IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 7.105 ⁜ Vol. 9, Issue 5, May 2022

DOI: 10.17148/IARJSET.2022.95107

CROP PRICE PREDICTION USING MACHINE LEARNING TECHNIQUES

Neeraj Soni[1], Jagruti Raut[2]

Student, M.Sc. (CS), Department of Computer Science, Viva College of Arts, Commerce, Science Virar(W), India[1]

Assistant Professor, Department of Computer Science, Viva College of Arts, Commerce, Science Virar(W), India[2]

Abstract: Farmers play a very important role in the agriculture sector. When the prices fall after the harvesting of the crops, farmers face massive amounts of losses because this country's GDP is affected by the price fluctuations of agricultural products. Crop price evaluation and estimation are done to take an intelligent decision before farming a specific type of seasonal crops. Predicting the price of a crop will help in making better decisions which results in minimizing the losses and managing the risk of price fluctuations. In this research paper, we had predicted the price of different crops by analysing the previous rainfall and Wholesale Price Index (WPI) data. We used the Decision Tree Regressor and Random Forest, a Supervised machine learning algorithm to analyse the previous data and Time Series Analysis to estimate the crop price for upcoming twelve months.

Keywords: Decision Tree Regressor, Random Forest, Machine Learning, Time Series Forecasting

I. INTRODUCTION

India is understood as an Agrarian country as about 55% of India's population depends on agriculture or related activities for his or her livelihood. Agriculture features a significant contribution to the first sector of India's economy. At present farmers face massive loss, its reasons are infertile soil, global climate change, oversupply within the market and there are tons of other uncertainties involved. Prediction of expected price in the future is very much needed to manage and sell the products at the right time to maximize revenue and minimize loss.[2]. In past years, price prediction was done by judging the farmers experience on particular crop and field. Suppose that we have the previous data available in which various corresponding price predictions are recorded and these recorded price predictions are used to classify future price predictions [8]. The price forecasting is important for farmers also as they base their production and marketing decisions on the expected prices that may have financial repercussions many months later [9]. This prediction is based on data obtained from government sources and it uses Machine Learning techniques. Data mining is also useful for predicting the crop price. Generally, data mining is the process of analysing data from different perspectives and summarizing it into useful information. Crop price prediction is an important agricultural problem. Each and every farmer always tries to know, how much price he will get from his expectation [6]. Considering Machine Learning/Deep Learning based models, with new price data arriving every day, updating the models might cause stability issues because of quality issues associated with the crop price data [7]. This website can also potentially benefit other agriculture-based industries to strategize sourcing of raw material at a reasonable price. The research is based on machine learning which will provide farmers and people in allied professions with a prediction which they can access anytime and will present them with predicted price trends in various crops 12 months in advance. Thus, a ML model with good prediction accuracy can prove to be a great asset for our feeding farmers.

II. RELATED WORK

[1] In this paper, machine learning techniques used to predict crop prices in advance. Also shown proper analysis of the crop and presented future scenario so that farmers can select right crop to strategize crop production.

[2] This paper aims at predicting both the price and profit of the given crop before sowing. The parameters considered for Price Prediction are Rainfall, Maximum-trade, Minimum Support Price, Yield. The parameter considered for Profit Prediction are Crop price, Yield, Cultivation Cost, Seed Cost. Naïve Bayes Algorithm was used to predict price of a crop and K-Nearest Neighbor (KNN) was used to predict profit of a crop.

[3] In this paper, the monthly prices of arecanut in Kerala are predicted using time-series model and machine learning models. The model SARMIA, Holt-Winter's Seasonal method, and LSTM neural network were used, and their performance was evaluated based on the value on the arecanut dataset with prices from 2007 to 2017. LSTM neural network model was found to be the best model that fits the data.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 7.105 ⁜ Vol. 9, Issue 5, May 2022

DOI: 10.17148/IARJSET.2022.95107

[4] In this paper, detailed forecasts for the next 6 month had been made. Decision Tree Regressor, a supervised learning algorithm was used. The parameters considered for prediction were rainfall data and crop price data.

[5] In this paper, Auto Regressive Integrated Moving Average (ARIMA) model was used for forecasting future prices of paddy. The performance of the model was examined by computing various measures of goodness of fit. The results showed that ARIMA model was the best suited model for forecasting.

III. PROPOSED SYSTEM & METHODOLOGY

We are creating the Crop Price Prediction Website for crop forecasting where we will take data from the Government website of more than 20 crops and represent the information about the rise and decrease in the costs of crops per month. Further, it will display various crop details like its type, location and export factors which will be convenient for the farmers to plan and manage crop harvesting. The result data visualization had been done with the help of chart and graphs. Initially, data from Government website data.gov.in will be taken comprising of rainfall and wholesale prices respectively month wise for every crop. Necessary pre-processing had been applied to remove noise from data and finally model had been trained. If found suitable, front and backend are going to be designed and therefore the ML model will be deployed at the backend. Requisite updating is going to be timely done on the dataset model are going to be redeployed. Here we do supervised learning because we've multiple inputs, an output and we are deriving a correlation between them. The two options suitable for this are Random Forest Regression and Decision Tree Regression .The algorithm will take inputs: The input parameters like months, year and monthly rainfall are passed to Jupiter notebook for training ML model and Visual Studio to develop frontend and backend. Data ingestion is to be through with the info collected from various sources. The injected data is to be prepared according to the requirement of the system. The Machine Learning model is to be designed and trained using the prepared data. Evaluation of the model is to be done using standard metrics. Model retraining has been done if the results are not satisfactory. When the desired results are achieved deploy the system.

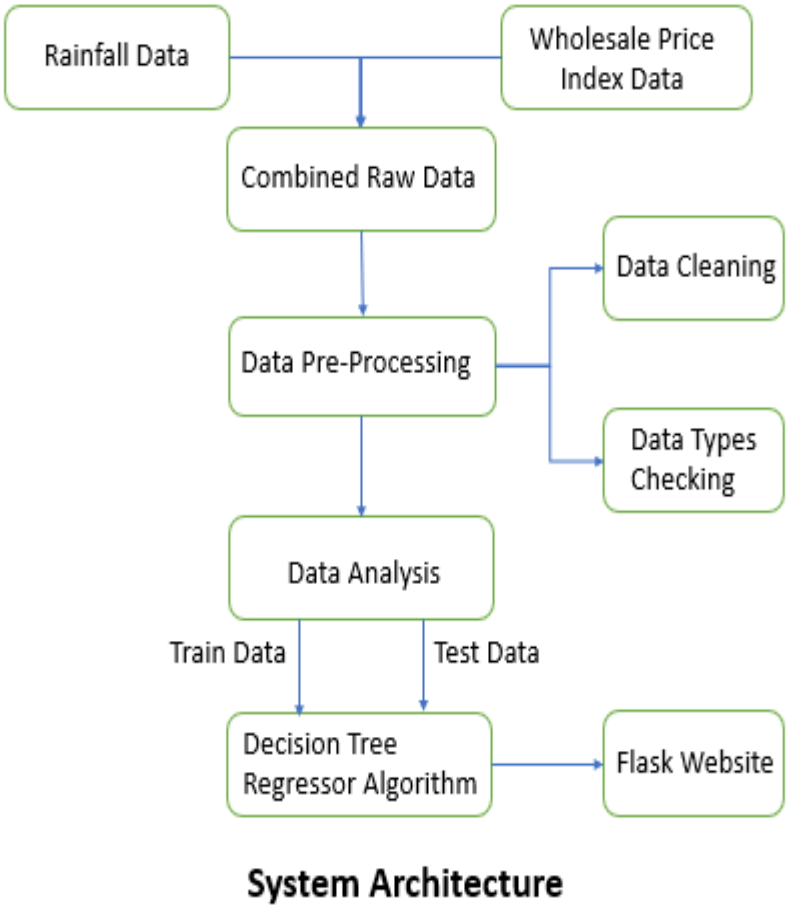


Fig. 1 System Architecture

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 7.105 ⁜ Vol. 9, Issue 5, May 2022

DOI: 10.17148/IARJSET.2022.95107

ALGORITHM SELECTION:

Random Forest Regressor: Random Forest Regressor is a supervised learning technique that uses an ensemble learning method for regression. Ensemble learning method is a technique that combines the predictions from the multiple machine learning algorithms to make a more accurate prediction than a single model.

Decision Tree Regressor: The dataset is going to be divided into multiple leaflets which are results of multiple decisions of yes and no then the new data are going to be calculated based on what leaflet they land then calculate the typical of that leaflet. It works on the idea of ensemble learning, which says that if we combine multiple algorithms or an equivalent algorithm multiple times then we will create a superior algorithm. Random Forest makes use of multiple decision trees to offer the output. As we've an enormous dataset. Random forest will first extract a little chunk of that data feed it to at least one decision tree regression model and the chain of that model this process is going to be repeated multiple times. we will control by specifying the worth of what percentage decision tree regressions we want. Now we got may be a huge dataset on which multiple decision trees are trained. Now the testing data is provided to each of the choice tree, and that they will give the output according to themselves and every one the gathering of output will be then averaged.

Time Series Forecasting:
Making predictions about the future is called extrapolation in the classical statistical handling of the time series data. More modern fields specialise in the subject and ask it as statistic forecasting. Forecasting involves taking models fit on historical data and using them to predict future observations. Descriptive models can borrow for the longer term (i.e. to smooth or remove noise), they only seek to best describe the information. An important distinction in forecasting is that the longer term is totally unavailable and must only be estimated from what has already happened.
The purpose of your time series analysis is usually two folds: to know or model the stochastic mechanisms that provides rise to an observed series and to predict or forecast the longer-term values of a series based on the history of that series.
The skill of a statistic forecasting model is decided by its performance at predicting the longer term. This is often at the expense of being able to explain why a specific prediction was made, confidence intervals and even better understanding the underlying causes behind the problem.

Training and Testing:
The dataset is split into two sets, training dataset and testing dataset. 80% for training the dataset & 20% for testing the dataset. The training dataset is used to train and fit the model while the test dataset is used to test the accuracy on an algorithm of trained model.

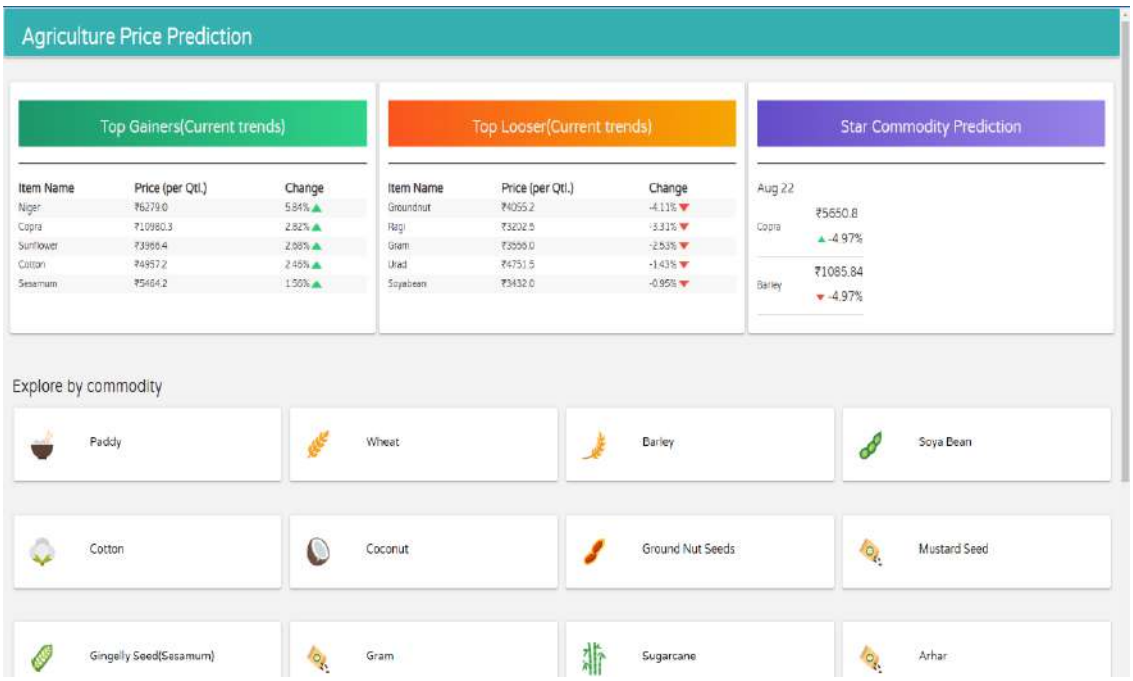


Fig. 2 Price Prediction Dashboard

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 7.105 ⁜ Vol. 9, Issue 5, May 2022

DOI: 10.17148/IARJSET.2022.95107

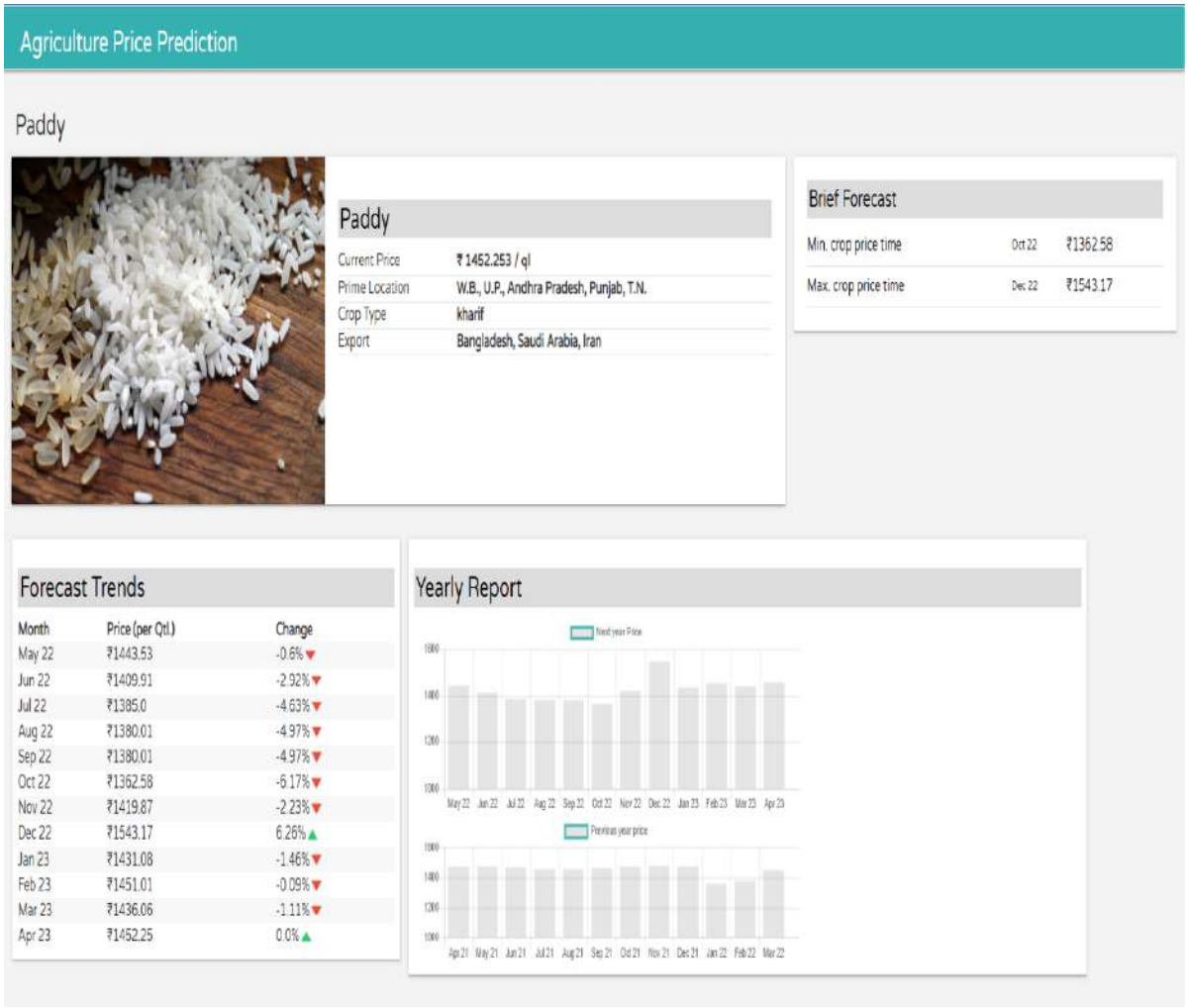


Fig 3. Commodity Details Dashboard

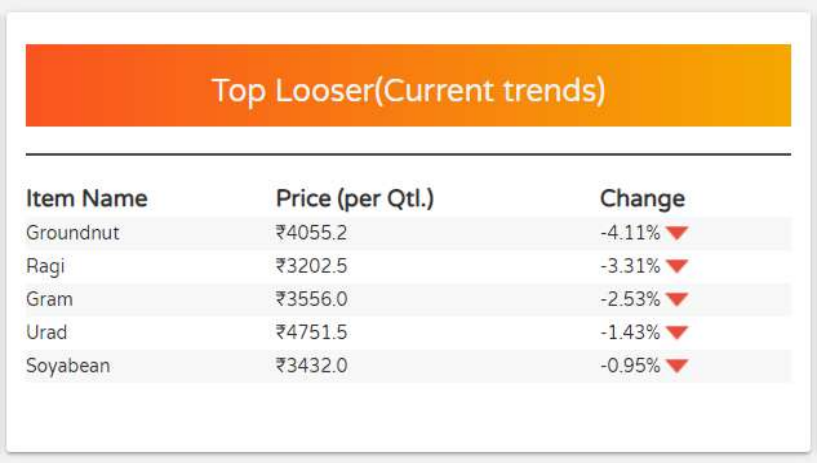
Fig 4. Top 5 Gainer

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 7.105 ⁜ Vol. 9, Issue 5, May 2022

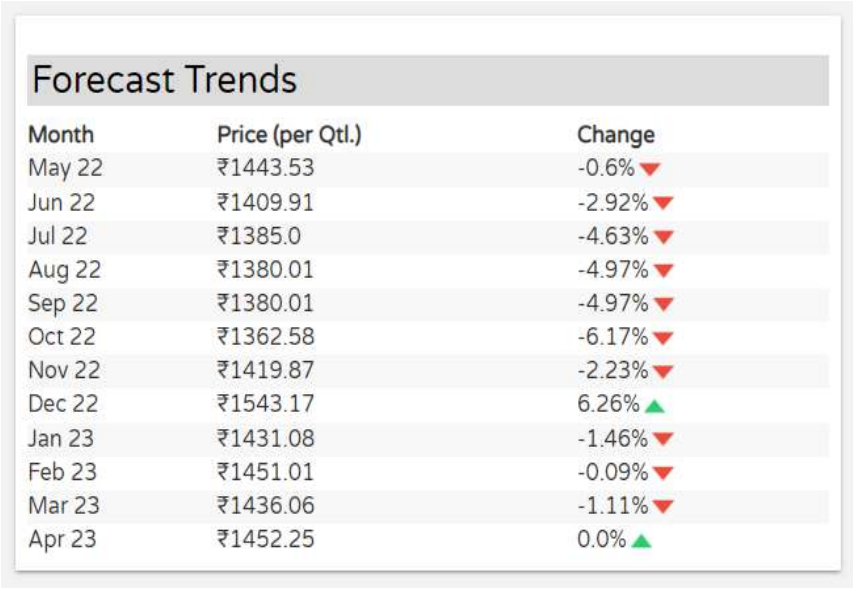
DOI: 10.17148/IARJSET.2022.95107



Top Looser(Current trends)

Item Name	Price (per Qtl.)	Change
Groundnut	₹4055.2	-4.11%
Ragi	₹3202.5	-3.31%
Gram	₹3556.0	-2.53%
Urad	₹4751.5	-1.43%
Soyabean	₹3432.0	-0.95%

Fig 5. Top 5 Looser



Forecast Trends

Month	Price (per Qtl.)	Change
May 22	₹1443.53	-0.6%
Jun 22	₹1409.91	-2.92%
Jul 22	₹1385.0	-4.63%
Aug 22	₹1380.01	-4.97%
Sep 22	₹1380.01	-4.97%
Oct 22	₹1362.58	-6.17%
Nov 22	₹1419.87	-2.23%
Dec 22	₹1543.17	6.26%
Jan 23	₹1431.08	-1.46%
Feb 23	₹1451.01	-0.09%
Mar 23	₹1436.06	-1.11%
Apr 23	₹1452.25	0.0%

Fig 6. 12 month Forecast Trends of selected Crops

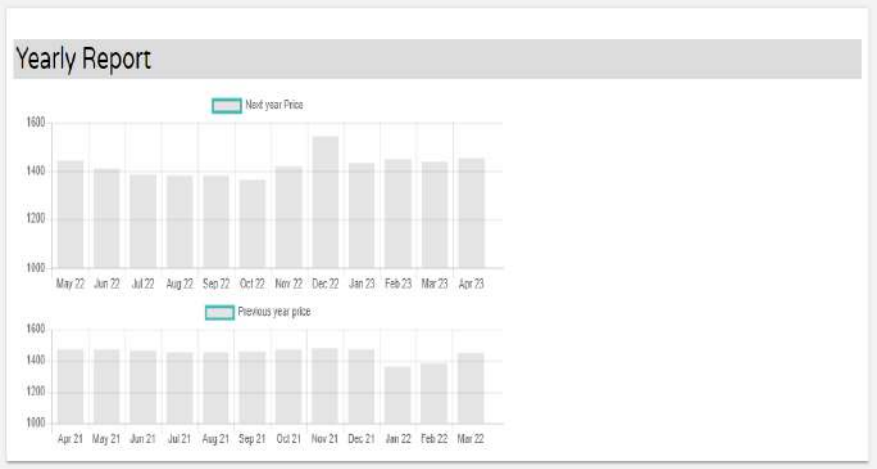


Fig 7. Yearly Report of Crop (Previous and Upcoming year)

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 7.105 ⁜ Vol. 9, Issue 5, May 2022

DOI: 10.17148/IARJSET.2022.95107

IV. RESULT AND ANALYSIS

i. Using Decision Tree Regressor:

As per the Government data Current Price of Wheat is approx. Rs. 1774.1/quintal. After the analysis on the given data sets of last 6 months, it concludes that at 20th October 2021 Minimum crop price is approx. Rs. 1876.9/quintal and at a month of 20th December 2021 Maximum crop price is Rs. 1972.65/quintal. From this, we have taken a Single month to check the accuracy of our model and as per the data sets Price of Wheat is Rs. 1695.8/quintal and the predicted price is Rs. 1621.45/quintal, so there is a Change of -4.75% approx. From above analysis, we conclude that there is a 92% Accuracy overall, and it may change slightly for each month.

ii. Using Random Forest Regressor:

As per the Government data Current Price of Wheat is approx. Rs. 1774.1/quintal. After the analysis on the given data sets of last 6 months, it concludes that at 20th October 2021 Minimum crop price is approx. Rs.1736.7/quintal and at a month of 20th December 2021 Maximum crop price is Rs. 1822.05/quintal. From this, we have taken a Single month to check the accuracy of our model and as per the data sets Price of Wheat is Rs. 1592/quintal and the predicted price is Rs. 1489.36/quintal, so there is a Change of -6.10% approx. From above analysis we conclude that there is a 88% Accuracy overall, and it may change slightly for each month.

By using both the machine learning Technique, we conclude that Decision Tree Regression is more accurate than the Random Forest Regression.

V. CONCLUSION

The research aims at predicting both the price and profit of the given crop before sowing. This web application runs on efficient machine learning algorithms and technologies having an overall user-friendly interface to the users. The training datasets so obtained provide the enough insights for predicting the appropriate price and demand in the markets. Thus, the system helps the farmers in reducing their difficulties and stop them by attempting suicides.

REFERENCES

[1]. I. Ghutake, R. Verma, R. Chaudhari, and V. Amarsinh, “An intelligent Crop Price Prediction using suitable Machine Learning Algorithm,” ICACC-2021, ITM Web of Conferences 40.

[2]. Rachana P S, Rashmi G, Shravani D, Shruthi N and Seema K R, “CROP PRICE FORECASTING SYSTEM USING SUPERVISED MACHINE LEARNING ALGORITHMS,” IRJET-2019.

[3]. K. M. Sabu, T. K. M. Kumar, “Predictive analytics in Agriculture: Forecasting prices of Arecanuts in Kerala,” Procedia Computer Science-2020.

[4]. R. Rohit, R. Vishnu, A. Kishore and D. Chakkarawarthi, “Crop Price Prediction and Forecasting System using Supervised Machine Learning Algorithms,” IJARCCE-2020.

[5]. A. Darekar and A. A. Reddy “Forecasting of Common Paddy Prices in India,” Journal of Rice Research 2017, vol.10, No.1

[6]. B. Sahithi, T. Saheli, D. Ramanika, N. Anil Kumar, “Crop Price Prediction and Forecasting System using Supervised Machine Learning Algorithms,” Journal of Software Engineering and Simulation 2020, vol.6, Issue 1.

[7]. A. Jain, S. Marvaniya, S. Godbole, V. Munigala, “A Framework for Crop Price Forecasting in Emerging Economies by Analyzing the Quality of Time-series Data”, 2020.

[8]. M. Kaur, H. Gulati, H. Kundra, “Data Mining in Agriculture on Crop Price Prediction: Techniques and Applications,” International Journal of Computer Applications (0975 – 8887), vol.99, no.12.

[9]. G. Jha and K. Sinha, “Agricultural Price Forecasting Using Neural Network Model: An Innovative Information Delivery System.”

