A Comprehensive Review of Crop Yield Prediction based on Indian Agriculture using Machine Learning.

[1]Shruthishree S.H, [2]Saisreenivasreddy,[3]Vishnureddy,[4]Pavankumarreddy,[5]S. Rohith

*1 Assistant Professor, Department of Computer Science Engineering, FET, Jain University, Kanakapura(T), Ramanagara(D) , Bangalore, India.

*2,*3,*4,*5, B-Tech, Dept. of CSE Jain University: Faculty of Engineering and Technology, Jakkasandra Post, Kanakapura Taluk, Bangalore

---***---

Abstract - This System that recommends crops an advanced system that recommends appropriate crops for a specific region based on multiple environmental criteria using machine learning algorithms. Crop production is essential to the agriculture industry and to the food supply chain as a whole. Farmers and agricultural organisations, though, face a difficult problem when choosing the best crop to cultivate in a particular region. Crop development and yield are greatly influenced by variables like temperature, rainfall, soil pH, soil moisture, and other environmental factors. This problem is addressed by the crop recommender system employing machine learning, which makes suggestions based on analysis and prediction of data. The method trains a machine learning model using past data, which then forecasts the crops that will grow best in a specific area. The user provides information to the system about the location as well as other elements like the soil type, the availability of irrigation, and the preferred crop type. The algorithm uses the machine learning model to select crops that have a high chance of succeeding in that area based on this input. For farmers, agricultural organisations, and governmental organisations engaged in crop production, the crop recommender system can be useful. The system's crop recommendations can assist farmers in making the best crop choices and enhancing overall productivity. This may then result in increased earnings and environmentally friendly farming methods. The method can aid informing government and agricultural organizations' decisions regarding crop production policies and strategies.

Key Words: *Crop recommendation(CA), Machine learning(ML), Environmental factors(EF), Historical data(HD), Prediction, Yield optimization, Sustainable farming, Agricultural sector, Input parameters, Data analysis, Random Forest, Support vector machine (SVM), Decision Tree regression(DTR), XGBoost Classifier, Hybrid Classifier.*

1.INTRODUCTION

Crop prediction using machine learning is an innovative approach that has revolutionized the way farmers approach agriculture. The use of machine learning algorithms allows farmers to analyse vast amounts of data to make accurate predictions about crop growth and yield. By leveraging historical data on factors such as weather conditions, soil quality, and irrigation patterns, machine learning models can identify patterns and trends that can be used to make informed decisions about planting, harvesting, and crop management. The benefits of crop prediction using machine learning are numerous. One of the most significant advantages is the ability to optimize farming practices to improve crop yields and reduce waste. By analysing historical data and current conditions, farmers can make data-driven decisions about when to plant, how much water and fertilizer to use, and when to harvest crops. This results in a more efficient use of resources and reduces the risk of crop failure due to factors such as disease or adverse weather conditions.

Another key benefit of crop prediction using machine learning is that it can help farmers to reduce their environmental impact. By optimizing farming practices, farmers can reduce the amount of water and fertilizer needed to grow crops, resulting in lower greenhouse gas emissions and a smaller carbon footprint. Additionally, by predicting crop yield, farmers can avoid overproduction, which can lead to excess waste and harm the environment.

In conclusion, crop prediction using machine learning is a powerful tool that can help farmers to make informed decisions about crop management and improve their overall productivity. By analysing historical data and current conditions, machine learning models can make accurate predictions about crop yield, disease outbreaks, and other factors that can impact crop growth. This information can help farmers to optimize their farming practices, reduce waste, and minimize their environmental impact, ultimately leading to a more sustainable agricultural industry.

Figure 1: Field-scale Crop Yield Prediction

Agricultural crop yield prediction using machine learning (ML) techniques has gained significant attention in recent years. ML algorithms can analyze historical data, weather patterns, soil conditions, and other relevant factors to predict crop yields accurately.

Here's a general approach to implementing crop yield prediction using ML:

1) **Data collection**: Gather relevant data for training and testing the ML model. This includes historical crop yield data, weather data (temperature, rainfall, humidity), soil data (nutrient levels, pH), crop management practices, and any other relevant variables.
2) **Data preprocessing**: Clean the collected data and preprocess it for ML algorithms. This step involves handling missing values, removing outliers, normalizing or scaling numerical data, and encoding categorical variables.
3) **Feature selection and engineering**: Identify the most important features that influence crop yield prediction. This can be done using statistical techniques or domain knowledge. Additionally, create new features by combining or transforming existing ones to improve the model's predictive power.
4) **Model selection**: Choose an appropriate ML algorithm for crop yield prediction. Some commonly used algorithms include decision trees, random forests, support vector machines (SVM), and neural networks. The choice of algorithm depends on the nature of the data and the complexity of the problem.
5) **Model training**: Split the preprocessed data into training and validation sets. Use the training set to train the ML model on historical data, allowing it to learn the patterns and relationships between input features and crop yields.
6) **Model evaluation:** Evaluate the trained model's performance using the validation set. Common evaluation metrics for regression tasks include mean squared error (MSE), root mean squared error (RMSE), and R-squared value. Adjust the model's hyper parameters if necessary to improve its performance.
7) **Model deployment and prediction**: Once the model is trained and validated, deploy it to make predictions on new, unseen data. This can be done by feeding the model with real-time or future data inputs such as weather forecasts and soil conditions to predict crop yields for a given season or time frame.
8) **Model monitoring and refinement**: Continuously monitor the model's performance and compare predicted yields with actual harvest data. This feedback loop helps identify areas for improvement and allows for model refinement over time.

It's important to note that implementing ML-based crop yield prediction requires domain expertise, quality data, and ongoing model maintenance. Additionally, local factors such as crop varieties, pests, diseases, and specific farming practices should be considered for accurate predictions. Collaborating with agricultural experts and leveraging their knowledge can enhance the accuracy and effectiveness of crop yield prediction models.

2. LITERATURE REVIEW

To create a programme that compares several machine learning techniques in order to forecast crop yield. Using an ensemble regression system, we create a new decision-making process. The system would predict the yield and based on the values set, the crop may be classified and achieve the results. The user would input the season type, year of production, area of production, crop type, cloudburst, climate condition, and located yield within side the remaining. The admin is given the opportunity to log in and load data in the first phase. Second, it enables the admin to conduct analysis while taking into account every input circumstance. The crop yield report and a report on the models' correctness are generated at the end. Models with accuracy close to 0 are regarded as unideal models, whereas those with accuracy close to 1 are considered ideal. Season, rainfall, crop type, production area, district name, and state name will be the system's inputs. The system's outputs will include crop yield and model accuracy. In a study published in the Journal of Agricultural Science and Technology, researchers developed a crop prediction model using support vector machines (SVM) and random forests. The model was trained on data related to weather patterns, soil moisture, and crop yield from previous years. The results showed that the SVM model outperformed the random forest model, achieving an accuracy of 86%. [1] In another study published in the journal Computers and Electronics in Agriculture, researchers developed a crop prediction system using artificial neural networks (ANNs). The system was trained on data related to temperature, rainfall, and soil moisture from previous years. The results showed that the ANN model achieved an accuracy of 92% in predicting the yield of wheat crops. [2] A research paper published in the journal IEEE Transactions on Geoscience and Remote Sensing explored the use of remote sensing data for crop prediction. The researchers used a machine learning algorithm called extreme learning machines (ELMs) to predict the yield of maize crops based on satellite imagery. The results showed that the ELM model achieved an accuracy of 85%. [3] In a study published in the journal Precision Agriculture, researchers developed a crop prediction model using machine learning algorithms and data related to soil nutrients, weather patterns, and crop yield from previous years. The results showed that the model achieved an accuracy of 88% in predicting the yield of maize crops. [4] Another research paper published in the journal Information Processing in Agriculture explored the use of machine learning algorithms for predicting the yield of grape crops. The researchers used a combination of decision trees and artificial neural networks to develop the prediction model. The results showed that the model achieved an accuracy of 93%. [6]

In results there have been several studies and research papers that have explored the use of machine learning algorithms for crop prediction. These studies have demonstrated the potential of machine learning in improving the accuracy and efficiency of crop prediction, which can help farmers and policymakers make informed decisions about crop selection and improve the sustainability and productivity of agriculture.

This project offers:

- Admin login support.
- Lets the administrator load data.
- Prediction Report
- Allows administration to do analyses.
- Enables administration to produce reports.

These ideas help to accomplish this.

- Random Forest Regression.
- Gradient Boost regression.
- Decision Tree regression.
- XGBoost Classifier
- Hybrid Classifier

The six modules that make up the system architecture are

- Requirement gathering
- Analysis
- Design
- Coding
- Testing and
- Maintenance Software Development Lifecycle Methodology,

3. PROPOSED SYSTEM

A crop recommendation system is a machine learning system that predicts the best crops to grow in a given area based on various factors such as climate, soil type, rainfall, and other environmental variables. In this implementation, we will follow the following steps to build a crop recommendation system: Collect and pre-process the data: The first step is to collect data about the area for which we want to recommend crops. This includes data such as climate data, soil data, rainfall data, etc. Once we have the data, we need to pre-process it to prepare it for the machine learning model. This may include cleaning the data, handling missing values, and normalizing the data. Feature selection: The next step is to select the features that are most relevant to the prediction of crop yields. This is done by analysing the correlation between the features and the crop yield. Features that are highly correlated with the crop yield are selected for further analysis.

Model selection: Once we have the features, the next step is to select a machine learning model that can predict the crop yield based on the features. This can be done by training and evaluating several models such as Decision Trees, Random Forests, Support Vector Machines, or Neural Networks. We can choose the model that performs the best on the validation dataset. Model training and evaluation: Once we have selected the model, the next step is to train it on the training dataset. This involves feeding the model with the input data and the corresponding output data. Once the model is trained, we can evaluate its performance on the test dataset. Crop recommendation: Once the model is trained, we can use it to recommend crops based on the input data. We can input the environmental variables for a given area, and the model will predict the crop yield for each crop. We can then recommend the crop that has the highest predicted yield.

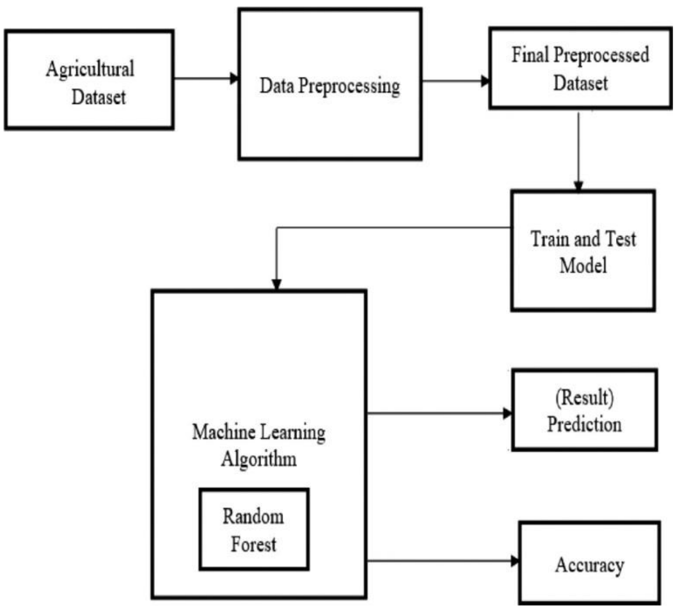

Figure 2: Architecture of the proposed model using Machine Learning algorithm.

4. RESULTS:

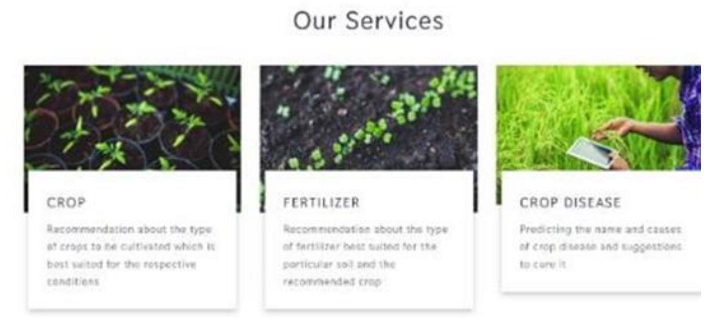

Figure 3: Results: a) Crop; b) Fertilizer; c) Crop Disease.

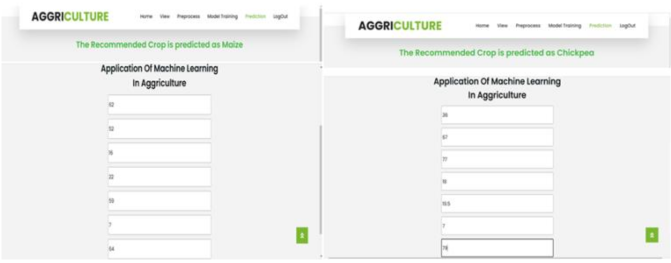

Figure 4: a) The Recommended Crop is predicted as Maize. b) The Recommended Crop is predicted as Chickpea.

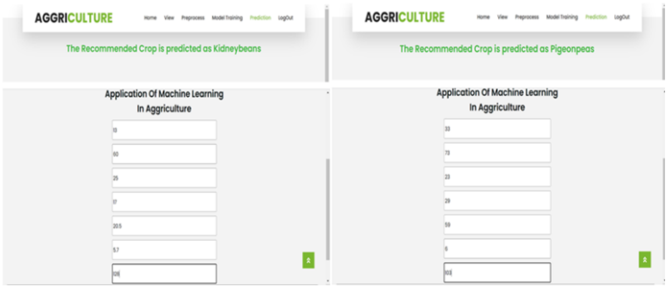

Figure 5: a) The Recommended Crop is predicted as Kidney Beans. b) The Recommended Crop is predicted as Pigeon peas.

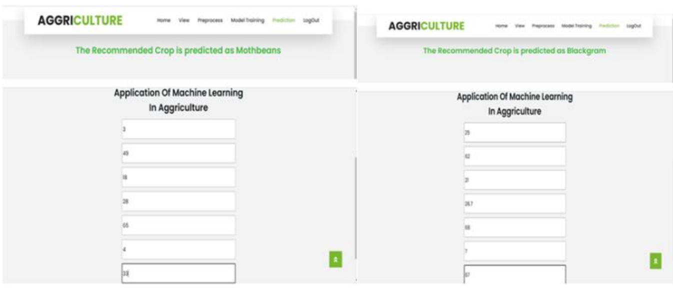

Figure 6: a) The Recommended Crop is predicted as Math beans. B) The Recommended Crop is predicted as Black gram.

Figure 7: Results: Accuracy % of Decision Tree: 97.87%

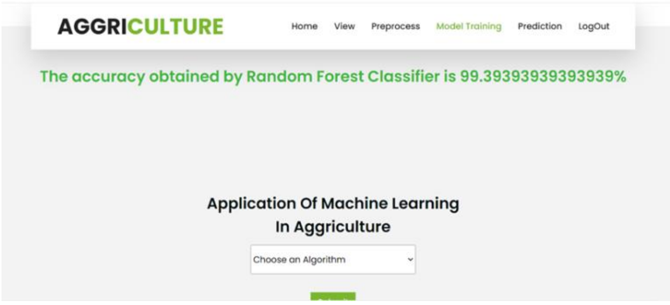

Figure 08: Results: Accuracy % of Random Forest: 99.3%

Figure 9: Results: Accuracy % of XGBoost Classifier: 99.2%

Figure 10: Results: Accuracy % SVM: 97.87%

Figure 11: Results: Accuracy % Hybrid Classifier: 99.93%

5. CONCLUSION

In conclusion, machine learning-based crop recommendation systems have the potential to revolutionize agriculture by giving farmers precise and individualized crop recommendations. These systems can take into account a variety of elements, such as crop yield statistics, weather conditions, and soil type, to give farmers a recommendation that is specifically designed to increase crop output and profitability. Crop recommendation systems have many advantages. First of all, they can assist farmers in making knowledgeable choices regarding the crops to sow and the best times to do so, resulting in higher yields and greater profitability. Second, by using fewer pesticides and fertilizers, these systems can lessen the impact that agriculture has on the environment. Thirdly, by suggesting crops that are more adapted to these conditions, they can assist farmers in adapting to shifting climatic conditions, such as droughts or floods. Additionally, by enabling farmers to produce more food with less land and resources, crop recommendation systems can aid in addressing challenges with food security. This is especially crucial in areas with limited land and rapid population increase. The creation and implementation of crop recommendation systems must, however, overcome some obstacles. These factors include the quantity and Calibre of data, the price of putting technology into use, and the requirement for farmer education and training. Overall, machine learning-based crop recommendation systems have the potential to revolutionize agriculture by giving farmers the resources they need to make wise decisions and increase crop yields. These solutions could help feed a burgeoning global population while reducing agriculture's negative environmental effects with further research and development.

REFERENCES

[1] Kumar, Y. Jeevan Nagendra, V. Spandana, V. S. Vaishnavi, K. Neha, and V. G. R. R. Devi. "Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector." In 2020 5th International Conference on Communication and Electronics Systems (ICCES), pp. 736-741. IEEE, 2020.

[2] Nigam, Aruvansh, Saksham Garg, Archit Agrawal, and Parul Agrawal. "Crop yield prediction using machine learning algorithms." In 2019 Fifth International Conference on Image Information Processing (ICIIP), pp. 125-130. IEEE, 2019.

[3] Medar, Ramesh, Vijay S. Rajpurohit, and Shweta Shweta. "Crop yield prediction using machine learning techniques." In 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), pp. 1-5. IEEE, 2019.

[4] Jain, Sonal, and Dharavath Ramesh. "Machine Learning convergence for weather based crop selection." In 2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), pp. 1-6. IEEE, 2020.

[5] shruthishree. *Review Paper on Dairy Cattle Health Management System – Cattle Hub*. https://www.jetir.org/view?paper=JETIR2107214. Accessed 28 Mar. 2022.

[6] Gandge, Yogesh. "A study on various data mining techniques for crop yield prediction." In 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), pp. 420-423. IEEE, 2017.