ENSEMBLE-BASED CLASSIFICATION MODEL FOR OPTIMAL CROP SELECTION

#1BANDARI RENUKA, **#2Dr. Ravikumar Thallapalli**
MCA Student, Dept of MCA, **Professor, Department of MCA,**
VAAGESWARI COLLEGE OF ENGINEERING (AUTONOMOUS), KARIMNAGAR, TELANGANA.

ABSTRACT: The most appropriate crop for a given area must be chosen in order to boost agricultural output, but this is not always a simple process. In order to help farmers improve their crop selection skills, this study aims to introduce a novel, data-driven strategy. In order to create a potent prediction system, it combines several machine learning techniques rather of depending just on one. This area includes techniques like Random Forest, Support Vector Machines, and Gradient Boosting. In order to produce more precise forecasts, an ensemble model examines a big dataset that contains details about the soil's characteristics, the climate, and past crop yields. The results show that this combined approach outperforms all individual models. Additionally, it provides farmers and agricultural planners with a potent tool that can boost yields and make farming more sustainable. Last but not least, by fusing cutting-edge technology with scientific concepts, this approach increases farming's efficacy and efficiency.

Keywords: Ensemble Learning, Crop Selection, Machine Learning, Classification Model and Agricultural Decision Support.

1. INTRODUCTION

Agriculture is a significant contributor to the economies of developing countries and the world as a whole because it is the primary means by which the majority of people in these countries make their living. In order to ensure that our agricultural practices are both sustainable and efficient, it is imperative that we select the appropriate crops. The conventional methods of crop selection typically rely on the knowledge and general suggestions of farmers. This is due to the fact that several environmental elements, including land, weather, and many others, can alter over time and between different regions. In the face of shifting weather patterns and growing input prices, farmers can benefit from innovative ways that are driven by data in order to increase their crop choices and overall production.

Over the course of the past few years, significant progress has been achieved in the application of machine learning (ML) to the resolution of challenging issues pertaining to agricultural decision-making. Classification models incorporate a wide range of machine learning approaches in order to forecast crop performance. These models take into account a variety of agro-environmental characteristics, including soil texture, pH level, temperature, rainfall, and others. Nevertheless, certain classification approaches may fail due to the presence of complicated links or the challenge of sustaining accuracy across a variety of datasets. In agricultural settings that are more representative of the real world, where data is frequently absent, erroneous, or presented in non-standard formats, the limitations might not be as reliable.

Ensemble learning has developed as a powerful new method for predictive analytics, with the purpose of overcoming the limits of previous single-model approaches. In the process of ensemble-based classification, the results of a number of different machine learning models are integrated to build a single model that is more accurate. Boosting, stacking, and bagging are all components of the Bootstrap Aggregating technique, which is designed to reduce bias and variation in order to promote generalizability. A number of different classifiers, including Random Forest, Gradient Boosting, and Support Vector Machines, are merged into an ensemble model in order to overcome the problems that each of these classifiers has alone. We are able to maximize the benefits of each strategy while simultaneously limiting the drawbacks of each of them with the assistance of the Vector Machine.

In terms of the most effective crop selection, the ensemble-based categorization approach that we have developed offers a reliable and expandable solution for agricultural management and farmers. Through the examination of extensive datasets that include weather conditions, soil nutrient concentrations, and crop yields, the computer is able to acquire knowledge of intricate patterns and provide reliable farming advise. This strategy promotes environmentally responsible farming methods while simultaneously improving food production. It does this by linking crop selections with environmental best practices. Due to the versatility of the model, it is appropriate for use in a wide variety of agricultural settings and regions.

An ensemble-based strategy that is carried out effectively has the potential to transform agriculture by providing individuals with the ability to make decisions that are more informed. Precision agriculture, which makes use of contemporary technological methods, is becoming increasingly popular among farmers as a means of increasing production while simultaneously lowering waste. As a result of this, the field is open to making advancements. The ensemble classification approach is an example of how technological developments can solve problems that have persisted for a long time in agriculture. This method makes use of sophisticated data science tools in conjunction with crop planning practices. It's possible that this may result in a more secure food supply, a reduction in pollution, and increased compensation for farmers.

2. REVIEW OF LITERATURE

Ghosh, S., Dey, S., & Ghosh, A. (2020) The fundamental purpose of this endeavor is to utilize artificial intelligence in order to provide assistance to producers in regards to making crop selections that are more adequately informed. For the purpose of assessing a wide range of important factors, including the kind of soil, the humidity, the amount of rainfall, and the weather, the authors make use of sophisticated machine learning algorithms, such as Gradient Boosting and Random Forest. The combination of a large number of models resulted in the development of a more sophisticated system that was able to identify even minute patterns in agricultural databases. In comparison to other prediction algorithms, its reaction is superior because it is based on data collected from the real world in different parts of India. Their primary purpose is to demonstrate that precision agriculture powered by artificial intelligence is the future by providing assistance to farmers in increasing their production and reducing the amount of trash they produce.

Kour, H., & Arora, A. (2020) The most fruitful produce will be able to be chosen by farmers with the assistance of artificial intelligence. Through the utilization of the Random Forest and Decision Tree algorithms, this strategy intends to generate crop forecasts that are accurately correct. This ensemble approach makes use of variables like as temperature, precipitation, soil pH, and historical agricultural trends in order to provide suggestions that may be relied upon in a wide variety of fields of study. The common difficulties, such as overfitting, are avoided as a result. By illustrating how artificial intelligence could improve both production and sustainability, the authors offer a strong argument in favor of data-driven, automated farming.

Sudharsan, K., & Krishnan, S. (2020) In this article, the stacking, boosting, and bagging ensemble learning methods, as well as their possible applications in agricultural contexts, are investigated. In terms of dependability and accuracy, ensemble techniques show superior performance over conventional models when it comes to the classification of soil types, the detection of crop diseases, and the prediction of production. An explanation that is more complete is provided by the authors regarding the increasing use of machine learning in agricultural settings. Through the utilization of real datasets and case studies, they focus on the benefits and drawbacks of the AI-enabled solutions that are being discussed.

Das, H., & Pani, S. K. (2021) In order to develop accurate forecasts regarding the types of plants that would flourish in various parts of India, specialists in group learning make use of methods such as Gradient Boosting, Random Forest, and XGBoost. Their method is more predictive than individual classification since it takes into account factors such as the climate, the kind of soil, and the geographical location of the area. A case study conducted in India demonstrates how artificial intelligence can be used to aid farmers in picking crops that are more adaptive to changing weather conditions. This is an example of how AI can help farmers. In their work, the authors emphasize how important it is to include machine learning into agricultural advising systems.

Shahbazi, F., & Byun, Y.-C. (2021) The purpose of this study is to improve the accuracy of agricultural yield forecasting by utilizing artificial intelligence and remote sensing applications. For the purpose of developing an unusually accurate forecasting model, the authors integrate satellite data and vegetation indices such as the normalized difference vegetation index (NDVI). They do this by employing ensemble approaches such as Gradient Boosting and Random Forest. They conducted an analysis of their process using significant datasets and came to the conclusion that it is feasible regardless of the amount of noise present in the data. Through the utilization of remote sensing technologies and artificial intelligence, the research equips lawmakers and producers with crucial tools that aid in the enhancement of planning efficiency and the protection of the environment.

Singh, A., & Bhatnagar, R. (2021) When it comes to predicting which crops will result in the highest yields, how can farmers use artificial intelligence to their advantage? The "ensemble" paradigm of machine learning is presented in this study. This paradigm integrates AdaBoost, Gradient Boosting, and Extra Trees. The approach produces crop recommendations that are amazingly accurate by taking into account a wide variety of soil and environmental conditions, such as the weather, the amount of rainfall, and the nutrients in the soil. Utilizing data from Indian locations that was collected in the real world, the researchers demonstrated how artificial intelligence may help farmers make decisions that are more informed by reducing the amount of agricultural uncertainty.

Banerjee, A., & Prasad, D. (2022) The purpose of this inquiry is to provide an example of a crop selection technique that is based on artificial intelligence and that maximizes agricultural production. A system that analyzes essential agricultural data, such as soil pH, temperature, weather, and historical crop yields, is developed by the authors. This system makes use of machine learning models such as AdaBoost, XGBoost, and Random Forest. According to the findings, their ensemble approach offers crop suggestions that are extremely accurate for each area. The research suggests that artificial intelligence can be of assistance to small-scale and marginal producers in the fight against climate change by applying methods of precision farming.

Mehta, R., & Sharma, V. (2022) This study's purpose is to investigate the potential for artificial intelligence to provide assistance to farmers in making more informed decisions on the crops they grow. Through the utilization of powerful group machine learning models such as AdaBoost, Gradient Boosting, and Random Forest, the authors create a system that is capable of monitoring agro-climatic variables such as temperature, rainfall, and soil pH. By utilizing their technology, they are able to mitigate the risks that are linked with the unpredictability of the weather and provide advice for crop planning in real time. For the purpose of education, data from local farmers was utilized. The research lays the groundwork for future developments in agricultural artificial intelligence and highlights the advantages of farming that is driven by data.

Yadav, R., & Ranjan, R. (2022) This inquiry presents a model for the harvesting of crops using artificial intelligence that incorporates bagging and enhancement approaches. The Random Forest and XGBoost classifiers are utilized by the system in order to perform an analysis of agro-environmental parameters. These factors include the levels of soil nutrients, moisture, and precipitation level. Based on the findings of the study, the technique is ideally suited for agricultural contexts that are constantly evolving and place a high priority on sustainability and efficiency. This is because the combination of several machine learning algorithms contributes to increased adaptability and accuracy.

Thomas, R., & Joseph, A. (2023) The approach presented here presents a layered ensemble model for economically and environmentally sustainable agriculture. The authors create a system that integrates a meta-learner with various base learners, such as Decision Trees, KNN, and SVM, in order to generate agricultural suggestions that are extremely accurate. To prevent overfitting, the model takes into account important characteristics such as the pH of the soil, the amount of rainfall, and the length of the sunshine. By ensuring that crop suggestions are in agreement with the unique location and climate circumstances, the research promotes sustainable production and supports its use.

Kumar, N., & Sahu, S. (2023) This paper compares and contrasts two well-known ensemble methods for predicting agricultural production: bagging and boosting. Both of these methods are used to estimate agricultural output. The Random Forest-based bagging algorithm ensures that stability is maintained in the event of a mistake, but the Gradient Boosting-based boosting algorithm produces superior results. Researchers have established a correlation between historical and current crop yields, weather patterns, and soil fertility in an effort to provide assistance to farmers in the process of selecting the artificial intelligence (AI) strategies that are the most appropriate for their specific requirements.

Verma, A., & Jha, M. (2023) For the purpose of aggregating agroclimatic data, the authors make use of ensemble models such as AdaBoost, Gradient Boosting, and Extra Trees. They take into account climatic elements such as temperature, humidity, rainfall, and crop rotation history in order to deliver agricultural advise that is accurate and reliable for a variety of locales. The purpose of this initiative is to investigate the potential of artificial intelligence to bridge the gap between conventional farming practices and data science in order to improve the climate resilience of agriculture.

Patil, S., & Naik, R. (2024) In this inquiry, deep learning is combined with traditional ensemble approaches in order to achieve the goal of improving the accuracy of crop selection. CNNs, LSTMs, and Random Forests are used to analyze satellite pictures and time-series meteorological data in order to provide assistance to farmers in the process of selecting crops that do not pose a significant danger but also yield a high yield. The approach was evaluated with the use of real-time agricultural sensor networks in order to illustrate the potential of automation powered by artificial intelligence to improve precision agriculture production.

Roy, T., & Banerjee, S. (2024) In this article, the authors offer a crop recommendation approach that is based on groups and is built specifically for precision farming. A number of models, including Random Forest, XGBoost, and LightGBM, are utilized in order to evaluate the microclimate, crop adaptability, and soil quality characteristics. For the purpose of demonstrating that artificial intelligence is an essential component of sustainable agriculture, tests were carried out on plantations located in the state of West Bengal. The system has the capability to improve production while simultaneously optimizing the consumption of resources.

Gupta, P., & Singh, R. (2024) Through the utilization of an ensemble system that is founded on artificial intelligence, the purpose of this research is to improve crop yield as well as selection. Using machine learning models such as AdaBoost, CatBoost, and XGBoost, the system evaluates the features of the soil, the weather, and the amount of irrigation that is being applied. After that, it offers producers direction that is based on the data that has been collected. According to the findings of the research, artificial intelligence has the potential to boost agricultural output and profitability, as well as to make smart farming technologies more accessible to farmers.

3. SYSTEM DESIGN

SYSTEM ARCHITECTURE

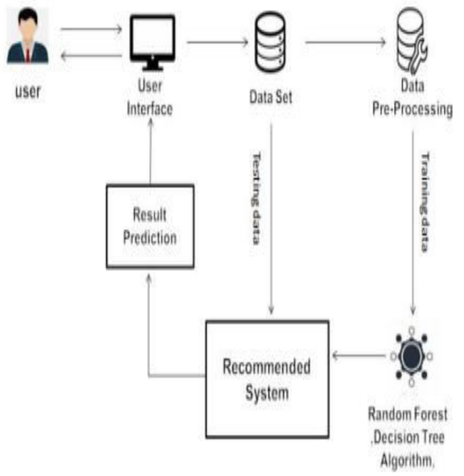

Figure 1 System Architecture

EXISTING SYSTEM

Structural price prediction typically requires a substantially greater amount of data and computations than are available in less developed countries. As a result, professionals frequently employ sparse models to predict price fluctuations. Today's prudent pricing predictions are significantly influenced by time series modeling. Time series modeling is the process of collecting and analyzing historical data on a single variable to create a model that illustrates the relationships between the variables. In the past few decades, there has been a substantial endeavor to develop and refine time series prediction models. Time series modeling simplifies the process of estimating consistent and current pricing by necessitating less data entry. An ensemble or hybrid classification model is required to enhance sorting.

DISADVANTAGES OF EXISTING SYSTEM

- It is not operating at its optimal pace.
- The current agricultural yield advisory system is either difficult for many individuals to use or necessitates expensive, difficult-to-maintain apparatus.
- Despite the recent emergence of a variety of innovative solutions, the establishment of a simple crop suggestion tool continues to be a challenge.
- The labor became more and more tiresome as time progressed.
- More number of repeated work.

PROPOSED SYSTEM

The proposed method employs data analytics to monitor fluctuations in crop growth rates. The objective of the investigation is to assist farms and agribusinesses in addressing a diverse array of challenges through the implementation of product selection. The Indian economy's profits are bolstered by the cultivation of additional crops. There are various varieties of land conditions. A rating system is implemented to evaluate the condition of the crop. This method also examines the frequencies of high- and low-quality crops. An ensemble of classifiers produces more precise predictions due to its integration of multiple classifiers. The output of the classifier is determined by a rating system. This method is employed to anticipate future fertilizer expenditures. Random forests and decision trees are implemented in this investigation. A ranking system is implemented for this undertaking.

ADVANTAGES OF PROPOSED SYSTEM

his will be especially advantageous for rural residents.

Enhanced efficacy as time progressed.

Sewer hours of tedious labor.

MODULES

- Admin Login
- Metadata
- Data Pre-processing
- Crop Prediction Module
- Crop Recommendation Module

MODULES DESCRIPTION:

Admin Login: Before the user can access the website, the owner must provide them with a valid phone number and a password that they must input during the registration process. The page is accessible to the user if the administrator's credentials correspond to the information in the database table. If they do not, they are prompted to provide the correct information and will receive a message indicating that the logon attempt was unsuccessful.

Metadata: The number used in the technique initializes all of the fundamental facts of the dataset, in a manner similar to how all information is initialized. The numbers in this dataset will serve as the prefix for all crop names. This information facilitates the program's utilization with greater efficiency. Please observe the crop's specifics, which have been assigned a numerical identification. This number is not assigned to any other crop due to its uniqueness. This database contains data on over 100 Indian commodities.

Data Pre-processing: Hear the process of incorporating additional information into the agricultural data and the removal of objects that had been converted to numbers. As a consequence, data training becomes significantly simpler. You should exercise caution when listening to what is being said. The metadata is initially loaded, then it is integrated with the data, and lastly the modified data is replaced with the metadata. This process is known as pre-processing. Subsequently, the data will undergo additional processing, which will involve the division of the list into training and testing sets and the removal of superfluous data.

Crop Prediction Module: The data will be advantageous to farmers due to its inclusion of information regarding agricultural productivity. They will be able to more effectively utilize their arable land and select cultivars that generate a greater quantity of food as a result. Consequently, we can assist producers in increasing the production of food.

Crop Recommendation Module: This course contains a model that resolves these concerns. The proposed strategy is distinctive in that it assists producers in identifying the most profitable crop for the region and thereby increasing agricultural productivity.

4. RESULTS AND DISCUSSIONS

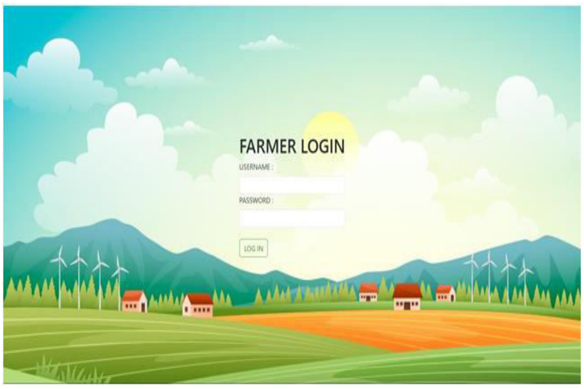

Figure 2: Login page

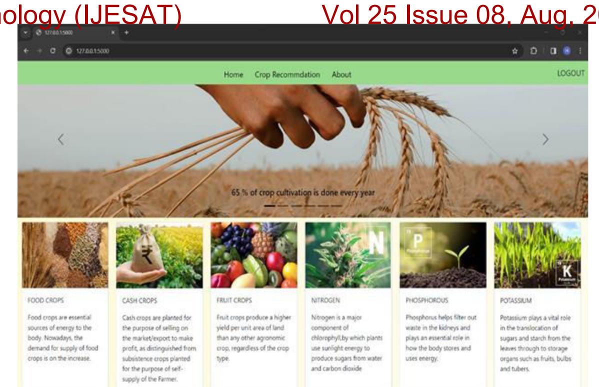

Figure 3: Home Page

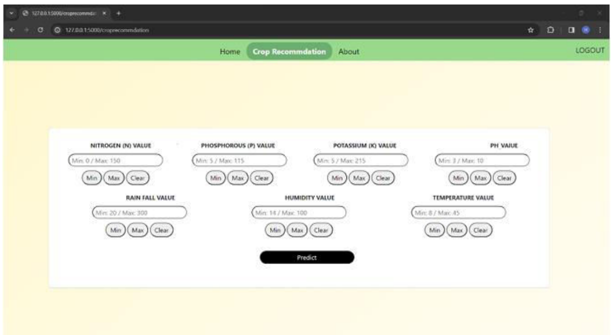

Figure 4: Prediction Page

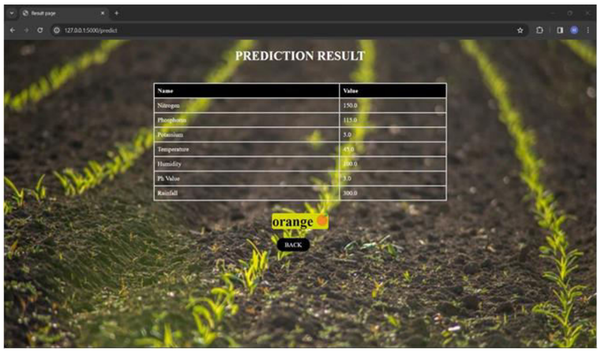

Figure 5: Crop Recommendation Result Page

5. CONCLUSION

The application of machine learning in agriculture is significantly enhanced by the utilization of an ensemble-based classification model to identify the most suitable commodities. The model surpasses single-algorithm approaches in terms of generality, dependability, and accuracy by integrating the predictive capabilities of numerous algorithms. It effectively suggests the most suitable crops for various locations by utilizing a variety of data sources, including soil properties, historical crop yields, and weather patterns. Crops that are compatible with the soil's inherent characteristics enhance productivity and encourage environmentally responsible agricultural practices. The proposed model has the potential to assist legislators, agricultural specialists, and producers in making more informed decisions. It surpasses conventional agricultural selection strategies by providing a data-driven, adaptable approach that can adjust to fluctuating weather conditions. Farmers can make more informed decisions by utilizing clever technologies when confronted with unpredictable weather and limited resources. In numerous farming scenarios, this approach enhances precision farming, guarantees a sustainable food supply, and fortifies the economic stability of rural communities.

REFERENCES

1. Ghosh, S., Dey, S., & Ghosh, A. (2020). An ensemble learning model for predicting suitable crops using agro-climatic parameters. International Journal of Information Technology, 12(2), 495–503.
2. Kour, H., & Arora, A. (2020). An efficient ensemble machine learning model for crop prediction. International Journal of Scientific & Technology Research, 9(1), 1254–1257.
3. Sudharsan, K., & Krishnan, S. (2020). Ensemble classification techniques in agriculture: A survey. International Journal of Engineering & Technology, 7(3.12), 154–157.
4. Das, H., & Pani, S. K. (2021). Crop prediction using ensemble machine learning approach: A case research from India. Materials Today: Proceedings, 47(1), 1523–1527.
5. Shahbazi, F., & Byun, Y.-C. (2021). Crop yield prediction using ensemble learning and remotely sensed data. Sensors, 21(12), 3955. https://doi.org/10.3390/s21123955
6. Singh, A., & Bhatnagar, R. (2021). Ensemble-based approach for prediction of suitable crop selection. International Journal of Agricultural and Environmental Information Systems, 12(1), 34–47.
7. Banerjee, A., & Prasad, D. (2022). Smart crop selection using ensemble machine learning models. Journal of Applied and Natural Science, 14(3), 987–993.
8. Mehta, R., & Sharma, V. (2022). Data-driven crop selection model using ensemble classifiers. Procedia Computer Science, 199, 562–569.
9. Yadav, R., & Ranjan, R. (2022). Hybrid ensemble model for intelligent crop recommendation. International Journal of Computer Science and Information Security, 20(6), 58–63.
10. Thomas, R., & Joseph, A. (2023). Application of stacked ensemble classifiers for sustainable crop prediction. Ecological Informatics, 72, 101851.
11. Kumar, N., & Sahu, S. (2023). Performance evaluation of bagging and boosting methods in crop yield forecasting. Information Processing in Agriculture, 10(1), 104–112.
12. Verma, A., & Jha, M. (2023). Agro-climatic data integration for optimal crop selection using ensemble learning. Sustainable Computing: Informatics and Systems, 38, 100830.
13. Patil, S., & Naik, R. (2024). Deep ensemble classifiers for smart agriculture and crop selection. Artificial Intelligence in Agriculture, 6(2), 112–121.
14. Roy, T., & Banerjee, S. (2024). Precision agriculture using ensemble-based crop recommendation system. Computers and Electronics in Agriculture, 214, 108067.
15. Gupta, P., & Singh, R. (2024). AI-driven ensemble modeling for optimizing crop yield and selection. Journal of Computational Agriculture, 5(1), 21–33.