periyaruniversity.ac.in
ORCID iD: https://orcid.org/0000-0002-3970-262X
*Corresponding author

Received: 08 October 2024; Revised: 07 January 2025; Accepted: 18 February 2025; Published: 08 June 2025

Abstract: This paper introduces a novel approach for forecasting the price trends of agricultural commodities to address the issue of price volatility faced by both farmers and consumers. The accurate forecasting of food prices is particularly crucial in emerging nations such as India where food security is a top priority. To achieve this goal, the paper presents an ensemble learning-based approach for predicting the agricultural commodity price (ACP) trend. Using dataset namely rainfall and wholesale pricing index (WPI), the study compares the performance of various individual and ensemble regression models. The findings of this work demonstrated that the novel competitive ensemble regression (CER) approach outperforms traditional individual regression models in predicting price fluctuations trend accurately. This approach has the high potential and more precise prediction to afford farmers and dealers, also make the model suitable for the financial industries.

Index Terms: Agricultural Commodity, Wholesale Pricing Index, Competitive Ensemble, Ensemble Regression Models, Price Trend Prediction.

1. Introduction

Most of the Indian farmers sustain the severe losses due to a lack of knowledge in the market field. This research information gap often upgrades them to sell their agricultural products at prices subsequent in foremost monetary losses. For making solution for this information gap is dangerous for empowering farmers to be more educated and making decision profitable in the market field. Several elements can inspiration for the prices of agricultural product. These include i) various condition based on climatic, ii) insect prevalence, iii) risks in production, iv) worldwide demand and supply, various government policies, and factors of economic. The combination of these various elements commonly causes the agricultural prices to fluctuate suggestively [1-3]. While handling the price issues for forecasting the accurate and reliable is critical and making detailed decisions about selling the agricultural products. In the previous ACP prediction models uses a variety of predictive models such as linear and nonlinear. Machine learning has recently emerged as a better solution in this area. It provides enhanced features that improve the accuracy of ACP predictions. [4].

Machine learning (ML) techniques is one of the predictive analytic techniques in past decade. Most of the literature's works have established the advantage over time series models for financial assets using prediction [5]. Most of the researchers mainly used three popular machine learning algorithms such as first one is Artificial neural networks (ANN), second is Generalized neural networks (GRNN), and the third is Support vector regression (SVR).

Data-driven models such as i) Random forests (RF), ii) Gradient boosting machines (GBM), and iii) other nonparametric approaches used to find the stochastic correlations in the dataset [6]. Box-Jenkins processes and regression models are well-thought-out techniques for conventional statistics to be less effective than ANNs [7].

Latest studies presented the Deep learning (DL) and ML algorithms provides the most effective solutions for problems for prediction [8, 9]. Neural networks outclassed statistical approaches for ACP [10]. The predictive models such as i) gradient boosting decision tree (GBDT), ii) MLR, iii) RF, iv) GPR, v) Lasso, K-nearest neighbor (KNN), and vi) SVR have also been successful for improving agricultural product for prediction of accuracy. However, these kinds of models have some limitations [11-13].

The prevailing prediction techniques for forecasting agricultural commodity data commonly rely on particular models, which is a remarkable disadvantage. To address this problem, ensemble learning approaches, which provides the increasingly use of variety of fields in the agricultural research, have arisen. Ensemble learning helps to combine many learners to generate as well as a more powerful and provide complete predictive model [14-17]. The basic concept is that several base learners can work together to correct errors, compensating for less accurate predictions provided by individual learners. Boosting, bagging, and stacking algorithms are common approaches to ensemble learning. This paper describes a novel ACP prediction method based on competitive ensemble regression (CER) learning that improves the accuracy and reliability of agricultural commodity forecasts. This CER model uses historical ACPdata and monthly rainfall to predict future ACP trends.

The remaining part of this research work is planned in this fashion: Section 2 provides an overview of the existing state of research on agricultural commodities and rainfall. Section 3 details the regression methods and their evaluative metrics used in this work. Section 4 describes the proposed CER methodology. The results of the CER model are presented and discussed in Section 5. Section 6 presented the summary of the research work.

Table 1. Literature review based on machine learning models

Author and Year	Commodity	Models Explored	Model Category	Measures
Yuzhen Zhang et al., [18]	Crop yield	Bagging, Boosting, Stacking, RF, ERT, XGBoost, LightGBM, CatBoost	Regression	RMSE,
Andres M. Ticlavilca et al., 2010 [19]	Corn, Cattle, Hogi	ANN and Multivariate Relevance Vector Machine	Regression	RMSE, RMSPE
Darra, N. et al., 2023 [20]	Tomato	ARD regression, SVR	Ensemble Regression	RMSE, R2
Yan Guo, et al., 2022 [21]	Corn	LR, RF, XGBoost, LightGBM, LSTM, AttLSTM, AttLSTM-ARIMA-BP	Regression	MAPE, RMSE, MAE, R^2
Yeong Hyeon Gu, et al., 2022 [22]	Radish, Cabbage,	GCN-LSTM, LSTM, DA-RNN, STL-AttLSTM, DIA-LSTM	Time Series Models	RMSE, MAPE
Isaac Kofi Nti, Adib Zaman, et al., 2023 [23]	Crop Yield	Various tree-based ensemble learning models	Classification	Accuracy, recall, precision, F1-score
Lisha Mao, et al., 2022 [24]	Cabbage, Carrot	ARIMA, NAIVE	Big data and Machine learning techniques	MSE, RMSE, MAE, MAPE
Anil Kumar Mahto, et al., 2021 [25]	Soybean and Sunflower seeds	ARIMA, MA, ARMA, ANN, SVM	Regression	MAPE, RMSPE
Xiang Wang, et al., 2022 [26]	Hog	LightGBM, XGBoost, GBDT, WOA, CEEMDAN, SVR, EEMD	Regression and Ensemble	MSE, MAE, RMSPE, MAPE, R^2, DA
Ranjit Kumar Paul, et al., 2022 [27]	Brinjal	ARIMA, SVR, RF, GRNN, GBM	Neural Network and Regression	ME, RMSE, MAE, MAPE
Neeraj Soni, Jagruti Raut 2022 [28]	Crop	RF, DT	Regression	RMSE
M. C, Dhanraj RK, 2023 [29]	Soil dataset and crop production dataset	Stacking-based ensemble deep learning (MAML)	Classification	MAPE, ACC, IR
N. Nabipour, et al., 2020 [30]	Petroleum, Non-metallic minerals	DT, Bagging, RF, Adaboost, GB, XGBoost, ANN, RNN, LSTM	Ensemble	MAPE, MAE, RMSE, MSE
Zhiyuan Chen [31]	Chicken, Chilli, Tomato	ARIMA, SVR, Prophet, XGBoost, LSTM	Regression	MSE
H.P. Suresha, et al., 2022 [32]	Wholesale prices (tomato, onion, potato)	ARIMA, RNN, LSTM, VAR, RFR, XGBoost	Regression and Ensemble	RMSE
Ahmed A. Adeniran, et al., 2019 [33]	Carbonate reservoir	ANN, SVM, KNN and NCC	Ensemble	RMSE, CC
A.H. Utomo, et al., 2022 [34]	Weather data and agricultural commodity prices	PNN	Big Data	Accuracy
Sussy Bayona-Ore, et al., 2021 [35]	Agricultural commodities	BPNN, GA, ANN-GA, RBFNN, SVM, ARIMA	Neural Network	MAE, MAPE, MSE, RMSE, MAD, R^2, SMAPE, MIV, SPA
Mohsen Shahhosseini, et al., 2022 [36]	Multiple Dataset	Bagging, boosting, and stacking/blending	Ensemble	RMSE, MAE, MAPE, MSE,
Thomas van Klompenburg, et al., 2020 [37]	Temperature, rainfall, and soil	CNN, DNN, LSTM	Classification	Accuracy

2. Related Works

The major aim of this section is to provide a summary of various studies that have explored models for predicting agricultural commodity prices or related factors such as crop yield and weather data. Table 1 presents an extensive range of topics including a ML models for prediction of commodity price. It includes a wide-ranging of predictive models that examine the several commodities for prediction and novel methodology, such as one is ensemble learning and second one is stacking-based ensemble DL. In addition, some studies undertake valuable relative assessments, providing information about the comparative strengths and limitations of several prediction approaches. The study found that those limitations, such as varying assessment criteria and limited hyperparameter evidence in specific models. Also, highlight the benefits of ensemble approaches for prediction of commodity price. As it overcomes the limitations of independent models, it is favoured choice for better-quality accuracy in forecasting ACP.

3. Methods and Materials

This section provides the necessary background information to understand this work

Model	Description and Application
Support Vector Regression (SVR)	SVR optimizes the cost function, fitting data points with the straightest line (or plane), utilizing the kernel trick for effective non-linear regression through implicit input transformations into high-dimensional feature spaces.
Gaussian Process Regression (GPR)	It uses a Bayesian approach, inferring a probability distribution over potential functions fitting the data. It employs a Gaussian process indicated as a multivaiate Gaussian distribution as a prior.
Decision Tree Regression (DTR)	These models learn from data by creating decision rules on variables to segregate classes in a tree-like structure. They are versatile and can be applied to both classification and regression tasks.
Random Forest Regression (RFR)	These models implement an ensemble of decision trees, deriving results from majority votes among the decision trees. They are effective tools for both classification and regression tasks.
Extreme Gradient Boosting Regression (XGBoost Regression)	XGBoost Regression is an ensemble-based regression method that pools weak predictive learners to enhance accuracy. Popular techniques within this category include XGBoost, LightGBM, and others.
Grid Search	It's an optimization algorithm called grid search for hyperparameters; it essentially demands that you supply a list of values for every hyperparameter that you want to tune and train models for all those combinations of the values.
Random Search	Random search optimizes hyperparameters by sampling values from finite distributions instead of fixed lists. It trains the model multiple times with random combinations and selects the best one based on performance. This method is efficient for large hyperparameter spaces.

A. Performance Measure

This section describes the mathematical equations for the performance metrics used in this work. Here, actualYData is the actual target value in the dataset whereas predictedYData is the predicted value of the target data. Mean Absolute Error (MAE) is the mean difference between the actual and forecasted data. It is denoted as in equation 1.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\text{actualYData} - \text{predictedYData}| \tag{1}$$

Mean Squared Error (MSE) is the mean of the squared difference between the actual and forecasted data of the given dataset. It calculates the residuals' variance as in equation 2.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\text{actualYData} - \text{predictedYData})^2 \tag{2}$$

Root Mean Squared Error (RMSE) is the square root of the MSE. It calculates the residuals' standard deviation as in equation 3.

$$RMSE = \sqrt{MSE} = \frac{1}{n}\sqrt{\sum_{i=1}^{n} (\text{actualYData} - \text{predictedYData})^2} \tag{3}$$

R-squared (R2) measures proportion of the variance in the dependent variable that is predictable from the independent variable(s) in a regression model. It is represented in equation 4.

$$R^2 = 1 - \frac{\sum (\text{actualYData}-\text{predictedYData})^2}{\sum (\text{actualYData}-\text{predictedYData})^2} \tag{4}$$

Mean Absolute Percentage Error (MAPE): The mean of the percentage differences between the actual and predicted values, calculated as a percentage of the actual values. It provides a measure of the accuracy of predictions relative to the actual values of the dataset. It is represented in equation 5.

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \frac{\text{abs (actualYData} - \text{predictedYData)}}{\text{actualYData}} \quad (5)$$

Percentage of Error (PE): The total absolute difference between the actual and predicted values, expressed as a percentage of the actual values. It indicates the magnitude of error relative to the actual values. It is given in equation 6.

$$\text{Percentage of Error (PE)} = \frac{\sum_{i=1}^{n} |\text{actualYData} - \text{predictedYData}|}{\sum_{i=1}^{n} \text{actualYData}} * 100 \quad (6)$$

4. Proposed Work

The proposed method has been broken down into several phases, such as data collecting, pre-processing, utilizing regression models to forecast the wholesale price index of crops, and using the roulette wheel selection (RWS) operator to create a competitive ensemble regression model. Figure 1 depicts an ensemble machine learning process for building a competitive ensemble regression model to predict the WPI value. It begins with the splitting of the dataset into a training set and a test set for proper model benchmarking. Various base regressors like SVR, DTR, GPR, RFR, and XGBoost are used, and hyperparameter tuning is used to achieve maximum performance. The optimal models are combined into an ensemble regression model to achieve maximum predictive accuracy. The predictive model, after training, accepts new data and delivers the predicted WPI value. The method delivers strong and stable predictions by utilizing ensemble learning and optimization methods. The key contribution of this research work is briefed as follows. The proposed competitive ensemble regression model is used as the forecasting approach to predict the Indian crop price. It consists of five basic benchmark regression models, and ensemble strategy to achieve more accurate and robust performance under various crop types. The entire process is depicted in the Algorithm – 1.

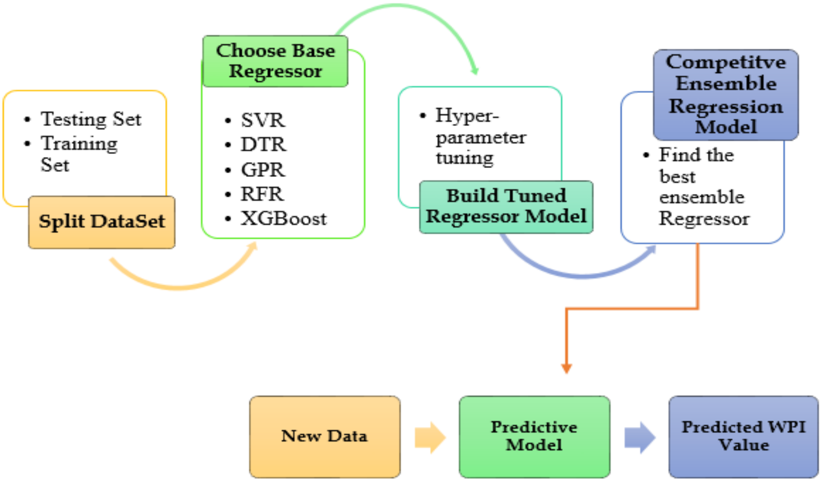

Fig.1. Overall workflow

Algorithm-1: Prediction using CER()

Input: Dataset (Crop_WPI, Rainfall)
Output: Forecasted values(Crop_WPI)
Begin
Split Dataset: Split the WPI dataset into training and testing sets.
Model Training Loop: a. For each base regressor model 'R' in modelList:
a. Train the R model.
b. Predict Crop's WPI values.
c. Calculate R2, RMSE, and PercentError for the training set.
d. Calculate fitness of the training model using the RWS operator (equation 1).
Model Testing Loop: a. For each trained model in trainModelList:
a. Predict testWPI for the crop.
b. Calculate RMSE, R2, and PercentError for the testing set.
c. Calculate fitness of the testing model using the RWS operator (equation 1).
Return Forecasted WPI for each crop
End

4.1. Stage 1: Data Collection and Data Pre-processing

The dataset taken for study is a time series data from 2012 to 2022. It contains four columns Month (represented as a number from 1 to 12), Year, Rainfall (the amount of rainfall received each month in millimeters), WPI (the Wholesale Price Index for the respective months), and data pre-processed for regression modeling. Appropriate imputation techniques were used to substitute missing data in order to address the missing value issue, which is essential for reliable model results.

4.2. Stage 2: Training phase of Prediction Algorithms

The prediction algorithm involves two key stages – training and testing. The dataset is divided into 70% training and 30% testing subsets. Top regressors like random forest (RF), support vector regressor (SVR), and XGBoost regressor are employed. The training phase creates the prediction model using the training dataset, addressing overfitting through performance validation. The trained model is then tested using the dedicated testing dataset, ensuring general performance evaluation. Iterative 10-time cross-validation is performed to identify overfitting issues.

4.3. Ensemble Learning: Competitive Ensemble Regression Model

In the competitive ensemble learning stage, ten algorithms for regressor were employed to forecast the crop's WPI. Optimal hyperparameters are determined through experimental trial-and-error for enhanced performance. Its values are tabulated in the Table 2. To identify the best-fitted regressor model for forecasting, the Roulette Wheel Selection (RWS) operator is used. The population is 'n' regression models. For each regression is represented as 'i' with a corresponding fitness value is denoted as 'fit(i)', the probability is defined as 'p(i)' of selection is figured as equation 7.

$$\text{pi} = \frac{\text{fit}_i}{fit_j} \tag{7}$$

Table 2. Hyperparameters details for base regressor models

Regression Model	Hyperparameters	Optimal Values
Decision Trees Regression (DTR)	max_depth	Integer: 10
	min_samples_split	Integer: 10
Gaussian Process Regression (GPR)	kernel	RBF
	regularization	Float: 10.0
Support Vector Regression (SVR)	kernel	RBF,
	regularization (C)	Float: 10.0
Random Forest Regression (RFR)	n_estimators	Integer: 200
	max_depth	Integer: 10,
XGBoost Regression (XGB)	learning_rate	Float: 0.3
	max_depth	Integer:10

Here, p(i) is called as the probability that regression model is defined as the 'i' will be selected, f(i) is represented as fitness, and fit(j) represents the sum of fitness values for all regressors in given the population is given equation 8.

$$\begin{gathered}\text{Fitness function (f)= argmin(g(x)),}\\ \text{where g(x)= (RMSE-MIN(RSME) / MAX(RMSE)-MIN(RMSE)+R2)/2.}\end{gathered} \tag{8}$$

This proposed fitness function is a combination of two evaluation measures, normalized RMSE and R2, for accurate evaluation of prediction models. This function gives equal importance to both the accuracy and goodness-of-fit of the predicted values.

5. Observational Studies

The experiments aimed to evaluate the performance of the proposed CER model using datasets taken from reliable sources: the official government website for commodities wholesale prices[1] and monthly rainfall[2] in India. Following significant preprocessing and cleaning process, data were fitted and trained with the base regression models. The two distinct ensemble learning methodologies namely classical ensemble regression model and competitive ensemble regression model were tested using the aforementioned datasets. The RMSE measure was used in the study to determine which predictive models performed best and which were least successful. In the crop dataset Ragi, Maize, Barley, Paddy, and Wheat there are five crops. Table 3 shows the performance characteristics of individual prediction models across

[1] https://eaindustry.nic.in/

[2] https://www.thehindubusinessline.com/data-stories/visually/now-120-years-of-indian-rainfall-figures-at-your- fingertips/article62176162.ece

various datasets. Various models, including the DecisionTree Regressor, GaussianProcess Regressor, and Support Vector Regressor (SVR), are tested for each dataset namely Paddy, Maize, Barely, Wheat, and Ragi) using Mean RMSE, Mean R2, MAE, MAPE, and Mean Percent Error.

Table 4 compares the performance metrics of ensemble prediction models on the same datasets. Mean RMSE, Mean R2, MAE, MAPE, and Mean Percent Error are the same measures used to evaluate models like the XGB Regressor and RandomForest Regressor. These tables provide a detailed comparison of the predictive capabilities of single and ensemble models across several datasets. Surely, it provides insights about the usefulness for each one.

For these datasets, the DecisionTree Regressor model has consistently low Mean RMSE and high Mean R2 across all datasets. Therefore, it is came known that it has good predictive accuracy and fit. The GaussianProcess Regressor model shows greater mean RMSE and varied mean R2, indicating mixed prediction accuracy across datasets. Whereas, Support Vector Regressor model shows a higher Mean RMSE and lower Mean R2. It has less favourable prediction capabilities.

Table 3. Performance of single prediction models

Dataset	Model	Mean RMSE	Mean R2	MAE	MAPE	Mean Percent Error
Paddy	DecisionTree Regressor	0.575	0.995	0.87	0.01	0.005
	GaussianProcess Regressor	60.39	-30.49	115.61	0.8	0.4
	SVR	15.865	-0.04	14.73	0.1	0.095
Maize	DecisionTree Regressor	0.975	0.985	1.47	0.01	0.005
	GaussianProcess Regressor	53.605	-40	102.97	0.8	0.395
	SVR	15.86	-0.05	10.48	0.08	0.095
Barely	DecisionTree Regressor	4.715	0.835	5.72	0.04	0.02
	GaussianProcess Regressor	56.815	-22.61	107.63	0.78	0.385
	SVR	18.245	-0.005	13.23	0.1	0.105
Wheat	DecisionTree Regressor	4.61	0.76	6.19	0.04	0.02
	GaussianProcess Regressor	56.81	-0.35	108.17	0.79	0.395
	SVR	14.545	-0.01	12.41	0.09	0.095
Ragi	DecisionTree Regressor	3.445	0.975	4.44	0.02	0.01
	GaussianProcess Regressor	78.5	-0.10	144.44	0.76	0.38
	SVR	39.445	-0.32	34.84	0.2	0.175

Table 4. Performance of ensemble prediction models

Dataset	Model	Mean RMSE	Mean R2	MAE	MAPE	Mean PE
Paddy	XGB Regressor	0.87	1	0.84	0.01	0.005
	RandomForest Regressor	0.72	1	0.68	0	0
Maize	XGB Regressor	5.38	0.79	5.29	0.04	0.03
	RandomForest Regressor	3.215	0.94	2.55	0.02	0.015
Barely	XGB Regressor	6.19	0.825	7.13	0.05	0.035
	RandomForest Regressor	5.9	0.84	7.39	0.05	0.03
Wheat	XGB Regressor	3.38	0.93	3.83	0.03	0.02
	RandomForest Regressor	4.17	0.875	5.05	0.04	0.025
Ragi	XGB Regressor	4.355	0.98	4.11	0.02	0.015
	RandomForest Regressor	5.24	0.97	5.79	0.03	0.02

XGB Regressor model shows lower Mean RMSE and higher Mean R2 across datasets than RandomForest Regressor, indicating greater predictive performance. RandomForest Regressor model has slightly greater Mean RMSE and lower Mean R2. The classical ensemble model is characterised as combining numerous models and averaging their forecasts to increase prediction accuracy.CER model is defined as the combination of well performing models to increase forecast accuracy and robustness. For these crop datasets, it is noted the DecisionTree Regressor model and XGB Regressor model have consistent performance. Hence, these models are ideal candidates for inclusion in the proposed ensemble model.

Table 5 compares the performance metrics of the Classical Ensemble model and the CER model across different datasets. The RMSE based comparison between the Classical Ensemble and CER models across different datasets indicates considerable variances in their performance. It is graphically shown in Figure 5.

Table 5. Comparison of proposed competitive ensemble regression (CER) prediction models

Dataset	Model	Mean RMSE	Mean R2	MAE	MAPE	Mean Percent Error
Paddy	Classical Ensemble	0.7225	0.9975	0.855	0.01	0.005
	CER	0.575	0.995	0.87	0.01	0.004
Maize	Classical Ensemble	0.9225	0.9925	1.155	0.01	0.006
	CER	0.975	0.985	1.47	0.09	0.005
Barely	Classical Ensemble	3.1775	0.8875	3.38	0.025	0.0175
	CER	4.715	0.835	5.72	0.04	0.02
Wheat	Classical Ensemble	5.0475	0.8125	5.505	0.04	0.025
	CER	3.38	0.93	3.83	0.03	0.02
Ragi	Classical Ensemble	5.4525	0.83	6.425	0.045	0.0275
	CER	3.445	0.975	4.44	0.02	0.01

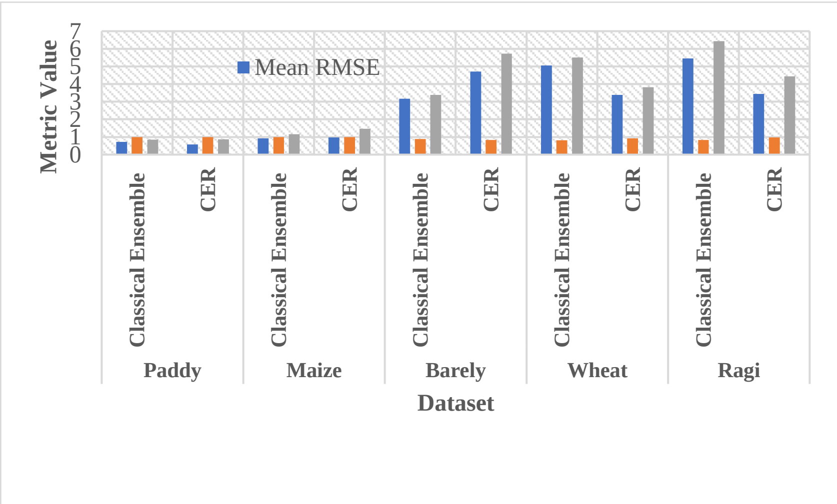

Fig.2. Comparative performance of ensemble models

Table 6. Forecasted price (in Rs) for commodities July 2024 -June2025)

Month	Paddy	Maize	Barley	Wheat	Ragi
July	3158.04	2797.81	2640.40	3131.28	8309.28
August	3167.30	3013.63	2696.00	3141.84	8295.31
September	3170.92	3013.63	2803.84	3166.90	8133.60
October	3171.00	2974.39	2876.30	3202.42	8156.83
November	3162.22	3025.40	2945.38	3252.55	8259.71
December	3150.70	3025.40	3031.32	3301.78	8162.99
January	3091.38	2695.79	2512.34	3059.31	7935.63
February	3100.80	2656.55	2512.34	3059.31	8088.07
March	3115.18	2656.55	2554.46	3086.76	8229.64
April	3126.52	2776.23	2605.01	3117.25	8357.90
May	3140.93	2776.23	2605.01	3145.92	8233.91
Jun	3146.35	2797.81	2628.60	3162.83	8176.66

On the rice dataset, the CER model has a significant 20.4% improvement in RMSE over the classical ensemble model. The CER model performs significantly better on the wheat and ragi datasets, with RMSE improvements of 32.9% and 36.8%, respectively. Notably, the Barley dataset exhibits the largest disparity, with the CER model predicting a significant 48.3% reduction in RMSE compared to the classical ensemble model. Conversely, in the maize dataset, the CER model shows a 5.7% increase in RMSE, indicating somewhat lower predictive ability. These findings highlight the ability of the CER model to improve forecasting accuracy on various datasets. Table 6 show the crops forecast prices for next one year based on the base value for India.

6. Conclusions

This work investigated a highly accurate forecasting model for crop price trend with competitive ensemble learning using a RWS operator in the output of four regressor models for rainfall and crop dataset. Instead of using a forecasting model with a single regressor, applying a competitive learning on each output of the model with multiple regressors was foundto play a significant role in improving model performance and making the model more robust.The results of the proposed methodology showed how competitive ensemble learning performed better than other single prediction models. A drawback of the study is that these datasets have a limited number of features. This can be overcome by adding more relevant features to the dataset for better accuracy in the future.

References

[1] D. Yang and Z. Liu, "Does farmer economic organization and agricultural specialization improve rural income? Evidence from China," *Economic Modelling*, vol. 29, no. 3, pp. 990–993, 2012.

[2] S. Yao, Y. Guo, and X. Huo, "An empirical analysis of the effects of China's land conversion program on farmers' income growth and labor transfer," *Journal of Environmental Management*, vol. 45, no. 3, pp. 502–512, 2010.

[3] L.-J. Chen, C. Ye, S.-W. Hu, V. Wang, and J. Wen, "The effect of a target zone on the stabilization of agricultural prices and farmers' nominal income," *Journal of Agricultural and Resource Economics*, vol. 38, no. 1, pp. 34–47, 2013.

[4] P. Flach, *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*. Cambridge: Cambridge University Press, 2012.

[5] D. Kumar and S. K. Rath, "Predicting the trends of price for Ethereum using deep learning technique," in *Artificial Intelligence and Evolutionary Computations in Engineering Systems*, S. Dash, C. Lakshmi, S. Das, and B. Panigrahi, Eds., Singapore: Springer, 2020, pp. 103–114.

[6] G. Milunovich, "Forecasting Australia's real house price index: A comparison of time series and machine learning methods," *Journal of Forecasting*, vol. 39, no. 7, pp. 1098–1118, 2020.

[7] P. J. Werbos, "Generalization of backpropagation with application to a recurrent gas market model," *Neural Network*, vol. 1, no. 4, pp. 339–356, 1998.

[8] G. P. Zhang and M. Qi, "Neural network forecasting for seasonal and trend time series," *Eur. J. Oper. Res.*, vol. 160, pp. 501–514, 2005.

[9] G. P. Zhang and D. M. Kline, "Quarterly time-series forecasting with neural networks," *IEEE Trans. Neural Netw.*, vol. 18, pp. 1800–1814, 2007.

[10] Y. Weng, X. Wang, J. Hua, H. Wang, M. Kang, and F. Y. Wang, "Forecasting horticultural products price using ARIMA model and neural network based on a large-scale data set collected by web crawler," *IEEE Trans. Comput. Soc. Syst.*, vol. 6, pp. 547–553, 2019.

[11] J. E. Sheehy, P. L. Mitchell, and A. B. Ferrer, "Decline in rice grain yields with temperature: Models and correlations can give different estimates," *Field Crop Res.*, vol. 98, no. 2–3, pp. 151–156, 2006.

[12] Y. Zhang and S. Na, "A novel agricultural commodity price forecasting model based on fuzzy information granulation and Mea-SVM model," *Mathematical Problems in Engineering*, vol. 2018, pp. 1–10, 2018.

[13] R. K. Paul, M. Yeasin, P. Kumar, P. Kumar, M. Balasubramanian, H. S. Roy, et al., "Machine learning techniques for forecasting agricultural prices: A case of brinjal in Odisha, India," *PLoS ONE*, vol. 17, no. 7, p. e0270553, 2022.

[14] A. Chatzimparmpas, R. M. Martins, K. Kucher, and A. Kerren, "StackGenVis: Alignment of data, algorithms, and models for stacking ensemble learning using performance metrics," *IEEE Trans. Vis. Comput. Graph.*, vol. 27, pp. 1547–1557, 2021.

[15] Z. Li, Z. Chen, Q. Cheng, F. Duan, R. Sui, X. Huang, and H. Xu, "UAV-based hyperspectral and ensemble machine learning for predicting yield in winter wheat," *Agronomy*, vol. 12, p. 202, 2022.

[16] S. V. Razavi-Termeh, A. Sadeghi-Niaraki, and S. M. Choi, "Spatial modeling of asthma-prone areas using remote sensing and ensemble machine learning algorithms," *Remote Sens.*, vol. 13, p. 3222, 2021.

[17] M. Ustuner and F. B. Sanli, "Polarimetric target decompositions and light gradient boosting machine for crop classification: A comparative evaluation," *ISPRS Int. J. Geo-Inf.*, vol. 8, 2019.

[18] Y. Zhang, J. Liu, and W. Shen, "A review of ensemble learning algorithms used in remote sensing applications," *Applied Sciences*, vol. 12, no. 17, p. 8654, 2022.

[19] A. M. Ticlavilca, D. M. Feuz, and M. Mckee, "Forecasting agricultural commodity prices using multivariate Bayesian machine learning regression," NCCC-134, 2010.

[20] L. Madaan, A. Sharma, P. Khandelwal, S. Goel, P. Singla, and A. Seth, "Price forecasting & anomaly detection for agricultural commodities in India," in *Proceedings of the Conference on Computing & Sustainable Societies - COMPASS 19*, 2019.

[21] Y. Guo, D. Tang, W. Tang, S. Yang, Q. Tang, Y. Feng, and F. Zhang, "Agricultural price prediction based on combined forecasting model under spatial-temporal influencing factors," *Sustainability*, vol. 14, no. 17, p. 10483, 2022.

[22] Y. H. Gu, D. Jin, H. Yin, R. Zheng, X. Piao, and S. J. Yoo, "Forecasting agricultural commodity prices using dual input attention LSTM," *Agriculture*, vol. 12, no. 2, p. 256, 2022.

[23] V. Talasila, C. Prasad, G. Reddy, and A. Aparna, "Analysis and prediction of crop production in Andhra region using deep convolutional regression network," *Int. J. Intell. Eng. Syst.*, vol. 13, no. 5, pp. 1–9, 2020.

[24] L. Mao, Y. Huang, X. Zhang, S. Li, and X. Huang, "ARIMA model forecasting analysis of the prices of multiple vegetables under the impact of the COVID-19," *PLoS ONE*, vol. 17, no. 7, 2022.

[25] A. K. Mahto, M. A. Alam, R. Biswas, J. Ahmed, and S. I. Alam, "Short-term forecasting of agriculture commodities in context of Indian market for sustainable agriculture by using the artificial neural network," *Journal of Food Quality*, vol. 2021, pp. 1–13, 2021.

[26] X. Wang, S. Gao, Y. Guo, S. Zhou, Y. Duan, and D. Wu, "A combined prediction model for Hog Futures prices based on WOA-LIGHTGBM-CEEMDAN," *Complexity*, vol. 2022, pp. 1–15, 2022.

[27] R. K. Paul, M. Yeasin, P. Kumar, P. Kumar, M. Balasubramanian, H. S. Roy, et al., "Machine learning techniques for forecasting agricultural prices: A case of brinjal in Odisha, India," *PLoS ONE*, vol. 17, no. 7, 2022.

[28] N. Soni and J. Raut, "Crop price prediction using machine learning techniques," 2022.

[29] D. Zhang, S. Chen, L. Liwen, and Q. Xia, "Forecasting agricultural commodity prices using model selection framework with time series features and forecast horizons," *IEEE Access*, vol. 8, pp. 28197–28209, 2020.

[30] M. Nabipour, P. Nayyeri, H. Jabani, A. Mosavi, E. Salwana, and S. S., "Deep learning for stock market prediction," *Entropy*, vol. 22, no. 8, p. 840, 2020.

[31] Z. Chen, H. S. Goh, K. L. Sin, K. Lim, N. K. Chung, and X. Y. Liew, "Automated agriculture commodity price prediction system with machine learning techniques," *Advances in Science, Technology and Engineering Systems Journal*, vol. 6, no. 4, pp. 376–384, 2021.

[32] H. P. Suresha, M. Uppaladinni, and K. K. Tiwari, "Indian commodity market price comparative study of forecasting methods - A case study on onion, potato and tomato," *Asian Journal of Research in Computer Science*, vol. 12, no. 4, pp. 147–159, 2021.

[33] A. A. Adeniran, A. R. Adebayo, H. O. Salami, M. O. Yahaya, and A. A. Abdulraheem, "A competitive ensemble model for permeability prediction in heterogeneous oil and gas reservoirs," *Applied Computing and Geosciences*, vol. 1, p. 100004, 2019.

[34] A. H. Utomo, M. A. Gumilang, and A. Ahmad, "Agricultural commodity sales recommendation system for farmers based on geographic information systems and price forecasting using probabilistic neural network algorithm," in *IOP Conference Series: Earth and Environmental Science*, vol. 980, no. 1, p. 012061, 2022.

[35] S. Bayona-Oré, R. Cerna, and E. T. Hinojoza, "Machine learning for price prediction for agricultural products," *WSEAS Transactions on Business and Economics*, vol. 18, pp. 969–977, 2021.

[36] M. Shahhosseini, G. Hu, and H. Pham, "Optimizing ensemble weights and hyperparameters of machine learning models for regression problems," *Machine Learning with Applications*, vol. 7, p. 100251, 2022.

[37] T. van Klompenburg, A. Kassahun, and C. Catal, "Crop yield prediction using machine learning: A systematic literature review," *Computers and Electronics in Agriculture*, vol. 177, p. 105709, 2020.

Authors' Profiles

R. Ragunath is a researcher pursuing his Ph.D. program in Computer Science at Periyar University, Salem, Tamil Nadu, India, specializing in Machine Learning. His research interests include Machine Learning, Deep Learning, Sustainable AI, and Explainable Artificial Intelligence (XAI).

Dr. R. Rathipriya is an Associate Professor in the Department of Computer Science at Periyar University, Salem, Tamil Nadu, India. With extensive experience in machine learning, artificial intelligence, and data science. She focuses on sustainable AI, ensemble modeling, and optimization techniques. She has authored numerous publications in reputed journals and actively contributes to advancing environmentally conscious AI practices. Their research interests include hyperparameter optimization, multi-criteria decision-making, and applications of AI in healthcare and agriculture. Scopus ID: 55091157400

How to cite this paper: R. Ragunath, R. Rathipriya, "Forecasting Agriculture Commodity Price Trend using Novel Competitive Ensemble Regression Model", International Journal of Information Technology and Computer Science(IJITCS), Vol.17, No.3, pp.97-105, 2025. DOI:10.5815/ijitcs.2025.03.07