INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Stacked Ensembles: Boosting Model Performance to New Heights Based on Regression for Forecasting Future Wheat Commodities Prices in Gujarat

Aditi M Joshi[1]*, **Sanjay G Patel**[2]

1 Kadi Sarva Vishwavidyalaya, Gandhinagar, 382015, Gujarat, India
2 Department of Computer Science and Engineering, Institute of Technology, Nirma University, India

OPEN ACCESS

Received: 05-08-2022

Accepted: 16-09-2022

Published: 10-11-2022

Citation: Joshi AM, Patel SG (2022) Stacked Ensembles: Boosting Model Performance to New Heights Based on Regression for Forecasting Future Wheat Commodities Prices in Gujarat. Indian Journal of Science and Technology 15(42): 2194-2203. https://doi.org/10.17485/IJST/v15i42.1623

* **Corresponding author**.

joshiaditi5993@gmail.com

Funding: None

Competing Interests: None

Copyright: © 2022 Joshi & Patel. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: The goal of this study is to develop an estimation method that will improve absolute and proportionate price predictions, helping farmers in their long-term efforts to boost output and profit. **Methods:** For the experimentation, Dataset is made up of wheat commodities price of different mandi's of Gujarat region which is collected from agmarknet website, run by Indian government and weather variables which are obtained from weather API (world weather online website) based on commodities region such as Max-Min Temp, Pressure, Wind, WindSpeed, Humidity respectively which has a direct impact on crop production. Various machine learning approaches such as SVR (Support vector regressor), DTR (Decision tree regression), PR(Polynomial regression), Multiple regression, and Boosting algorithms such as XgBoost, Light GBM experimented for Daily data, 10 days and 15 days average price estimation, from the experimentation, observed that DTR Can't be performed well with continuous numerical data. The presence of one or two outliers in the data can seriously affect the results of PR (sensitive to the outliers). The meta-learner attempts to minimize the weakness and maximize the strengths of every individual model such as Multiple regression and SVR (Support vector regressor). **Findings:** The development of a stacking model, which aids a farmer in more accurately predicting crop price, is motivated by the study of several agriculture-specific requirements like weather parameters, and fuel cost. Every individual regression model like the Decision tree regressor, Support vector regressor, and polynomial regressor is evaluated by the meta-learner to identify its strengths and weaknesses. **Novelty:** This research addresses the complexity of crop price forecasting along with unique feature set which includes (i) historical commodities prices (ii) weather data (iii) transportation factors iiii) Holiday Impacts.

Keywords: Agriculture; Commodity price; Wheat; Machine learning; Regression; Time Series; Forecasting

1 Introduction

Wheat is a vital food crop that helps to meet global food demands. Gujarat wheat data for 2020 was recorded at 3,326.800 Tons. This is an increase compared to the previous year's total of 2,407.400 Ton tones. Agricultural Production: Wheat: Gujarat data is updated yearly, with an average of 1,637.000 Ton from March 1981 to March 2020. In 2014, the statistics hit an all-time high of 4,694.000 Ton, surpassing the previous high of 351.200 Ton set in 1988. Agricultural Production: Wheat: Gujarat data is reported by the Directorate of Economics and Statistics, Department of Agriculture and Farmers Welfare, and is currently in CEIC[(1)].

Based on the various parameters that determine the crop prices, a unified system can be developed which can act upon the past data and predict the upcoming prices in near future. It will also involve a ubiquitous network of algorithms working in harmony to improve the accuracy of price prediction. Based on the input from the past data, the forecasting of agricultural produce for the profits of the farmer can be done. Maximization of agricultural yield can benefit the economy of the country drastically. With a unified approach, planning analysis becomes easier than ever. This will enable farmers to indulge in Smart farming. With maximum profits and keeping up with constantly changing demands, the farmer's contribution towards an even higher GDP will be notable.

This research serves as an answer to the question "How to make accurate predictions for crop prices based on farmer's region and weather conditions?"

1.1 Related Methodology

The most important factor in agricultural sustainable development, which strongly influences the economy, is the cost of agricultural produce.[(2)] The key issue is that rural Indian farmers still don't have access to the newest technologies, making it challenging for them to obtain the information they need and causing price swings. Despite the pandemic we are currently facing, the agriculture sector rose by 1.2% in January 2021 compared to the same month in 2020, and small farmers have been performing a vital role in the Agri - sector's recovery[(3)]. To improve farmer returns, the Indian government has undertaken several programs. Making informed choices about when to plant, what kind of crop to grow, when to sell, and where to sell their goods is one of the biggest issues facing farmers today.

Numerous forecasting strategies have been created in recent years to address the variety and complexity of managerial forecasting problems that are on the rise. The prediction problem has traditionally been solved using linear statistical approaches (such as ARIMA models), but more recently, with the development of machine learning techniques, the focus of the solution has mainly changed from statistical methods to the machine learning domain[(4)]. Research has been done in recent decades to estimate and predict the price of agricultural commodities. To improve predictions, the researchers explored a wide range of various techniques and factors. However, there remains a research gap in how to dynamically choose the best and most consistent algorithm for the raw data and criteria. The historical research reveals that several agricultural commodity price projections were made using various algorithms shown in Table 1. The least-square estimation, regression and correlation, statistical inference, and non-linear functional forms were investigated in the simple regression approaches. Over the past three decades, one of the most widely used linear models for time series forecasting has been the autoregressive integrated moving average (ARIMA). The well-known Box-Jenkins model-building methodology and the statistical characteristics of the ARIMA model contribute to its popularity. ARIMA models can also be used to build other exponential smoothing techniques. The ARIMA was thoroughly investigated. The fundamentals of ANN, including its model, modes of behaviors, and structure, as well as fundamental learning principles like the delta rule and Hebbian learning principles, were initially investigated. The effectiveness of ANN in comparison to ARIMA was then tested using business data and a straightforward ANN algorithm[(3,5)].

While earlier research has developed algorithms to forecast produce price changes, they do not explicitly consider the spatiotemporal reliance of future pricing on historical data. As a result, they utilize traditional classical and statistical prediction models (such as decision trees, and ARIMA), which frequently exhibit poor performance and are computationally expensive (we validate this in our experiments). As The ARIMA model can be used to forecast periodic changes, but it cannot forecast changes caused by unforeseen factors. These issues reduce the accuracy (and hence, the practicality (or utility) of these techniques. In this study, we suggest solutions to these flaws. It has been observed that combining approaches deliver better results than using individual models. To help the millions of struggling farmers and improve their living conditions by building the best stacked-Model utilizing historical along with the weather. Long-term increase in output and profit through intelligence strategy[(9,11)].

Table 1. Related Methodology and Deficiencies in Previous Research

Citation	Algorithms / techniques / models	Research gap
(6)	BP Neural Network, Improved BP Neural Network Model	The low prediction resulted from the prediction methods' emphasis on the linear relationship between the prices. Backpropagation algorithms are susceptible to noisy data.
(7)	Weighted Linear Regression	The historical data's knowledge could not be applied effectively to manage the factors that affected agricultural revenue, which led to the failed forecast. Correcting incorrect data in the system and alerting relevant systems are future concerns.
(8)	ANN, RNN, LSTM	Authors could extrapolate their maximum temperature prediction to cover precipitation and wind speed, among other factors. The price is affected by a variety of factors, including its qualities, demand, seasonal trends, and price offers for comparable goods from a wide range of suppliers.
(9)	ARIMA, SARIMA, Prophet	The widely used time-series forecasting methods (ARIMA, SARIMA, and Prophet) were found to be greatly impacted by data quality problems, which are a common problem in developing nations like India. Learn more about context-based rules and how to use model stacking and other meta-learning approaches to select the optimal strategy from the available ones and improve performance. The daily manual entering of crop prices at marketplaces in developing nations like India increases the risk of human-caused errors like inaccurate data entry or days without data entry. In addition to these human flaws, the price swings themselves make it difficult to develop a stable and reliable forecasting solution.
(10)	Random forest ensemble learning, Decision tree regressor	The random forest regressor takes fewer data and is less computationally demanding than ANN. For Naive Bayes to perform better than Random Forest, a lot of data must be collected.
(11)	ARIMA, SARIMA	The historical wholesale prices that were used in the investigation may have been affected by the effects of inflation. To lessen the impact of inflation, it is then recommended to translate past price values into current actual prices using an appropriate price index.
(12)	Decision tree regressor	Calculate the pricing for the upcoming 12 months. Can be used to estimate short-term prices for planning on a monthly or quarterly basis.
(5)	Long-Short Term Memory (LSTM)	The technique can be improved in the future by taking more factors like soil type and water level into account. Additionally, include a schedule and instructions for fertilization, which will aid farmers with no prior knowledge of crops. Additionally, the system can be altered to accept data from IoT gadgets without relying on raw data. Other than that, this system can also be created for different platforms.
(13)	A neural network, Multiple Linear Regression	Weather-related factors are not considered for the study. Future improvements may involve including other parameters in addition to the soil parameters.
(14)	ARIMA, XGBoost, LASSO, Random Forest	The potential distortion of market prices is a risk and a negative effect of price forecasting. This could result in an oversupply, reduced pricing, and a protracted waiting period for the farmers to sell their products if all the farmers who receive the forecasts choose to sell at one specific market. While preparing for deployment in the actual world, contingencies for such a result must be planned.

2 Methodology

2.1 Recognition of Foremost Features (15–18)

Authors have recognized foremost features that affects the crop price as mentioned in Figure 1 and to select the most significant features correlation heatmap plotted which is shown in Figure 2. Methods of correlation are used to determine the "association" of two or more variables. Each sampling unit has two observations. The two variables are said to be correlated when a change in one variable is followed by a change in the other. Thus, correlation is a statistical method that may be used to establish a relationship between two variables if there is a cause between them.

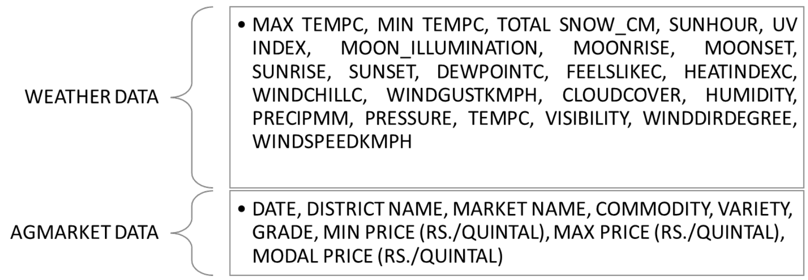

Fig 1. Foremost features

Pearson correlation formula

$$r = \frac{\Sigma\left(x_i - \overline{x}\right)\left(x_i - \overline{y}\right)}{\sqrt{\Sigma\left(x_i - \overline{x}\right)^2 \Sigma\left(yi_i - \overline{y}\right)^2}} \tag{1}$$

Where,

r = correlation coefficient

x_i = values of the x-variable in a sample

$\overline{x}$ = mean of the values of the x-variable

y_i = values of the y-variable in a sample

$\bar{y}$ = mean of the values of the y-variable

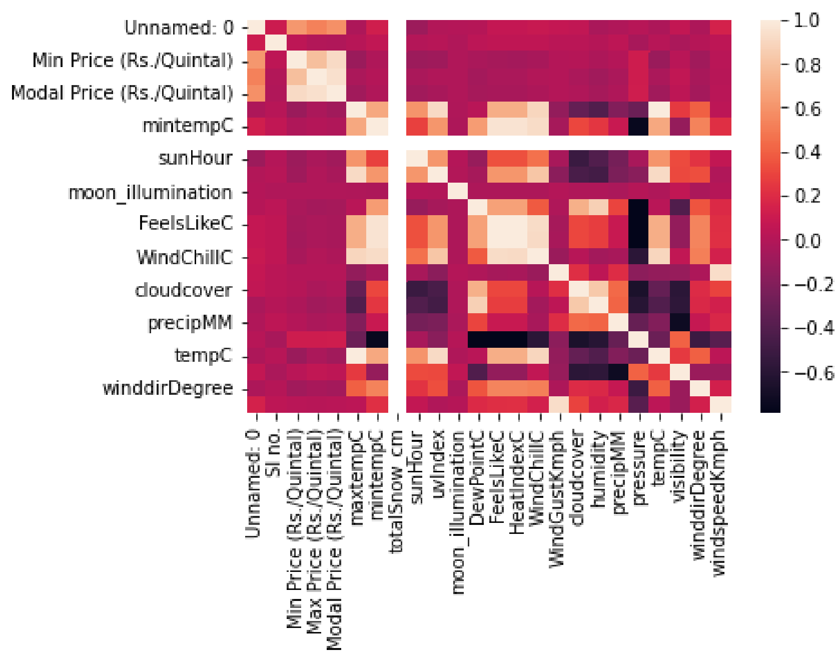

Fig 2. Parameters & Correlation Heatmap

A correlation matrix in Figure 2 shows the relationship between the variables in a set of data, where value 0 represents the no correlation exists between the variables and removed these identified irrelevant variables from data. while 0.8,0.6 and similar values whose range is between $\pm$ 0.9 or $\pm$ 0.8 shows the strong relationship among the variables.

2.2 Dataset

Machine Learning models are data-hungry and necessitate a large amount of data to produce the best model or a high-fidelity system. Even if you've developed excellent machine learning algorithms, the quality of your data is just as crucial as the quantity. Real-world datasets, on the other hand, are more complex, jumbled, and unstructured. The number, quality, and relevance of the dataset determine the performance of any Machine Learning or Deep Learning model. Finding the correct balance is a difficult task. The Agmarknet website, which is run by the Indian government and holds daily mandi prices and arrival information for several commodities in the nation, was used to compile a dataset of a wheat commodity for the Gujarat region. The ultimate dataset was chosen based on correlation coefficients between the variables and included weather parameters and commodities data. Agmarnet's repository of time-series data has many missing values, which might be the reason why the model performs less well. In data analysis and modelling, missing data can lead to issues. The soft impute data imputation approach [19] examined, against other imputation techniques like linear interpolation, last observation carried forward and moving average. Soft impute performance among them is better with an average RMSE of 48.26.

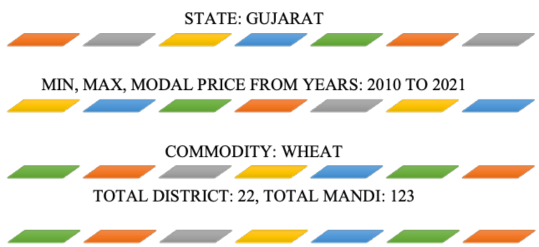

Fig 3. Statistics for the Wheat dataset

2.3 Experimental Framework and Methods [15]

Mandi has no impact on agricultural prices, it was discovered after analyzing various mandi-wise prices. Therefore, it is appropriate to use all the data at once, which provides us with plenty of data. Using a new stacked model technique, we first imputed the data into multiple regression models and SVR, and then we used the output of these models, weather data, and past days data (5, 10, and 15 days of Time Frame), to create a meta-model. Meta model forecasts crop price using XGB. Machine learning algorithms that learn from the results of other machine learning algorithms are most frequently referred to as meta-learning. Here, stacked model is assortment of three base models SVR, DTR and XGB.

Crop price (CP), a separate feature, is provided to the stacked meta model together with the day, the month, and the year.

In 1 day, cp1 to cp5 means the last 5 days.

In 10 days - cp1 means the last 10 days, cp2 means the last 11-20 days, cp3 means the last 21-30 days and so on.

In 15 days - cp1 means the last 15 days, cp2 means the last 16-30 days, cp3 means the last 31-45 days and so on.

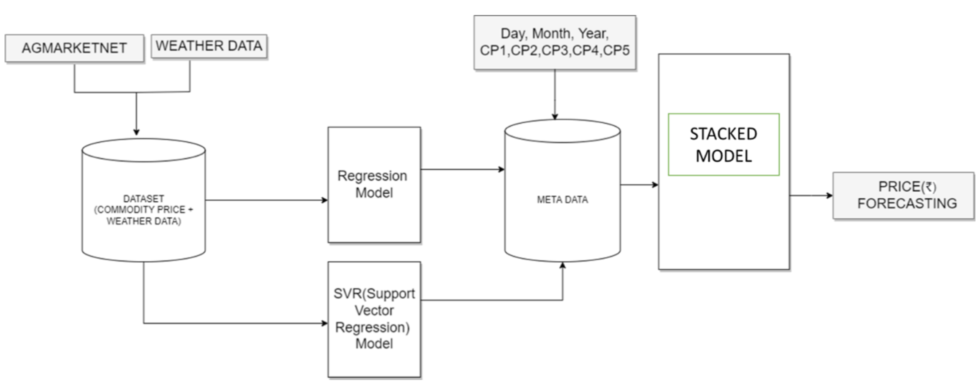

Fig 4. Proposed Framework

Consequently, it can be inferred from the previous experiments that our innovative stacked-model method outperforms more established ML and statistical models. This is due to the typical ML models' inability to effectively update the weights according to the loss function when they are trained. Contrarily, with the stacked model technique, the model receives the loss functions of the two prior models (SVR and LR) as input, which aids in updating weights and improving convergence to the global minima.

3 Results and Discussion

Table 2. Comparative study amongst other similar approaches

Citation	Actual Price	Forecasted Price	Difference in price	Difference in price (%)
(7)	760.00	650.00	110.00	14.47
	660.00	800.00	140.00	21.21
	785.00	650.00	135.00	17.20
	580.00	400.00	180.00	31.03
	710.00	850.00	140.00	19.72
	1020.00	1180.00	160.00	15.69
	540.00	430.00	110.00	20.37
	580.00	790.00	210.00	36.21
	1040.00	880.00	160.00	15.38
(6)	96.90	103.20	6.30	6.50
	98.60	106.60	8.00	8.11
Author's Result (For Daily Price Forecasting)	2059.00	2078.81	19.81	0.96
	1598.13	1628.37	30.25	1.89
	1983.80	1905.30	78.50	3.96
	2156.89	2004.65	152.24	7.06
	2212.75	2061.86	150.89	6.82
	1694.44	1638.84	55.60	3.28
Author's Result (For 10 Days Forecasting)	2059.00	2078.81	19.81	0.96
	1598.13	1628.37	30.25	1.89
	1983.80	1905.30	78.50	3.96
	2156.89	2004.65	152.24	7.06
	2212.75	2061.86	150.89	6.82
	1694.44	1638.84	55.60	3.28
Author's Result (For 15 Days Forecasting)	2059.00	2078.81	19.81	0.96
	1598.13	1628.37	30.25	1.89
	1983.80	1905.30	78.50	3.96
	2156.89	2004.65	152.24	7.06
	2212.75	2061.86	150.89	6.82
	1694.44	1638.84	55.60	3.28

Table 2 shows the analysis of the proposed systems performance amongst the other similar approaches[6,7]. Previous approaches result shows higher error rates compared to author's daily, 10 days average and 15 days average forecast.

Several alternative methodologies, including Decision Tree Regression, Polynomial Regression, Multiple Regression, Support Vector Regression (SVR), Gradient Boosting Regressor, and Stacked Model, are used to evaluate performance evolution Mean Absolute Percentage Error (MAPE), and root mean square error (RMSE) were calculated to evaluate the outcome. The forecasting accuracy increases with decreasing the gap between estimated and actual prices; a MAPE of less than 5% is seen as a sign that the forecast is respectably accurate. According to a general rule of thumb, RMSE values between 0.2 and 0.5 indicate that the model can reasonably predict the data reliably. As R2 grows closer to 1, the degree of model fitting also gets closer. One of the methods most frequently used to assess the accuracy of forecasts is the root mean square error, also known as root mean square deviation.

As indicated in Figure 5, the RMSE (root mean squared error) can be used to assess the performance of the models.

If we look at the table, the stacked model outperforms other methods for data averaged over 10 and 15 days, respectively. In contrast, polynomial regression results indicate a greater RMSE, indicating that the model does not adequately match the data. It has been noted that a linear relationship between the independent and dependent variables in the data set is not necessary for Polynomial Regression to work. Polynomial regression is employed when the linear regression model is unable to appropriately represent the optimal conclusion and the data points in the data. However, in polynomial regression, one or two data outliers may have a considerable effect on the results of the nonlinear analysis. These rely too much on extreme cases.

In conclusion, stacked models outperformed support vector regression and polynomial regression for average prediction times of 10 and 15 days, respectively. Compared to average ensembles or normal individual models, stacked ensembles typically give a more reliable forecasting performance.

It is evident from a comparison of Figure 5 and Figure 6 that stacked models are more accurate and effective at forecasting crop prices, especially over the short term. For the past 10 days, it has been noted that the polynomial regression and support vector regression RMSEs are the greatest, which indicates that the data has likely been grossly overfitted, although the daily predictive RMSE is nearly identical and is around 33%.

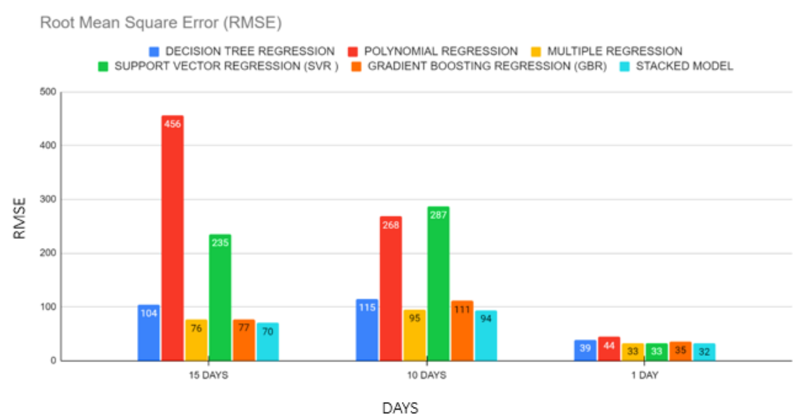

Fig 5. The result of RMSE comparison among different methods

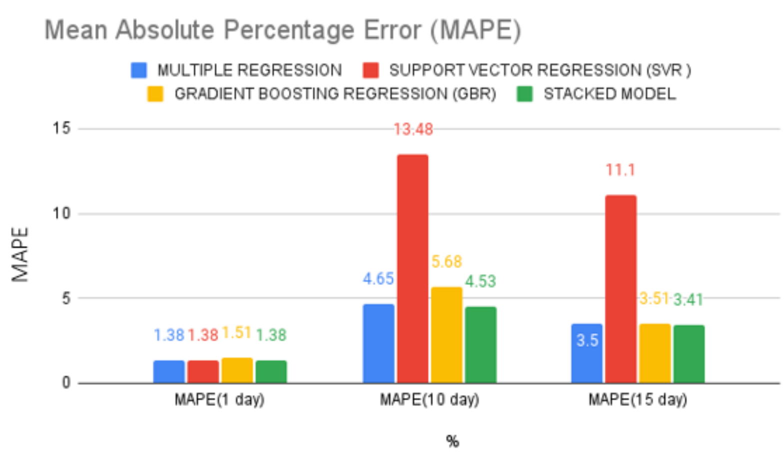

Fig 6. The result of MAPE comparison among different methods

No single algorithm consistently outperformed others on the commodities and prediction frequencies, according to the obvious trial data. With 1.38% inaccuracy on daily data, the performance of stacked models and regression techniques for wheat data is nearly identical. However, with an average error of 4.53% for the data averaged over 10 days, stacked models and multiple regression outperform other algorithms. Support vector regression does not perform well for 10 days when we compare the same commodity data with daily and 15-day average data.

The model with reduced MAPE for daily data, which indicates how well the machine learning model can forecast values, is the Stack Model when we consider the same commodity data with daily, 10 days, and 15 days of data from Gujarat. Gradient Boosting performs better than Multiple Regression and Support Vector Regression in the instance of Wheat Daily Data for Gujarat Region. In comparison to stacked models and multiple regression, model performance is lower at predicting values, as seen by higher MAPE values approximately 11% observed of SVR for 15 days of average prediction.

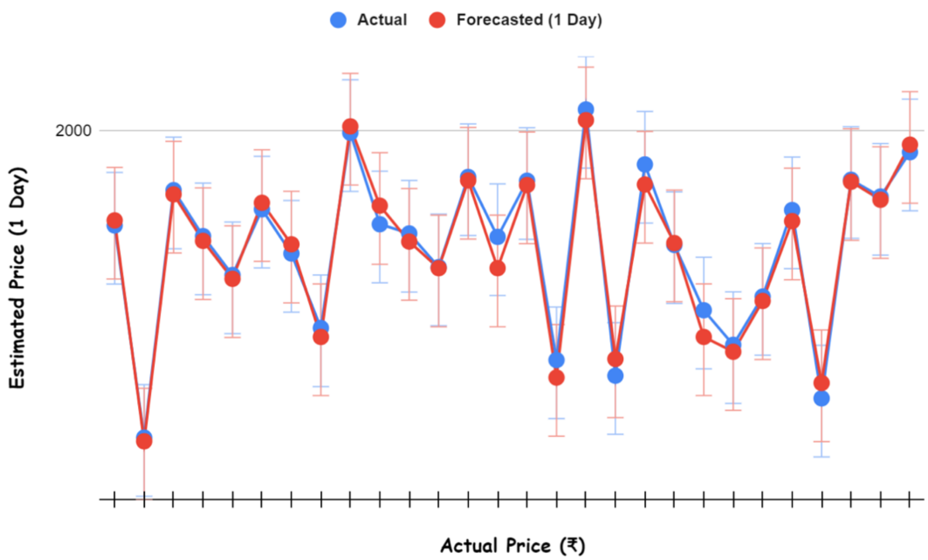

Fig 7. Comparison of Predicted and Actual value (1 Day of Time Frame) for Stacked Model.

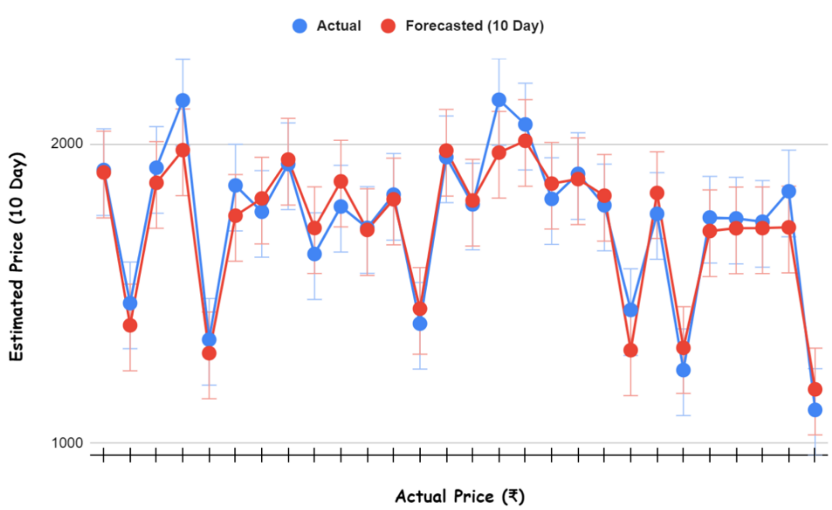

Fig 8. Comparison of Predicted and Actual value (10 Days of Time Frame) for Stacked Model.

Most of the researchers have focused LSTM (long short-term memory) [2,4], Neural network [20], which needs a huge amount of data as compared to the meta-model presented in the current study, and Traditional statistical approaches like ARIMA,

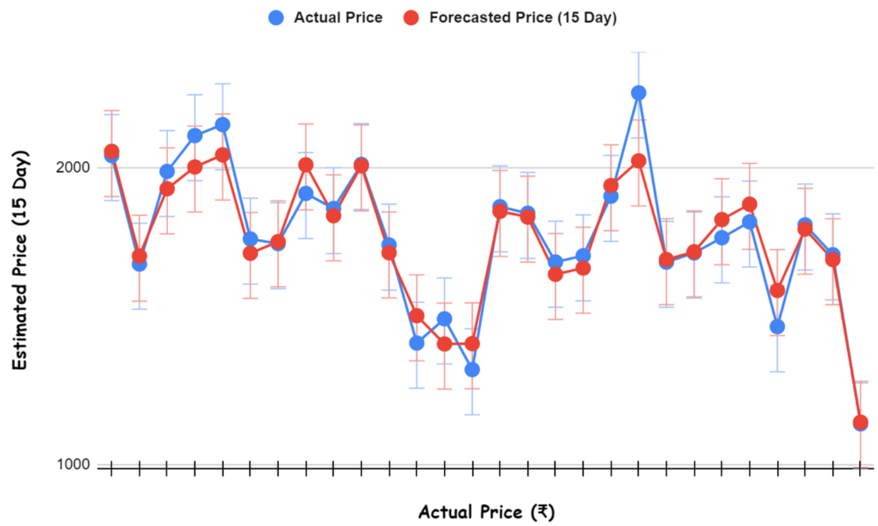

Fig 9. Comparison of Predicted and Actual value (15 Days of Time Frame) for Stacked Model.

SARIMAX[3,5,11], which required high computing capacity of classical Machine learning models such as autoregressive models, Feature engineering is performed manually and not able to learn more complex data patterns ultimately. Almost all the published research work reported improvement in accuracy with Bi – Direction LSTM. Forward and backward propagation is the window of time where the current price is affected, which is less computationally intensive than SARIMAX. An increase in differential order will take more time to run models like ARIMA, SARIMAX.[2–5,11,20] and Logarithmic Complexity increases with the SARIMAX, Same doesn't apply to LSTM.

4 Conclusions

Data was enriched quantitatively. Holiday influences caused some daily pricing data to be missing. The experiment hindered from some daily commodity pricing data that remained static for several days. Authors transformed the daily average price data into the 10 days and 15 days average price data to resolve the mentioned holiday issues. From experimentation, authors conclude that our innovative stacked-model technique outperforms classic machine learning SVR, Multiple regression and statistical models ARIMA, SARIMAX respectively. The reason for this is tradition approaches are unable to update weights efficiently according to the loss function when trained. In the stacked learners, the model receives the loss function of the preceding two models (SVR and Multiple Regression) as an input, which allows the model to update weights and converge more quickly to the global minima. Because XGB has learned the shortcomings of the two base models, the proposed stacked model technique is quite accurate for short-term price estimation of agricultural commodities. In the future, the model can be trained on a variety of commodities, which will aid recognition of complex patterns and the facilities of more accurate information. Additionally, it might be enhanced by include features like soil type and water level, fertilization calendar and instructions, which would help farmers who aren't familiar with crops.

5 Data Availability

On request, the corresponding author will provide access to the data utilized in this work.

References

1) Agricultural Production: Wheat: Gujarat" | Economic Indicators | CEIC Ceicdata.com. 2022. Available from: https://www.ceicdata.com/en/india/production-of-foodgrains-in-major-states-wheat/agricultural-production-wheat-gujarat.
2) Guo Y, Tang D, Tang W, Yang S, Tang Q, Feng Y, et al. Agricultural Price Prediction Based on Combined Forecasting Model under Spatial-Temporal Influencing Factors. *Sustainability*;14(17):10483. Available from: https://doi.org/10.3390/su141710483.
3) Bayona-Oré S, Cerna R, Hinojoza ET. Machine Learning for Price Prediction for Agricultural Products. *WSEAS TRANSACTIONS on BUSINESS and ECONOMICS*. 2021;18:989–977. Available from: https://doi.org/10.37394/23207.2021.18.92.

4) Chen Z, Goh HS, Sin KL, Lim K, Chung NKH, Liew XY. Automated Agriculture Commodity Price Prediction System with Machine Learning Techniques. *Advances in Science, Technology and Engineering Systems Journal*. 2021;6(4):376–384. Available from: https://www.astesj.com/publications/ASTESJ_060442.pdf.
5) Selvanayagam T, Suganya S, Palendrarajah P, Manogarathash AMP, Gamage D, Kasthurirathna. Agro-Genius: Crop Prediction using Machine Learning. *International Journal of Innovative Science and Research Technology*. 2019;4(10):243–249. Available from: https://ijisrt.com/assets/upload/files/IJISRT19OCT1880.pdf.pdf.
6) Minghua W, Qiaolin Z, Zhijian Y, Jingui Z. Prediction model of agricultural product's price based on the improved BP neural network. *2012 7th International Conference on Computer Science & Education (ICCSE)*. 2012. Available from: https://doi.org/10.1109/ICCSE.2012.6295150.
7) Corraya AD, Corraya S. Regression Based Price and Yield Prediction) of Agricultural Crop". *International Journal of Computer Applications*. 2016;152(5):1–7. Available from: https://www.ijcaonline.org/archives/volume152/number5/corraya-2016-ijca-911804.pdf.
8) Tran TTK, Lee T, Shin JY, Kim JS, Kamruzzaman M. Deep Learning-Based Maximum Temperature Forecasting Assisted with Meta-Learning for Hyperparameter Optimization. *Atmosphere*. 2020;11(5):1–21. Available from: https://doi.org/10.3390/atmos11050487.
9) Jain A, Marvaniya S, Godbole S, Munigala V. A Framework for Crop Price Forecasting in Emerging Economies by Analyzing the Quality of Time-series Data. *arXiv*. 2020. Available from: https://doi.org/10.48550/arXiv.2009.04171.
10) Prassanna SA, Harshanand BA, Srishti B, Chaitanya R, Soundari SKN, Sriram M, et al. Crop Value Forecasting using Decision Tree Regressor and Models. *European Journal of Molecular & Clinical Medicine*. 2020;07(02):3702–3722. Available from: https://ejmcm.com/article_2585_86ae450aad4837565e30767586a4d64b.pdf.
11) Rathnayake RMYL, Razmy AM, Alibuhtto MC. A Comparison Of Price Fluctuations For The Green Chilli And Tomato In Dambulla Dedicated Economic Center In Srilanka. *5th Annual Science Research Sessions*. 2016. Available from: https://www.researchgate.net/publication/349699325_AN_ANALYSIS_OF_COLOMBO_OPEN_MARKET_AVERAGE_RETAIL_PRICES_OF_BIG_ONIONS_IN_SRI_LANKA_USING_TIME_SERIES_FORECASTING_TECHNIQUES.
12) Dhanapal R, AjanRaj A, Balavinayagapragathish S, Balaji J. Crop price prediction using supervised machine learning algorithms (article retracted). *Journal of Physics: Conference Series*. 2021;1916:012042. Available from: https://iopscience.iop.org/article/10.1088/1742-6596/1916/1/012282.
13) Lavanya M, Parameswari R. A Multiple Linear Regressions Model for Crop Prediction with Adam Optimizer and Neural Network Mlraonn. *International Journal of Advanced Computer Science and Applications*. 2020;11(4):253–257. Available from: https://dx.doi.org/10.14569/IJACSA.2020.0110434.
14) Yadav A, Verma A, Govindaraj A, Sarbajna V. Increasing small holder farmer income by providing localized price forecasts. In: 33rd Conference on Neural Information Processing Systems. 2019;p. 1–5. Available from: https://aiforsocialgood.github.io/neurips2019/accepted/track1/pdfs/51_aisg_neurips2019.pdf.
15) Agriculture Market. 2022. Available from: https://agmarknet.gov.in/.
16) Weather API. 2022. Available from: https://www.worldweatheronline.com/developer.
17) The Role of Weather And Weather Forecasting In Agriculture. 2018. Available from: https://www.dtn.com/the-role-of-weather-and-weather-forecasting-in-agriculture.
18) Economic Importance, Soil And Climatic Requirement, Varieties, Cultural Practices And Yield. 2022. Available from: http://ecoursesonline.iasri.res.in/Courses/Field%20Crops/AGRO302/Data%20Files/Contents/lec04.pdf.
19) Moritz S, Bartz-Beielstein T. imputeTS: Time Series Missing Value Imputation in R. *The R Journal*. 2017;9(1):207–218. Available from: https://journal.r-project.org/archive/2017/RJ-2017-009/RJ-2017-009.pdf.
20) Utomo MAAH, Gumilang, Ahmad. Agricultural Commodity Sales Recommendation System For Farmers Based on Geographic Information Systems and Price Forecasting Using Probabilistic Neural Network Algorithm. *IOP Conf Series: Earth and Environmental Science* . 2022;980:1–10. Available from: https://iopscience.iop.org/article/10.1088/1755-1315/980/1/012061/pdf.