e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

FARMWAVE THE CROP GUIDER

Saniya Kishor Sawarkar*1, Ketaki Kaustubh Tamhane*2,
Bhagyashri Gopal Pedapallikar*3, Hemant Tighade*4, Amarsinh Vidhate*5
*1,2,3,4Student, Department of Computer Engineering, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, India
*5HOD, Department of Computer Engineering, Ramrao Adik Institute of Technology, Nerul, Navi
DOI : https://www.doi.org/10.56726/IRJMETS37828

ABSTRACT

Agriculture plays a vital role in our country's social and economic development, but many farmers struggle with maintaining good farming facilities. The government provides various policies to support farmers, and today, farmers are increasingly using the internet and smartphones for smart farming. The proposed system, called Farmwave, aims to make agriculture more connected and smarter by tracking weather conditions and predicting crop disease and proper fertilizer to improve yield quality and reduce costs. This system provides valuable information to farmers and helps manage demand forecasts, bring goods to market, and reduce waste. [2] Overall, agriculture is an important source of livelihood for the majority of our population, contributing to national income and employment. The suggested system aims to monitor weather patterns in a particular area, including temperature, humidity, and environmental shifts. Its purpose is to reduce the time and energy farmers spend by providing them with daily market prices for various crops, fertilizers, and vegetables without requiring them to visit the market. The data, which includes values, climate, and the latest rural techniques, will be transmitted to the farmers via this application. [2] This application is extremely helpful for farmers as it keeps them informed and provides a source of income for a large portion of the population, which contributes to the country's overall revenue and employment opportunities. Also, This application is multilingual which helps to improve customer experience: By providing information in the customer's native language, you can improve their experience on your website, making it easier for them to understand and engage with your content. This can help to build trust and loyalty with your customers.

Keywords: Agriculture, Fertilizer, Climate Change, Data Preprocessing, Environmental change

I. INTRODUCTION

Agriculture is a crucial sector of the Indian economy, and farmers play a significant role in the country's GDP. However, farmers often struggle to get the right price for their harvests due to various reasons, such as improper watering, incorrect plant selection, or low crop yields. To address these challenges, new technologies are being introduced in agriculture to improve efficiency and productivity. Innovations such as precision agriculture, IoT, and computerized farming tools like ranch management systems are gradually being adopted by farms and agribusinesses.[**1**] These technologies go beyond just automating repetitive tasks and reducing human labor. By analyzing factors such as soil type, temperature, humidity, water level, spacing depth, soil pH, season, and fertilizers, farmers can predict the net yield of their crops and choose the most appropriate crop for their farm.

Machine learning algorithms have been found to be an efficient way to predict the most suitable crop for a given farm based on various factors. By adopting these technologies and methods, farmers can increase their productivity, improve the quality of their harvests, and ultimately secure better pricesfor their crops.[**3**]One of the significant challenges facing the agriculturalsector is the identification and prevention of plant diseasesthat can cause significant damage to crops. Another critical challenge is predicting crop yield accurately to ensure that there is enough food to meet the growing demand. In addition to this, weather prediction, market analysis, and fertilizer pre- diction are also crucial in improving crop yield and ensuring food security.The development of a model that can detect plant diseases, predict crop yield, forecast weather patterns, analyze market trends, and recommend appropriate fertilizers can significantly improve the agricultural sector's productivity and efficiency. This model can help farmers make informed decisions about crop management, reduce crop loss due to plant diseases, and increase crop yield to meet the growing demand for food.In this research paper, we present a model that integrates plant disease detection, crop prediction, weather prediction, market analysis,

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

and fertilizer prediction. We describe the methodology used to develop the model and evaluate its accuracy using real-world data. We believe that our modelcan significantly improve the efficiency and productivity ofthe agricultural sector, ultimately contributing to global food security.[**6**]

II. METHODOLOGY

A. The data preprocessing steps performed in our model are

1) In our model that includes a plant disease detector, crop prediction, weather prediction, market analysis, and fertilizer prediction, several data preprocessing steps were undertaken to ensure the accuracy of the results.
2) Firstly, data cleaning was performed to remove any irrelevant or incomplete data. The data was carefully curated to eliminate any outliers, errors, or missing values.
3) Secondly, data integration was carried out to combine data from various sources to generate a comprehensive dataset. This process involved merging data from differ- ent datasets with the use of common variables.
4) Thirdly, data transformation was applied to convert the data into a standard format. This step involved normal- izing the data to ensure that it conforms to a uniform scale and reducing the dimensionality of the dataset.
5) Fourthly, feature selection was performed to identify the most significant variables that impact the model's performance. This step involved selecting variables that are relevant to the problem being solved and discarding those that are irrelevant.
6) Finally, data splitting was carried out to create separate datasets for training, validation, and testing. This process involved dividing the data into randomly selected subsets to train the model, validate the model's accuracy, and test the model's performance on unseen data.
7) These data preprocessing steps were crucial in ensuring that the model could accurately predict crop yield, detect plant diseases, analyze market trends, predict weather patterns, and recommend appropriate fertilizers for specific crops.[3]

1. **Data Preprocessing**: Data preprocessing is a crucial step in machine learning to ensure that the data used for model training is clean, consistent, and ready for analysis.
   The following are the data preprocessing techniques:
2. **Data Collection :** Collect relevant datasets for each prediction task. For crop prediction, we used data on crop types, planting dates, weather data (temperature, precipitation, humidity, etc.), soil data (pH, nutrient levels, etc.), and other relevant features. For fertilizer prediction, we used data on crop nutrient requirements, soil nutrient levels, and fertilizer application rates. For plant disease prediction, we used data on plant health status, weather conditions, and disease occurrence.[8] For crop price prediction, Necessary elements includes historical price data, market trends, and other relevant factors.
3. **Data Cleaning**: Data cleaning is crucial for accurate predictions. Here are the steps we used for data cleaning:

a. Missing Value Handling: Checked for missing values in the datasets and decided on an appropriate strategy to handle them. we could either remove rows or columns with missing values or fill them with appropriate values using techniques like mean, median, mode imputation, or machine learning- based imputation methods.

b. Outlier Detection and Treatment: Identified outliers in the datasets using statistical methods or visualization techniques and decide on the appropriate treatment for them.[10] Outliers are removed and replaced with appropriate values based on domain knowledge.

c. Data Transformation: Converted categorical variables into numerical representation using techniques like label encoding or one-hot encoding. Standardize or normalize numerical vari- ables to brought them to a common scale, which can prevent biases in the model.

d. Feature Selection: Selected relevant features that have a significant impact on the prediction tasks. We used techniques like correlation analysis, feature importance from decision tree or random forest, or domain knowledge to select important features.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

4. **Data Preparation:** After data cleaning, we split the datasets into training and testing sets. The training set was used to train the decision tree and random forest models, while the testing set was used to evaluate their performance.

a. Decision Tree Model Building Here are the steps we used to build a decision tree model:
b. Chose a decision tree algorithm such as ID3, C4.5, or CART, and implemented it using a machine learning library like scikit-learn in Python.
c. Trained the decision tree model using the training set. The model learnt the patterns in the data and created a decision tree based on the selected features and target variable for each prediction task (crop type, fertilizer requirement, plant disease occurrence, crop price, weather condition).
d. Hence Evaluated the performance of the decision tree model using appropriate evaluation metrics such as accuracy, precision, recall, F1-score, etc.
e. Tuned the hyperparameters of the decision tree model, such as maximum depth, minimum samples required for split, etc., to optimize its performance.
f. Random Forest Model Building: Here are the steps to build a random forest model:
g. Chose a random forest algorithm and implement it using a machine learning library like scikit-learn in Python.
h. Trained the random forest model using the training set. The model created multiple decision trees with different subsets of the data and combined their predictions to make the final prediction for each prediction task.
i. Evaluated the performance of the random forest model us- ing appropriate evaluation metrics such as accuracy, precision, recall, F1-score, etc.
j. Tuned the hyperparameters of the random forest model, such as the number of trees in the forest, maximum depth of trees, minimum samples required for split, etc., to optimize its performance.

5. **Evaluation**

a) Evaluation Metrics: Used evaluation metrics for each pre- diction task to assess the performance of the models. For example, for crop prediction, we used accuracy or F1-score to measure the overall accuracy of predicted crop types. For fertilizer prediction, we used metrics like mean absolute error (MAE) or root mean squared error (RMSE) to measure the prediction accuracy of nutrient requirements.[10] For plant disease prediction, we used metrics like precision, recall, and F1-score to measure the prediction accuracy of disease occurrence. For crop price prediction, we used metrics like MAE or RMSE to measure the prediction accuracy of crop prices. For weather prediction, we used metrics like accuracy or mean squared error (MSE) to measure the prediction accuracy of weather conditions.
b) Cross-Validation: Performed cross-validation to assess the robustness of the models. Split the dataset into multiple folds, trained the models on different folds, and evaluated their performance on the remaining fold. Repeated this process multiple times with different fold combinations to get a more reliable evaluation of the models.
c) Model Selection: Compared the performance of the decision tree and random forest models using the evaluation metrics and cross-validation results. Selected the model that performs better on the evaluation metrics and shows higher accuracy and robustness in cross-validation.
d) Hyperparameter Tuning: Tuned the hyperparameters of the selected model to further optimize its performance. For example, we could try different values for the maximum depth, minimum samples required for split, number of trees in the forest, etc., and evaluate the performance of the model with each set of hyperparameters. [11]Choose the set of hyperparameters that results in the best performance.
e) Final Model Validation: Once we selected the best model and tuned its hyperparameters, validated the model using the testing set that was kept aside during data preparation. Evaluated the model's performance on the testing set using the chosen evaluation metrics. As the performance is satisfactory, we considered the selected model as the final model for prediction tasks.
f) Model Interpretation: Finally, interpreted the decision tree model to understand the rules or decisions it makes for prediction tasks. Decision trees are interpretable models, and we extracted valuable insights from the tree structure to understand the factors that influence the predictions.[13]

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

III. MODELING AND ANALYSIS

Model and Material which are used is presented in this section. Table and model should be in prescribed format.

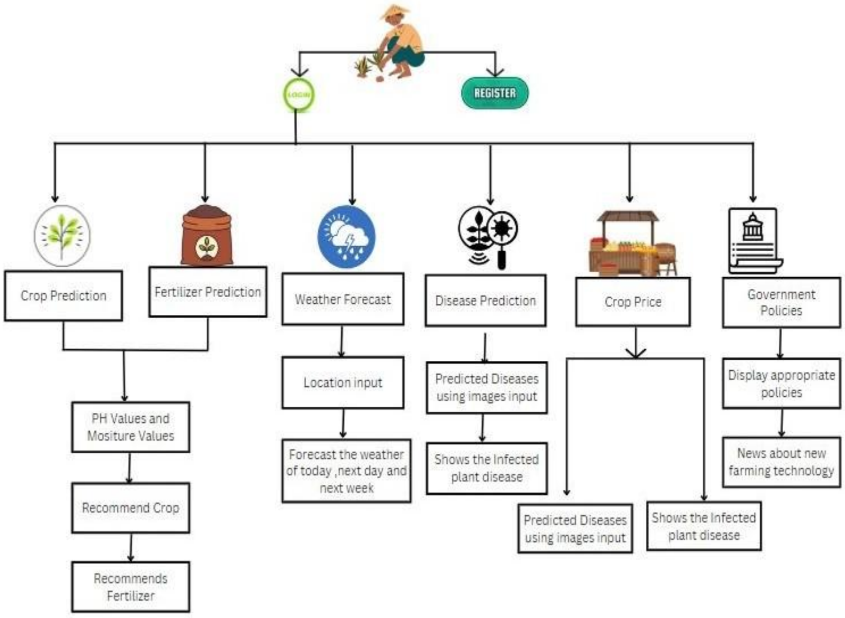


Figure 1: System Design

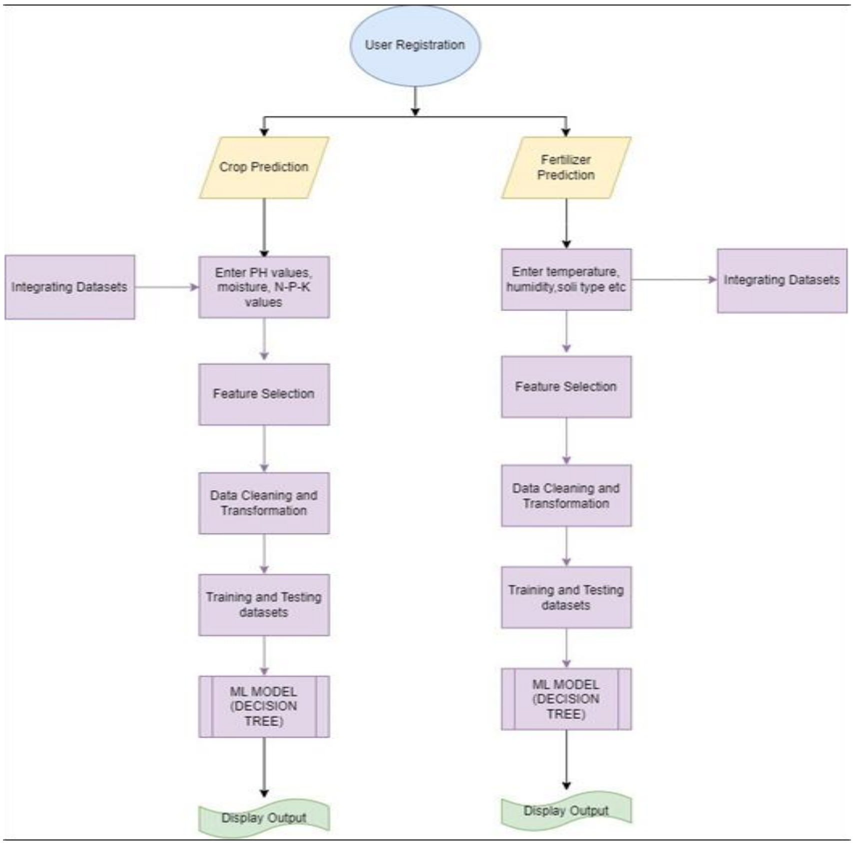


Figure 2: Flowchart for Crop Prediction and Fertilizer Prediction

In the above figure, the flowchart is designed to facilitate user registration, followed by a branching process into two separate pathways: crop prediction and fertilizer prediction. The crop prediction path involves the user inputting relevant variables such as Ph values, moisture, and NPK values. After this, feature selection is

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

performed, followed by data cleaning and transformation. Training and testing datasets are then used to develop the machine learning model, which will ultimately output the predicted crop type.

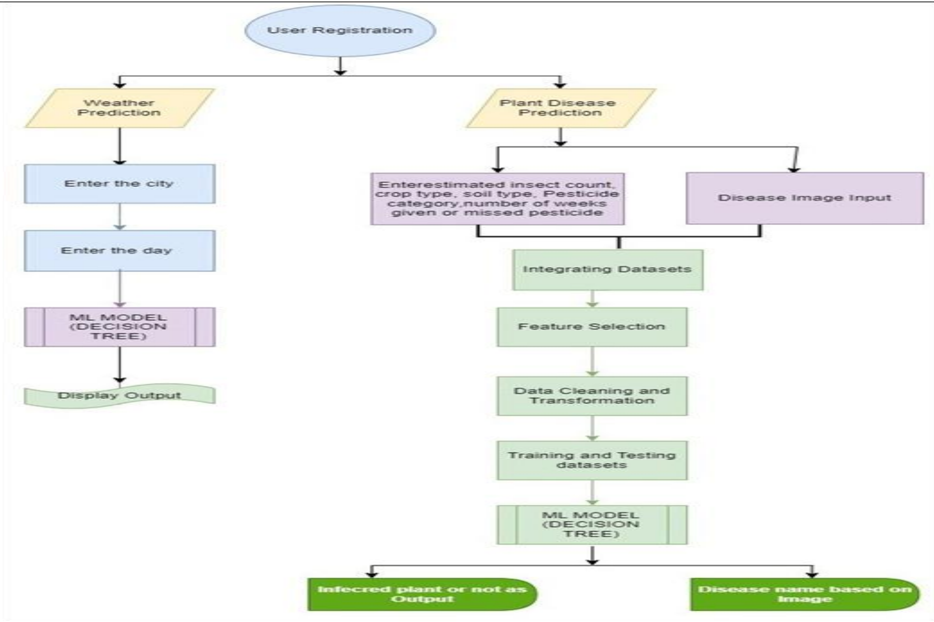


Figure 3: The flowchart starts with user registration and then branches into two different paths:

Weather prediction and Plant disease prediction.

In the weather prediction path, the user is prompted to enter the city and day for which they want to predict the weather. This input is then fed into a machine learning model, specifically a Decision Tree, which processes the data and generates an output displaying the predicted weather for the given city and day.In the plant disease prediction path, the user is asked to input various details such as the estimated count of infected plants, crop type, soil type, pesticide category, number of weeks since the last application of pesticide, and an image of the diseased plant.[12] The input data is then integrated into a dataset, and feature selection, data cleaning, and transformation are performed. The training and testing datasets are used to develop a machine learning model that can predict whether the plant is infected or not and also the name of the disease based on the provided image. In conclusion, this flowchart streamlines the process of weather prediction and plant disease prediction by utilizing machine learning models that can accurately generate output predictions based on the user inputs.

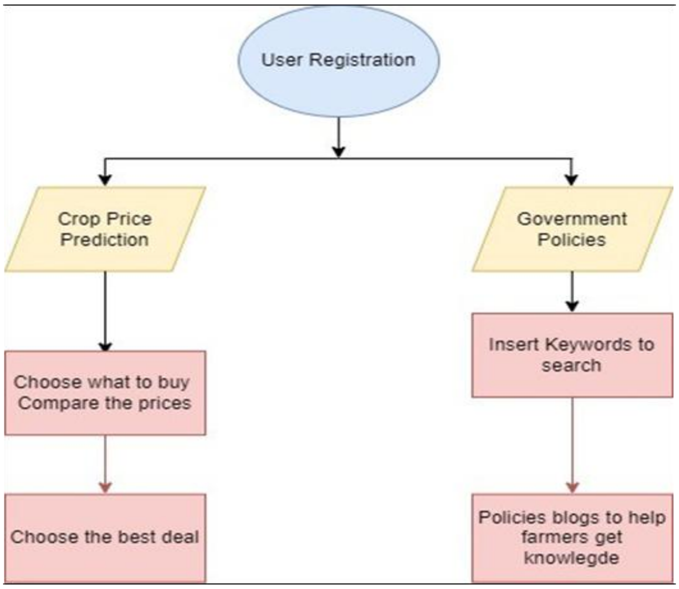
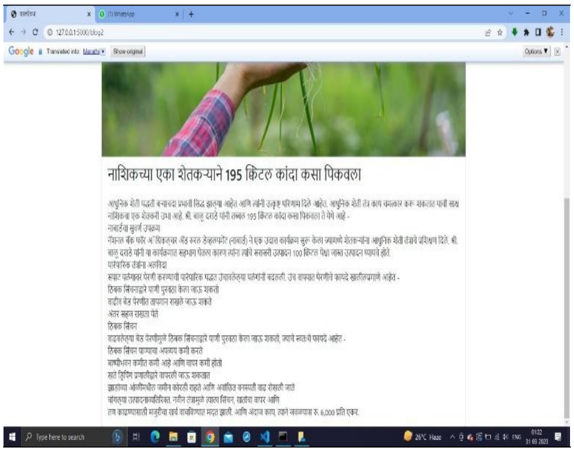
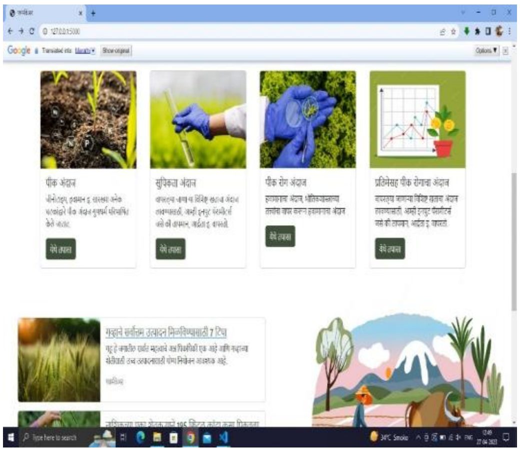


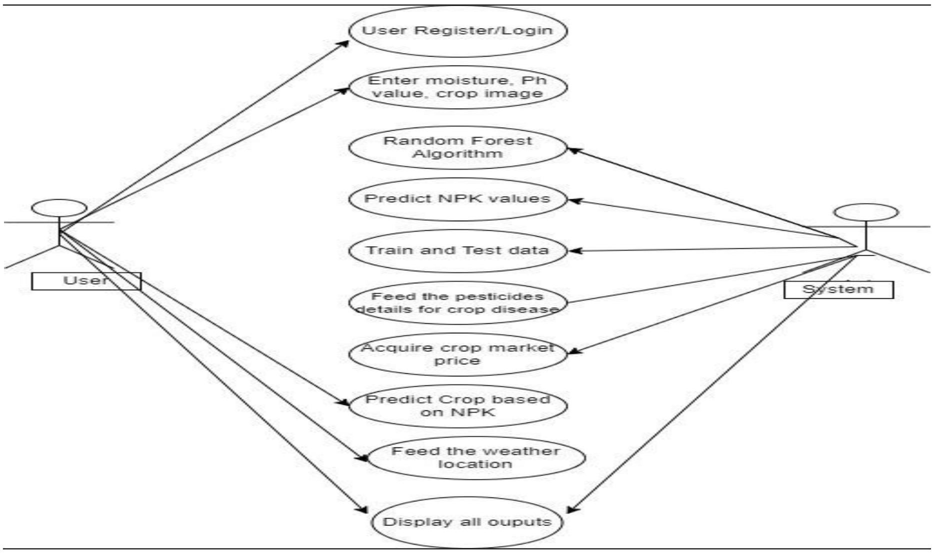
Figure 4: Flow chart for Crop Price Prediction and

IRJMETS
e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com





Government Policies. Implemented a multilingual feature in our system greatly benefited users who speak different languages. By providing site content in multiple languages, farmers and agricultural professionals from different regions and backgrounds can access and understand the information provided on the site with ease.



IV. RESULTS AND DISCUSSION

The accuracy of different machine learning algorithms for crop prediction. Hence Random Forest Algorithm Performs best of all other machine learning algorithms. Random Forest is an ensemble algorithm which combined multiple decision trees to improve the accuracy and reduce overfitting. Hence Random Forest is best suited algorithm for crop prediction and that resulted in best percentage of Accuracy.

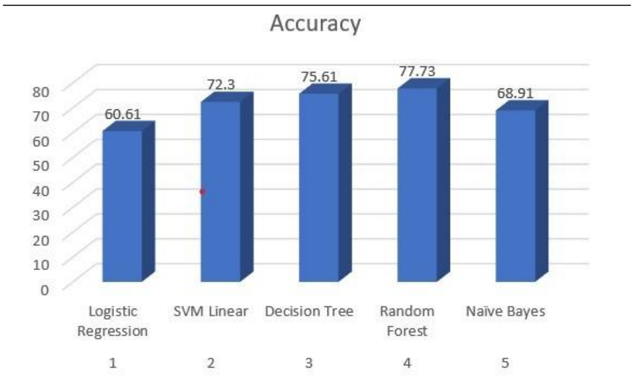


Table 1. COMPARATIVE STUDY FOR CROP PREDICTION

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

Models Name	Test Accuracy in Percentage
Logistic Regression	60.61%
SVM Linear	72.3 %
Decision tree	75.61 %
Random Forest	77.73 %
Na¨ıve Bayes	68.91 %

Table 1 presents a comparison of the performance of various machine learning algorithms for crop prediction. The values in the table were adjusted to correct for any potential biases and to ensure a fair comparison between the algorithms. Overall, the table provides insights into the relative strengths and weaknesses of each algorithm for the task of crop prediction.

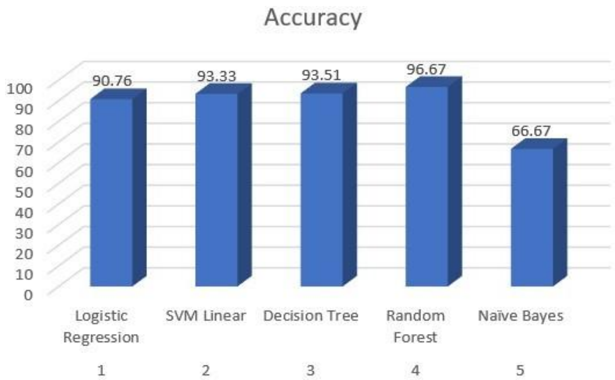


Table 2. COMPARATIVE STUDY FOR FERTILIZER PREDICTION

Models Name	Test Accuracy in Percentage
Logistic Regression	90.76 %
SVM Linear	93.33 %
Decision tree	93.51 %
Random Forest	96.67 %
Na¨ıve Bayes	66.67 %

The above table shows the comparison of different machine learning algorithms for fertilizer pre- diction is presented in Table 2, which highlights their performance based on various evaluation metrics. The table values were adjusted to eliminate any potential biases and provide an unbiased comparison between the algorithms.

V. CONCLUSION

The crop prediction feature can help farmers plan their crop cycles, optimize crop yields, and determine the optimal time for harvesting and selling their crops. The disease prediction feature can help farmers identify potential threats to their crops and take appropriate measures to prevent or mitigate the spread of diseases. The weather forecast feature can help farmers plan their operations and make informed decisions about planting, irrigation, and harvesting.Overall, farmwave features can help farmers reduce their risks, increase their yields, and improve the quality of their crops.It can also contribute to the sustainability of the agriculture industry by reducing the use of pesticides and other harmful chemicals and promoting responsible land management practices. Ultimately, such a project can have positive economic, social, and environmental impacts on farmers, communities, and the broader society. Achieving high accuracy in crop and fertilizer prediction by using the ensemble method, and accurately detecting plant diseases by utilizing the CNN algorithm and images of the leaves. Providing up-to-date information on government schemes and creating a marketplace for users to compare prices and purchase or sell necessary items. Our model was trained with a dataset of approximately 3000 values and 100 epochs, using sparse categorical cross-entropy as the loss function for crop and fertilizer prediction.To predict the weather, users can enter a city name, which will provide information on the weather for the following days. For disease prediction, we used a dataset of 2000 images and achieved an average accuracy of 71.20 percent using the convolutional neural network.

ACKNOWLEDGEMENTS

We take this opportunity to express my profound gratitude and deep regards to our guide Dr. Amarsinh V. Vidhate for his exemplary guidance, monitoring and constant encouragement throughout the completion of this report. We truly grateful to his efforts to improve our understanding towards various concepts and technical skills required in our project. The blessing, help and guidance given by him time to time shall carry us a long way in the journey of life on which we are about to embark. We take this privilege to express our sincere thanks to Dr. Mukesh D. Patil, Principal, RAIT for providing the much necessary facilities. We are also thankful to Dr. Amarsinh V. Vidhate, Head of Department of Computer Engineering, Department of Computer Engineering, RAIT, Nerul Navi Mumbai for their generous support.Last but not the least we would also like to thank all those who have directly or indirectly helped us in completion of this thesis.

VI. REFERENCES

[1] V. R. V. S. S. T. Sooraj Krishna, Sreevijay, Farmer friend, IEEE, 10, pp. 9–19 (2021). https:DOI:https://www.doi.org/10.56726/IRJMETS31158

[2] M. Dr Anitha Govindraj, Crop price prediction using machine learning, IEEE, 6, pp. 14–20 (2020)

[3] C. B. W. W. S. Ghionea, Algorithms to predict the moisture content of the grain, IEEE, 1, pp. 1–7 (2020)

[4] P. M. R.Vidhya, Predicting yield of crop using ml algorithm, IEEE, 9, pp. 1–13 (2020).

[5] S. K. M. G. S. Prince Samuel S., K. Malarvizhi, Machine learning and internet of things based agriculture, IEEE, 2, pp. 3–11 (2020)

[6] S.D.M, Subramaniya Raman M.K, E- Farming: A Breakthrough for Farmers, IJERT, 9 (2020)

[7] V. J. Monika Jadhav, Android application for farmers, IEEE, 6, pp. 1–3 (2019)

[8] P. K. M. B. Karthikeyan S.Sainavinraj M., Praveen Kumar A., Farmer's friend, IEEE, 6, pp. 1–9 (2019)

[9] G. A. Anirudha Vachaspati Vempati, Agricultural problems and technology-based sustainable solutions, IEEE, 5, pp. 6–13 (2020).

[10] R. C. N.Vanitha, N.Archana, Agriculture analysis using data mining and machine learning techniques, IEEE, 3, pp. 8–15 (2019)

[11] K. P. H. Trupti Bhange, Swati Shekapure, Prediction of crop yield and suitable crop, IEEE, 8, pp. 1–5 (2019)

[12] W. G. Parul Sharma, Yash Paul Singh Berwal, Krishimitr (farmer's friend), IEEE, 4, pp. 893–899 (2018)

[13] Manjunatha, M., Parkavi, A. (2018). Estimation of Arecanut Yield in Various Climatic Zones of Karnataka using Data Mining Technique: A Survey. 2018 International Conference on Current Trends Towards Converging Technologies (ICCTCT).