**An Approach for Crop Prediction in Agriculture: Integrating Genetic Algorithms and Machine Learning**

**ABSTRACT: -**

The agricultural sector in many South Asian countries, including Bangladesh and India, plays a pivotal role in the economy, with a significant portion of the population relying on it for their livelihood. However, farmers face challenges like unpredictable weather, soil variability, and natural disasters such as floods and erosion, leading to crop losses and financial difficulties. This often results in a decline in interest in agriculture despite government support. Our study focuses on predicting the classification of various crops, such as rice, jute, and maize, using a combination of soil and weather features. The predictive model leverages soil parameters like Nitrogen, Phosphorus, Potassium, and pH levels, alongside weather variables such as Temperature, Humidity, and Rainfall. We propose a hybrid approach that integrates machine learning with genetic algorithms, where a Random Forest Classifier is used for crop classification across 22 different crop types. The Genetic Algorithm is utilized to optimize hyperparameters, enhancing model performance and robustness. Additionally, we applied Explainable AI (XAI) techniques, including Local Interpretable Model-agnostic Explanations (LIME) and SHapley Additive exPlanations (SHAP), to interpret and validate the model’s predictions. By improving feature selection and model parameters, our approach addresses limitations associated with existing models, providing more reliable and accurate predictions. This system has the potential to reduce crop losses, improve agricultural productivity, and contribute to the sustainability and prosperity of the agricultural sector.

**INTRODUCTION**

1

1.Introduction :

Agriculture is the backbone of many South Asian countries, including Bangladesh and India, where a significant portion of the population depends on farming for their livelihoods. However, farmers in these regions face numerous challenges, such as unpredictable weather conditions, soil variability, and natural disasters like floods and erosion, which often result in crop losses and financial instability. Despite government support and subsidies, the agricultural sector continues to struggle, leading to a decline in interest among younger generations. In response to these challenges, accurate crop prediction has emerged as a critical area of research. By predicting the right crops to plant based on environmental conditions, farmers can make better decisions, mitigate losses, and improve productivity. This study aims to develop an advanced crop prediction model using machine learning techniques, integrating soil features like Nitrogen, Phosphorus, Potassium, and pH levels with weather variables such as Temperature, Humidity, and Rainfall. The proposed model leverages a **Random Forest Classifier** to classify 22 different crop types, enhancing the prediction process by utilizing **Genetic Algorithms (GA)** for hyperparameter optimization. Additionally, the model incorporates **Explainable AI (XAI)** techniques, such as **LIME** (Local Interpretable Model-agnostic Explanations) and **SHAP** (SHapley Additive exPlanations), to improve the transparency and interpretability of the predictions. These methods help in explaining the influence of individual features on the crop prediction, making the model more reliable and accessible for agricultural stakeholders. The integration of Genetic Algorithms with Random Forest enhances the performance of the model by optimizing its hyperparameters, ensuring more accurate and robust predictions. The combination of these advanced techniques promises to not only improve the accuracy of crop classification but also contribute to the sustainability and prosperity of the agricultural sector, offering a valuable tool to mitigate the impact of unpredictable environmental factors on crop yield.

**SCOPE OF THE PROJECT**

The scope of this project is focused on developing an advanced crop prediction model to support agricultural decision-making in South Asian countries, where the sector plays a vital role in the economy and the livelihood of millions. The project aims to integrate machine learning techniques with environmental data—specifically, soil features (e.g., Nitrogen, Phosphorus, Potassium, pH levels) and weather conditions (e.g., Temperature, Humidity, Rainfall)—to accurately predict the most suitable crops to plant in specific regions and conditions. The proposed system uses a **Random Forest Classifier** for crop classification and optimizes its performance through **Genetic Algorithms** (GA), which fine-tune the model’s hyperparameters for improved accuracy and efficiency.

**OBJECTIVE**

The objective of this project is to develop an advanced crop prediction system that can assist farmers in making informed decisions by accurately predicting the most suitable crops to plant based on environmental factors such as soil properties (e.g., Nitrogen, Phosphorus, Potassium, and pH levels) and weather conditions (e.g., Temperature, Humidity, and Rainfall). The model will employ a **Random Forest Classifier** for classifying 22 different crop types and will utilize **Genetic Algorithms** to optimize its hyperparameters, enhancing performance and accuracy. Additionally, the project aims to improve the model’s interpretability by integrating **Explainable AI (XAI)** techniques, such as **LIME** and **SHAP**, which will enable farmers to understand the factors influencing the model’s predictions. By providing transparent and reliable insights, the system will empower farmers to reduce crop losses, mitigate risks from unpredictable weather, and ultimately increase agricultural productivity. The model will be designed to be scalable, adaptable to changing agricultural conditions, and applicable across different geographical regions with varying climates and soil types. In doing so, the project seeks to contribute to sustainable farming practices, ensuring the long-term viability and profitability of the agricultural sector, thereby supporting the livelihoods of farmers and enhancing food security.

**EXISTING SYSTEM:**

In the realm of agricultural research, numerous existing algorithms have been applied to address crop classification and prediction challenges, especially in regions like South Asia, where agriculture significantly impacts the economy and livelihoods. Traditional approaches such as Decision Tree Classifier, Support Vector Machine (SVM), Naive Bayes, and K-Nearest Neighbors (KNN) have been widely used to predict crop types based on various features like soil composition, weather conditions, and other environmental factors. These algorithms, while effective to some extent, often face limitations in terms of scalability, accuracy, and interpretability, particularly when dealing with large, complex datasets or diverse crop types. For instance, models like Decision Trees can easily overfit, while SVMs may struggle with multi-class classification and require intensive computational resources. Additionally, these traditional algorithms may not be flexible enough to adapt to varying soil and climatic conditions, which are crucial in predicting crop yields in diverse agricultural landscapes. The proposed algorithm leverages a hybrid methodology where the Genetic Algorithm is employed to fine-tune the hyperparameters of the Random Forest model, enhancing its performance in classifying 22 different types of crops such as rice, jute, and maize. By optimizing key parameters like the number of trees, depth, and feature selection, the integration of GAs ensures that the model is not only more accurate but also more adaptable to the variability in soil and weather conditions. Furthermore, to ensure the model’s predictions are transparent and interpretable, we incorporated Explainable AI (XAI) techniques, specifically Local Interpretable Model-agnostic Explanations (LIME) and SHapley Additive exPlanations (SHAP).

**EXISTINGSYSTEM DISADVANTAGES:**

* Lack of Interpretability
* Overfitting and Model Complexity
* Scalability and Computation Costs
* Feature Dependency and Sensitivity
* Difficulty in Handling Missing Data
* Limited Adaptability to Changing Conditions Problems.

**LITERATURE SURVEY**

**Title:** Explainable Machine Learning for Crop Recommendation from Agriculture Sensor Data-a New Paradigm

**Author:** Samiran Das and Sujoy Chatterjee

**Year:** 2023

**Description:** The dwindling agricultural earnings and decrease in crop yield in recent years due to improper crop selection and fluctuation/ uncertainty in weather necessitate proper machine learning-based analysis. Machine learning methods can potentially alleviate the predicament caused by the lack of appropriate soil testing, consultation, and bias in manual suggestion. This work attempted to comprehend the agricultural sensor data and weather conditions and formulated the task in terms of supervised classification. The work obtained accurate suggestions in the presence of missing data, noise, etc. by using advanced machine learning methods. But recommendation alone is insufficient to convince farmers and other stakeholders to adopt this approach. Hence, this paper introduced explainable machine learning to completely comprehend the decision-making process. This work quantified the importance of features, explained individual prediction outcomes, and uncovered the rationale for decisions. The work employed state-of-the-art local interpretable model-agnostic, post-hoc explanation methods to provide in-depth insights. The insights obtained from the explanations can help the farmers develop a knowledge base and assist the farmers in choosing the appropriate sensors for the task. The human interpretable analysis enables the farmers to obtain satisfactory yields in these ever-changing and extreme weather conditions and environmental degradation.

**Title:** A Cloud Enabled Crop Recommendation Platform for Machine Learning-Driven Precision Farming.

**Author:** Navod Neranjan Thilakarathne, Muhammad Saifullah Abu Bakar, Pg Emerolylariffion Abas

**Year:** 2022.

**Description**: Modern agriculture incorporated a portfolio of technologies to meet the current demand for agricultural food production, in terms of both quality and quantity. In this technology-driven farming era, this portfolio of technologies has aided farmers to overcome many of the challenges associated with their farming activities by enabling precise and timely decision making on the basis of data that are observed and subsequently converged. In this regard, Artificial Intelligence (AI) holds a key place, whereby it can assist key stakeholders in making precise decisions regarding the conditions on their farms. Machine Learning (ML), which is a branch of AI, enables systems to learn and improve from their experience without explicitly being programmed, by imitating intelligent behavior in solving tasks in a manner that requires low computational power. For the time being, ML is involved in a variety of aspects of farming, assisting ranchers in making smarter decisions on the basis of the observed data. In this study, we provide an overview of AI-driven precision farming/agriculture with related work and then propose a novel cloud-based ML-powered crop recommendation platform to assist farmers in deciding which crops need to be harvested based on a variety of known parameters. Moreover, in this paper, we compare five predictive ML algorithms—K-Nearest Neighbors (KNN), Decision Tree (DT), Random Forest (RF), Extreme Gradient Boosting (XGBoost) and Support Vector Machine (SVM)—to identify the best-performing ML algorithm on which to build our recommendation platform as a cloud-based service with the intention of offering precision farming solutions that are free and open source, as will lead to the growth and adoption of precision farming solutions in the long run.

**Title:** Enhancing Crop Management: Ensemble Machine Learning for Real-Time Crop Recommendation System from Sensor Data.

**Author:** Nuzhat Prova, Sadia Hossain, Md Rezwane Sadik, Abdullah AI Maruf

**Year:** 2024.

**Description:** The agricultural industry is essential to the world’s food production, and it is critical to use cutting-edge technologies to increase crop productivity. We provide a revolutionary Crop Recommendation System (CRS) that utilizes cutting-edge technology to maximize crop output in response to the pressing need for improvement. Our study incorporates real-time monitoring of soil conditions, made possible by a custom hardware configuration that includes sensors for temperature, humidity, phosphorus, potassium, nitrogen, and pH measurements. First, we assembled a large dataset with 22 kinds of agricultural production components. Using many machine learning models, such as ensemble methods and baseline classifiers, we were able to classify crops with an astounding 99% accuracy rate. With the application of these insights, the CRS provides customized recommendations through an easy-to-use user interface for appropriate crops under particular climatic conditions. Our system’s innovative combination of hardware sensing capabilities and AI-driven decision-making promises to revolutionize crop management practices, offering actionable insights for agricultural stakeholders. Our system’s novel integration of AI-driven decision-making and hardware sensing capabilities promises to transform crop management techniques and provide agricultural stakeholders with useful insights.

**Title:**  Crop Prediction Model Using Machine Learning Algorithms

**Author:**  Ersin Elbasi, Chamseddine Zaki, Ahmet E.Topcu, Wiem Abdelbaki, Aymen I.Zreikat, Elda Cina, Ahmed Shdefat and Louai Saker.

**Year:** 2023

**Description**: — Machine learning applications are having a great impact on the global economy by transforming the data processing method and decision making. Agriculture is one of the fields where the impact is significant, considering the global crisis for food supply. This research investigates the potential benefits of integrating machine learning algorithms in modern agriculture. The main focus of these algorithms is to help optimize crop production and reduce waste through informed decisions regarding planting, watering, and harvesting crops. This paper includes a discussion on the current state of machine learning in agriculture, highlighting key challenges and opportunities, and presents experimental results that demonstrate the impact of changing labels on the accuracy of data analysis algorithms. The findings recommend that by analyzing wide-ranging data collected from farms, incorporating online IoT sensor data that were obtained in a real-time manner, farmers can make more informed verdicts about factors that affect crop growth. Eventually, integrating these technologies can transform modern agriculture by increasing crop yields while minimizing waste. Fifteen different algorithms have been considered to evaluate the most appropriate algorithms to use in agriculture, and a new feature combination scheme-enhanced algorithm is presented. The results show that we can achieve a classification accuracy of 99.59% using the Bayes Net algorithm and 99.46% using Naïve Bayes Classifier and Hoeffding Tree algorithms. These results will indicate an increase in production rates and reduce the effective cost for the farms, leading to more resilient infrastructure and sustainable environments. Moreover, the findings we obtained in this study can also help future farmers detect diseases early, increase crop production efficiency, and reduce prices when the world is experiencing food shortages.

**Title:** Crop Recommendation System Using Machine Learning Algorithm

**Author**: Prakalya Murali, Pradhusha Ayyasamy, Obuli.

**Year:** 2024**.**

**Description:** This study aims to develop an intelligent agricultural yield recommendation framework leveraging the capabilities of AI algorithms. The proposed framework takes yield efficiency and optimal growing seasons as crucial factors in generating appropriate crop recommendations. We have put forth four widely used models, namely Linear Regression (LR) and Multi-Layer Perceptron (MLP), which were trained and evaluated on a comprehensive dataset comprising historical agricultural data encompassing various features such as climatic factors, soil properties, and geographical variables. Furthermore, the data was segmented based on seasonal patterns to provide crop suggestions tailored to specific time periods. The performance of these models was assessed using standard metrics, and an ensemble approach was considered to enhance the system's robustness. Ultimately, the developed framework offers farmers and agricultural professionals a valuable tool for making informed decisions, optimizing crop selection, and enhancing overall agricultural productivity.

**PROPOSED SYSTEM**

The proposed system in this study aims to address the limitations of traditional crop prediction models by introducing a novel hybrid approach that combines machine learning with genetic optimization techniques, specifically focusing on improving the accuracy and robustness of crop classification. Our system utilizes a **Random Forest Classifier**, a widely recognized ensemble learning method known for its high performance and ability to handle large datasets with multiple features, to classify a diverse range of crops, including rice, jute, maize, and others. However, to enhance the efficiency and predictive capability of this model, we integrate a **Genetic Algorithm (GA)** to optimize its hyperparameters. The Genetic Algorithm, inspired by the principles of natural selection, iteratively searches for the best combination of parameters, such as the number of trees, maximum depth, and feature subsets, to ensure that the Random Forest model achieves optimal performance. This approach allows the model to adapt more effectively to varying soil and climatic conditions, which are crucial factors in agricultural crop prediction.

**PROPOSED SYSTEM ADVANTAGES:**

* Improved Accuracy and Performance
* Enhanced Interpretability with Explainable AI (LIME & SHAP)
* Optimized Hyperparameters through Genetic Algorithms
* Adaptability to Changing Agricultural Conditions
* Reduced Overfitting with Ensemble Learning (Random Forest).

**SYSTEM REQUIREMENTS**

**HARDWARE REQUIREMENTS**

The hardware requirements may serve as the basis for a contract for the implementation of the system and should therefore be a complete and consistent specification of the whole system. They are used by software engineers as the starting point for the system design. It should what the system do and not how it should be implemented.

* PROCESSOR : DUAL CORE 2 DUOS.
* RAM : 4GB DD RAM
* HARD DISK : 500 GB

**SOFTWARE REQUIREMENTS**

The software requirements document is the specification of the system. It should include both a definition and a specification of requirements. It is a set of what the system should do rather than how it should do it. The software requirements provide a basis for creating the software requirements specification. It is useful in estimating cost, planning team activities, performing tasks and tracking the teams and tracking the team’s progress throughout the development activity.

* Operating System : Windows 10
* Platform : Spyder3
* Programming Language : Python
* Front End : Spyder3

**SYSTEM ARCHITECTURE:**

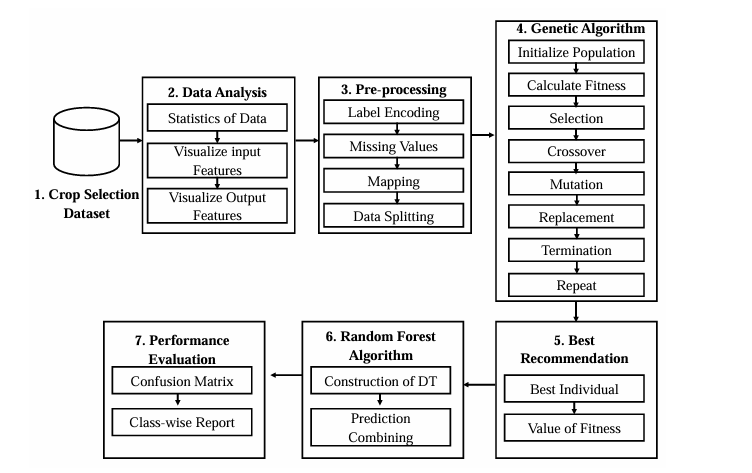


Fig: System Architecture

**REFERENCES**

[1] “The world factbook.” https://www.cia.gov/the-world-factbook/countries/bangladesh/. Accessed: 2023-8-10.

[2] “Bureau of south and central asian affairs.” https://2009-2017.state.gov/r/pa/ei/bgn/3452.htm. Accessed: 2023-8-10.

[3] “India economic survey 2018.” https://www.financialexpress.com/budget/india-economic-survey-2018-for-farmers-agriculture-gdp-msp-1034266/. Accessed: 2023-8-10.

[4] M. H. Imam, N. Nahar, R. Bhowmik, S. B. S. Omit, T. Mahmud, M. S. Hossain, and K. Andersson, “A transfer learning-based framework: Mobilenet-svm for efficient tomato leaf disease classification,” in 2024 6th International Conference on Electrical Engineering and Information

&Communication Technology (ICEEICT), pp. 693–698, 2024.

[5] Y. Akkem, S. K. Biswas, and A. Varanasi, “Smart farming using artificial intelligence: A review,” Engineering Applications of Artificial Intelligence, vol. 120, p. 105899, 2023.

[6] T. Mahmud, K. Barua, A. Barua, N. Basnin, S. Das, M. S. Hossain, and K. Andersson, “Explainable ai for tomato leaf disease detection: Insights into model interpretability,” in 2023 26th International Conference on Computer and Information Technology (ICCIT), pp. 1–6, IEEE, 2023.

[7] T. Mahmud, M. Ptaszynski, and F. Masui, “Exhaustive study into machine learning and deep learning methods for multilingual cyberbullying detection in bangla and chittagonian texts,” Electronics, vol. 13, no. 9, p. 1677, 2024.

[8] A. S. Khan, F. T. Khan, T. Mahmud, S. K. Khan, N. Sharmen, M. S.Hossain, and K. Andersson, “Integrating bert embeddings with svm for prostate cancer prediction,” in 2024 6th International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT), pp. 01–06, 2024.

[9] S.UmmeHabiba,F.Tasnim,M.S.HasanChowdhury,M.K.Islam,L.Nahar, T. Mahmud, M. S. Kaiser, M. S. Hossain, and K. Andersson, “Early prediction of chronic kidney disease using machine learning algorithms with feature selection techniques,” in International Conference on Applied Intelligence and Informatics, pp. 224–242, Springer, 2023.

[10] A. D. Bappy, T. Mahmud, M. S. Kaiser, M. Shahadat Hossain, and K. Andersson, “A bert-based chatbot to support cancer treatment follow up,” in International Conference on Applied Intelligence and Informatics,pp. 47–64, Springer, 2023.

[11] S. Das, T. Mahmud, D. Islam, M. Begum, A. Barua, M. Tarek Aziz, E. Nur Showan, L. Dey, E. Chakma, et al., “Deep transfer learning-based foot no-ball detection in live cricket match,” Computational Intelligence and Neuroscience, vol. 2023, 2023.

[12] T. Mahmud, M. Ptaszynski, J. Eronen, and F. Masui, “Cyberbullying detection for low-resource languages and dialects: Review of the state of the art,” Information Processing & Management, vol. 60, no. 5, p. 103454,2023.

[13] L. Bi and G. Hu, “A genetic algorithm-assisted deep learning approach forcrop yield prediction,” Soft Computing, vol. 25, pp. 10617–10628, 2021.

[14] V. Geetha, A. Punitha, M. Abarna, M. Akshaya, S. Illakiya, and A. Janani,“An effective crop prediction using random forest algorithm,” in 2020 International Conference on System, Computation, Automation and Networking (ICSCAN), pp. 1–5, IEEE, 2020.

[15] N. Jain, A. Kumar, S. Garud, V. Pradhan, and P. Kulkarni, “Crop selection method based on various environmental factors using machine learning,” International Research Journal of Engineering and Technology (IRJET),vol. 4, no. 02, pp. 551–558, 2017.

[16] S. Rani, A. K. Mishra, A. Kataria, S. Mallik, and H. Qin, “Machine learning-based optimal crop selection system in smart agriculture,” Scientific Reports, vol. 13, no. 1, p. 15997, 2023.

[17] S. Jain and D. 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.

[18] M.A.Devi, D. Suresh, D. Jeyakumar, D. Swamydoss, and M. L. Florence, “Agriculture crop selection and yield prediction using machine learning algorithms,” in 2022 second international conference on artificial intelligence and smart energy (ICAIS), pp. 510–517, IEEE, 2022.

[19] M.S.Rao,A.Singh,N.S.Reddy,andD.U.Acharya,“Cropprediction using machine learning,” in Journal of Physics: Conference Series, vol. 2161, p. 012033, IOP Publishing, 2022.

[20] E. Elbasi, C. Zaki, A. E. Topcu, W. Abdelbaki, A. I. Zreikat, E. Cina,A. Shdefat, and L. Saker, “Crop prediction model using machine learning algorithms,” Applied Sciences, vol. 13, no. 16, p. 9288, 2023.

[21] S.Raja, B.Sawicka, Z. Stamenkovic, and G. Mariammal, “Cropprediction based on characteristics of the agricultural environment using various feature selection techniques and classifiers,” IEEE Access, vol. 10, pp. 2362523641, 2022.

[22] D. Modi, A. V. Sutagundar, V. Yalavigi, and A. Aravatagimath, “Crop recommendation using machine learning algorithm,” in 2021 5th International Conference on Information Systems and Computer Networks (ISCON), pp. 1–5, IEEE, 2021.

[23] P. Murali, P. Ayyasamy, O. Obuli, et al., “Crop recommendation system using machine learning algorithm,” International Research Journal on Advanced Engineering Hub (IRJAEH), vol. 2, no. 05, pp. 1237–1242,2024.

[24] N. A. Patil and S. Mane, “Crop recommendation in precision agriculture using machine learning techniques,” 2024.

[25] K. Nischitha, D. Vishwakarma, M. N. Ashwini, and M. Manjuraju, “Crop prediction using machine learning approaches,” International Journal of Engineering Research & Technology (IJERT), vol. 9, no. 08, pp. 23–26, 2020.

[26] H. Karthikeya, K. Sudarshan, and D. S. Shetty, “Prediction of agricultural crops using knn algorithm,” Int. J. Innov. Sci. Res. Technol, vol. 5, no. 5,pp. 1422–1424, 2020.

[27] H. Cao, Y. Gu, J. Fang, Y. Hu, W. Ding, H. He, and G. Chen, “Application of stacking ensemble learning model in quantitative analysis of biomaterial activity,” Microchemical Journal, vol. 183, p. 108075, 2022.

[28] “Indian council of agricultural research.” https://icar.org.in. Accessed: 2023-8-10.

[29] M. L. Waskom, “Seaborn: statistical data visualization,” Journal of Open Source Software, vol. 6, no. 60, p. 3021, 2021.

[30] R. Guedrez, O. Dugeon, S. Lahoud, and G. Texier, “Label encoding algorithm for mpls segment routing,” in 2016 IEEE 15th International Symposium on Network Computing and Applications (NCA), pp. 113117, IEEE, 2016.