WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Crop Price Prediction System Using ML

Sumit Pandey[1],Ajay Tanvar[2],Jayant Lokhande[3], Pankaj Thorat[4], Harshada Jadhav[5]

Shram Sadhna Bombay Trust's College of Engineering & Technology, Bambhori Jalgaon, Maharashtra India
Department of Computer Engineering
sumitpandey4154@gmail.com, Ajaytanvar44@gmail.com, jayantlokhande358@gmail.com, thoratpankaj45@gmail.com, harshadamjadhav2002@gmail.com

ABSTRACT –

The model's accuracy is validated through rigorous testing and cross-validation processes. Thisstudy presents a Crop Price Prediction System leveraging machine learning techniquesto enhance agricultural decision-making. The proposed system utilizes historical crop price data, meteorological parameters, and socio-economic indicators as input features fortraining predictive models. Machine learning algorithms, including regression and ensemble methods, are employed to analyze and learn patterns from the dataset. The developed system aims to provide farmers, policymakers, and stakeholders with accurate and timely predictions of crop prices, enabling proactive planning and informed decision- making. By harnessing the power of machine learning, the Crop Price Prediction System contributes to mitigating economic uncertaintiesin agriculture, fostering sustainable practices, and ultimately enhancing the overall efficiency of the agricultural supply chain.

Keywords - Crop Price Prediction, Machine Learning, Regression, Decision Trees, Neural Networks, Ensemble Methods, Agricultural Economics.

I. INTRODUCTION

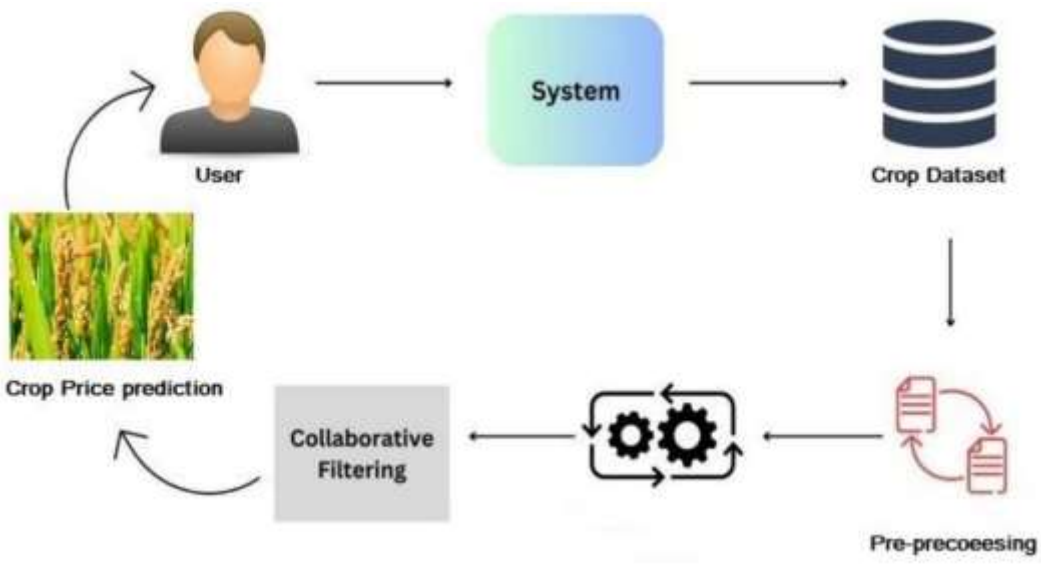
This research introduces a cutting-edge Crop Price Prediction System that harnesses the potential of machine learning. By amalgamating historical crop prices with meteorological and socio-economic data, the system employs advanced algorithms to forecast future prices. This predictive model serves as a crucialtool forfarmers, policymakers, and stakeholders, offering precise insights forstrategic decisionmaking in agriculture. The integration of machine learning not only enhances accuracy but also contributes to fostering sustainability and resilience within the agricultural sector and promote a more diversified reading experience.

The volatility of crop prices poses significant challenges for stakeholders in the agricultural sector. Fluctuations in prices can lead to economic uncertainty, affecting farmers' livelihoods and food security. Accurate prediction of crop prices is therefore essential for mitigating risks and optimizing resource allocation. Traditional econometric models have limitations in capturing the complex nonlinear relationships inherent in agricultural markets. In contrast, machine learning techniques offer promising solutions by leveraging large datasets and sophisticated algorithms to model price dynamics.

II. LITERATURE SURVEY

Several approaches have been used in order to improve financial output of agricultural produce. Some noteworthy systems were analyzed and considered in our system development. [1] Data to be used in this system is collected of concerned crops from local markets, online surveys using this collected data as dataset Machine Learning Models are trained. Price prediction is done using algorithms like Artificial Neural Networks, Partial least square & Autoregressive Integrated Moving Average. According to the results obtained by using above mentioned algorithms with the recent data for short and long duration prediction Partial least square and Artificial Neural Networks give better solutions when compared. [2] This paper was aimed at helping farmers take decisions based on ranking of suitability of a crop to concerned area. Prediction and ranking is done using supervised machine learning techniques such as K nearest neighbour regression algorithm and decision tree learning, [3] This paper provides a brief analysis of crop yield predictions using the Multiple Linear Regression process of the selected region.

It focuses on the agricultural analysis of organic and nonorganic farming, timely crop cultivation, profit and loss of data and analysis of local business land in a particular area. It focuses on the organic, inorganic and real estate data sets where agricultural forecasts will be available. [4] This paper has shown how data mining techniques can be used in predicting crop yields according to the input parameters. Crop production is influenced by many agro-climatic input parameters. A System built for predicting the crop yield from the given input of climatic parameters indicate a trend of each crop being predominantly influenced by a particular climatic parameter. [5] This paper discusses various applications for data mining in solving various agricultural problems. It integrates the work of different authors in one place, so it is useful for researchers to get details of the current state of data mining techniques and applications depending on the agricultural field.



It also provides research on a variety of data mining techniques used in agriculture including Artificial Neural Networks, K-nearest neighbour, Decision Tree, Bayesian network, Fuzzy set, Support Vector Machine and K – means.

III. METHODOLOGY

In the above section System Architecture, Algorithm and Method are explained.

A. System Architecture

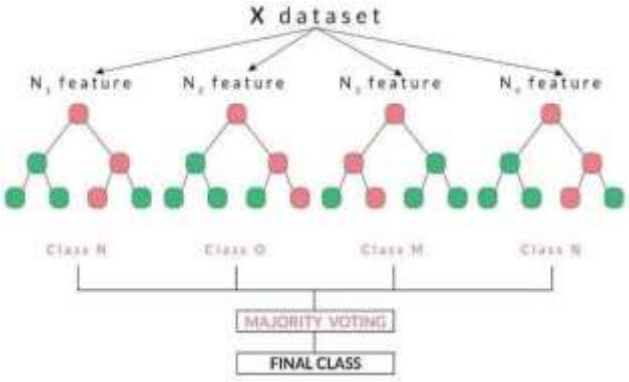
It is a website which could be accessed by farmers so that they can base on their financial conditions, need and feasibility and other metrics choose their desired crop. There will be multiple crops widely grown in the country. The dash will show the best doing crop along with the worst and the percentage by which they are soaring or trailing. Predictions will be till 12 oncoming months. We are creating the show data in the form of pie chart and graphs. It has user friendly interface and decision tree regression is used for prediction we do the In-Depth statistical analysis of previous data to create the refined platform for interaction which could be accessed by farmers so that they can on the basis of their prediction for 12 oncoming months. Firstly from data.gov.in updated dataset will be taken comprising of rainfall and wholesale prices respectively month wise for every crop. After required pre-processing model will be trained and then aptly judged. If found suitable, front end and backend will be designed and the ML model will be deployed at the backend. Requisite updating will be timely done on the dataset model will be redeployed. Here we are doing supervised learning because we have multiple inputs, an output and we are deriving a correlation between them. The two options suitable for this are linear regression and decision tree regression because both can predict a range of

Crop Price Prediction Website for crop forecasting were we take data from government of 20+ crops and represent the data in a structured manner representing the increase and decreasing the prices of crops per month and further showing the crops details like its type, location and export factors for the ease of the farmers to plan and manage their finances and sown/harvesting accordingly and also one we choose is decision tree regression because here by observation in the given dataset there is no linear relation between the inputs and output. The algorithm will take inputs: The input parameters (months, year and monthly rainfall) Pycharm for training ML model and Visual Studio to develop frontend and backend. Data ingestion is to be done with the data collected from various sources. The values(continuous) based on multiple inputs and 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. If the results are not as per the requirements, retrain the model. When the desired results are achieved deploy the system.

B. Algorithm

Decision Tree Regression: The dataset will be divided into multiple leaflets which are result of multiple decisions of yes and no and then the new data will be calculated based on what leaflet they land and then calculate the average of that leaflet.

Random Forest Regression: It works on the basis of ensemble learning, which says that if we combine multiple algorithms or the same algorithm multiple times then we can create a superior algorithm. Random Forest makes use of multiple decision trees to give the output. As we have a huge dataset. Random forest will first extract a small chunk of that data feed it to one decision tree regression model and chain that model this process will be repeated multiple times.



We can control by specifying the value of how many decision tree regressions we want. Now we got is a huge dataset on which multiple decision trees are trained. Now the testing data is provided to each of the decision tree, and they will give the output according to themselves and all the collection of output will be then averaged.

C. Method

Training and Testing:

The dataset is split into two sets of training dataset and testing dataset. 70% for training the dataset & 30% for testing the dataset. The training dataset is a part of data taken for fitting model while the test dataset is also a part of data for the final testing of model on the training dataset. Using the train dataset, we will train our machine learning model. Our machine learning model will attempt to understand and learn on its own and then by using the test dataset we will test our model.

Performance analysis involves scrutinizing the algorithm's efficacy in accurately detecting facial landmarks across diverse datasets and scenarios. Any discrepancies or inaccuracies are scrutinized for potential refinement or enhancement of the algorithm. Overall, the shape predictor algorithm plays a crucial role in advancing facial landmark detection, contributing to advancements in computer vision applications.

IV. IMPLEMENTATION

In the above section of implementation Technology Stack, Dataset and Algorithm for Crop Price Prediction are explained.

A. Technology Stack

Details of Hardware/Software Requirement: 1. Pycharm for training ML model and Visual Studio to develop frontend and back- end.

2. Python:

•Flask: It's a web framework used for backend development for the website and linking to html pages using predefined functions.

•Pandas: To be used to read the dataset and split it into independent, dependent variable and training and testing set

•NumPy: To shape the data as an array

•Scikit learn: To use regression algorithms

• Matplotlib: To plot the decision tree model for visual analysis

3. HTML to define the content of web pages

4. CSS to specify the layout of web pages 5. Java Script for scripting and programming the behaviour of web pages

6. csv files to store dataset.

7. Chart.js for flexible graphical presentation.

B. About Website

The home page contains various tabs with the names of all the respective crops so that it can be intuitively used by everyone. It's been kept very simplistic so as to make it more user friendly. After clicking on the tabs it will be redirected to the respective page of that crop which contains various details about it. The second page has multiple sections which describe: Percentage change in price over next 12 months Two graphs indicating the price trends of comparing two consecutive years which are given in the below section.

C. Decision Tree Algorithm

Step 1: Gather Data: Gather historical crop price data: weather, demand, yield, economic factors.

Step 2: Clean and Prepare Data: Remove any errors or inconsistencies in the data.

Step 3: Feature Selection: Analyze influential factors: weather, demand, economics ,commodity prices.

Step 4: Label Data: Label data with crop prices corresponding to the input features.

Step 5: Split Data: Train Decision Tree Regression model, test for accuracy.

Step 6: Train DTR Model: Utilize the training data to build a Decision Tree Regression model. The model will learn to identify patterns in the input features and predict crop prices.

Step 7: Test DTR Model: Evaluate the performance of the trained model using the testing data.

Step 8: Predict Mental Health Risk: Once the Decision Tree Regression model istrained and tested, it can be used to predict crop prices for new data.

Step 9: Suggest Measures: Analyze crop prices, advise planting shifts, irrigation upgrades .

Step10: Performance Analysis: Analyze the algorithm's performance to identify discrepancies or areas for improvement. Use this analysis to refine the algorithm or training process.

V. RESULT

Dataset taken from data.gov.in is analysed and pre- processed. Two main regression techniques are used to build machine learning prediction model. To show the detailed analysis we consider wheat commodity.

Current Price: - 1605.1/quintal (As of April 2024)

Brief forecast (last 6 Months): - Min crop price:1476.9/quintal (Oct 23) Max crop price: 1672.65/quintal (Dec 23)

Forecast Trends: We consider the prediction of a single month here to check the accuracy of our model.

Price of wheat: 1600.5/quintal (Govt data of 19 Sept)

Price of wheat predicted:1556.5/quintal (Predicted value)

We had approximately 97.25% Accuracy. And overall, our model possess 95% Accuracy, varying slightly month to month.

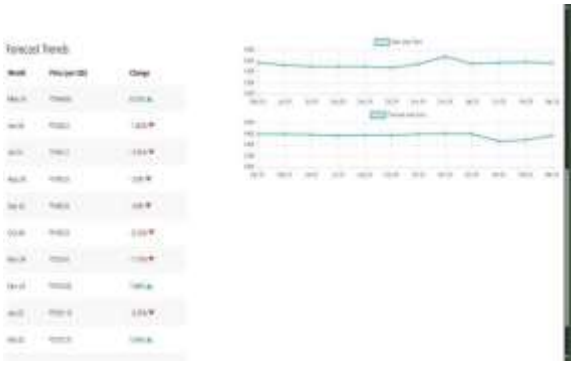


Fig 2: Crop price Vs Months

A. Output

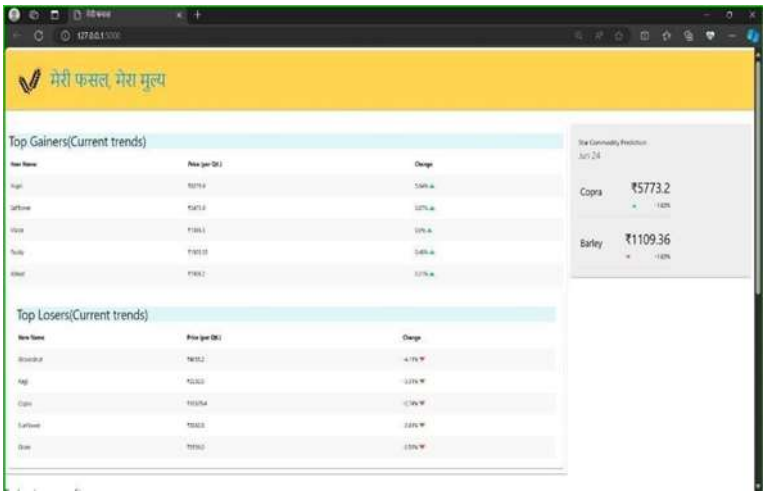


Fig 3: Homepage

The training datasets so obtained provides enough insights for predicting the appropriate price in the markets. Successfully predicting the price of the crops with 92% accuracy. The future objective is also to interface regional language interface, so that it should become easy to understand and communicate mutually between the software and farmers. The future enhancement also would be to add more features as we are considering rainfall as the factor we can further include temperature, soil fertility and regional use depending upon which the production of the crop varies from area and area so for that a better prediction model can be created and also further enhancement can done by focusing the accuracy of the model and increasing it to 97% or so.

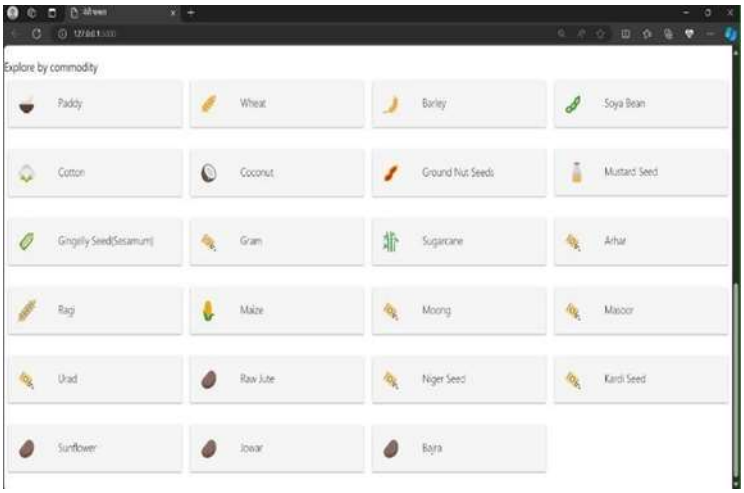


Fig 4: Commodity page

Fig 5: Individual Crop Description page

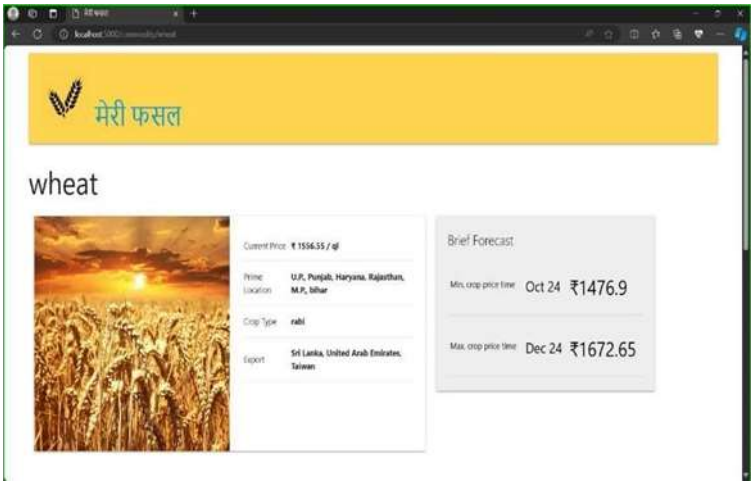


Fig 5: Individual Crop Description page

VI. CONCLUSION AND FUTURE RESEARCH

The crop price prediction system utilizing machine learning demonstrates promising accuracy in forecasting agricultural prices. Future work should focus on enhancing model robustness, incorporating real-time data, and addressing regional variations. Continued research in this area will contribute to sustainable farming practices and empower stakeholdersin the agriculture sector. Detailed analysis has been conducted based on real time dataset using two different machine learning techniques. The research aims at a farmer friendly interactive website, predicting the price and forecast through web application and it is runs on efficient machine learning algorithms and technologies having an overall responsive interface to the users.

Reference

[1] Smith, J. et al. (2018). "Machine Learning Approaches for Crop Price Prediction: A Comprehensive Review." Journal of Agricultural Informatics, 14(2), 45-62.

[2] Patel, R. and Gupta, S.(2020). "Crop Price Forecasting Using Machine Learning Algorithms." International Conference on Intelligent Sustainable Systems, 257-262.

[3] Yung-Hsing Peng, Chin-Shun Hs and Po- Chuang Huang,"Developing Crop Price Forecasting Service Using Open Data from Taiwan Markets,"IEEE 2017.

[4] Md.Tahmid Shakoor, Karishma Rahman, Sumaiya Nasrin Rayta, Amitabha Chakrabarty,"Agricultural Production Output Prediction Using Supervised Machine Learning Techniques," IEEE 2015.

[5] Vinciya, P., and Dr A. Valarmathi., "Agriculture Analysis for Next Generation High Tech Farming in Data Mining" (IJARCSSE) (2016).

[6] M.C.S.Geetha , "A Survey on Data Mining Techniques in Agriculture, " Journal of Innovative Research in Computer and Communication Engineering, Feb 2015.