International Journal of Research Publication and Reviews, Vol (6), Issue (1), January (2025), Page – 1173-1177

International Journal of Research Publication and Reviews

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

SMART AGRICULTURE SIMPLIFIED: MACHINE LEARNING-BASED CROP MANAGEMENT AND FARMING SOLUTIONS

U. Mohammed Rizwan[1], Mrs.D. Shanmuga priyaa[2]

[1]III B.Sc.CS, Department of Computer Science, Sri Krishna Adithya College of Arts & Science, Coimbatore.
[2]Assistant professor, Department of Computer Science, Sri Krishna Adithya College of Arts & Science, Coimbatore.

Abstract

The Agricultural Portal is an inventive platform that gives farmers easy access to vital agricultural tools, resources, and information in order to improve crop productivity. Among its many features are crop planning tools, soil health evaluations, pest and disease management insights, real-time weather forecasts, and market price updates. The creation and execution of the portal are highlighted in this document, along with a description of its features and useful uses.

The site is intended to increase farmers' profitability, productivity, and decision-making skills. It meets the various needs of farmers in various regions and farm sizes and is based on a strong, scalable, and flexible technological architecture. Even isolated rural villages may use it because of its user-friendly design, which guarantees accessible on a variety of devices, including tablets and smartphones. The Agricultural Portal uses machine learning techniques, including Random Forest and Naïve Bayes, to predict agricultural yields. These data-driven analytics allow for more educated crop production, marketing, and storage decisions, all while avoiding climate variability concerns. The gateway marks a huge development in agricultural technology by altering traditional farming processes, paving the door for more sustainable and efficient farming around the world.

Keywords: farmers, agricultural technology, machine learning, Random Forest, Naïve Bayes, crop production, crop yield prediction, predictive modeling, agricultural portal, and user-friendly interface.

I. INTRODUCTION

Agriculture serves as the backbone of economies worldwide, playing a vital role in sustaining populations and driving economic growth. With a growing worldwide population, the sector faces increasing pressure to supply rising food demand. Agriculture in India not only provides a living for more than half of the workforce, but it also contributes significantly to the country's GDP, accounting for roughly 17-18%. Despite its importance, the agricultural sector faces issues such as inefficiency in farming operations, climate instability, and market volatility. India's agricultural landscape is characterized by a wide range of agro-ecological factors, including soil, rainfall, temperature, and cropping patterns. These differences, combined with unpredictable weather patterns and natural calamities such as floods, impair agricultural activity, resulting in lower productivity. Addressing these difficulties requires the use of cutting-edge technologies such as machine learning (ML) and data analytics, which have the potential to transform traditional farming techniques. Historically, agricultural development efforts in India have concentrated post-harvest interventions, with little emphasis on preventive measures. While weather prediction technologies exist, their reach and accessibility are restricted, leaving many farmers without critical climate information. Accurate agricultural yield estimation, enabled by ML approaches, can reduce production losses under bad conditions while increasing yields in favorable scenarios. Such predictive capacities are critical for making sound decisions at the global, regional, and field levels. This project aims to address these gaps through the development of an Agricultural Portal—a comprehensive online platform designed to empower farmers. By leveraging ML algorithms, the portal facilitates the prediction of future crop yields and favorable weather conditions. Additionally, it serves as a marketplace where farmers can sell their produce directly to consumers, eliminating intermediaries and enhancing economic outcomes. The Agricultural Portal incorporates key elements essential for modern farming, including:

Market Intelligence: Real-time market data helps farmers determine optimal times and prices for selling their produce, ensuring better profitability. Weather Data: Accurate forecasts and alerts enable efficient planning for planting, harvesting, pest control, and irrigation. Best Practices: Access to updated agricultural tools and methods empowers farmers to adopt modern techniques, improving productivity and crop quality. Machine learning underpins the portal's predictive functionalities. Algorithms like Decision Tree, Naïve Bayes Gaussian, K-Nearest Neighbor (KNN), and Random Forest use historical and environmental data to reliably anticipate crop yields. Among these, the Decision Tree algorithm has the best accuracy, making it a key component of the portal's predictive powers. India's agricultural past is firmly ingrained in its cultural and economic history, with natural crops supporting human, animal, and environmental systems. However, modern methods, especially the emergence of manufactured hybrids, have created new difficulties such as environmental degradation and unhealthy lifestyles. The lack of information about ideal farming practices exacerbates problems such as resource depletion and climate change, endangering food security. The Agricultural Portal seeks to transform traditional farming practices by integrating cutting-edge technologies into agriculture. By offering predictive insights, weather forecasts, and a direct-to-consumer marketplace, the platform fosters a resilient, sustainable, and economically viable agricultural ecosystem. This initiative underscores the transformative potential of technology in agriculture, ensuring food security and improving the livelihoods of farmers worldwide.

II. LITERATURE REVIEW

[1] Analysis shows that temperature, precipitation, and soil type are the most frequently utilized features. Artificial Neural Networks are the most often employed algorithm in these models. Analysis indicates that the most popular deep learning technique in these research is Convolutional Neural Networks (CNN), followed by Long-Short Term Memory (LSTM) and Deep Neural Networks (DNN).

In [2], the costs of a few key crops were examined in order to use meta-learning for time-series prediction. Crop price and yield datasets are trained using Meta-Learning Based Adaptive Crop Price Prediction (MLACPP), a combination of Self-Organized Map (SOM) and LSTM (Long-Short Term Model). The experimental results demonstrate a significant improvement over the current crop price prediction methods in terms of prediction accuracy and cross-correlation entropy.

By using characteristics such as State, district, season, and area, the author in

[3] allows the user to anticipate the crop production in the desired year. The study predicts the yield using sophisticated regression techniques such as Kernel Ridge, Lasso, and ENet algorithms. It also applies the idea of stacking regression to improve the algorithms and provide a more accurate prediction. Root mean square error is the performance statistic employed in this project.

The author of the

[5] research suggests using hyperparameter adjustment in combination with forward feature selection to train the random forest classifier. This method's accuracy is 71.88%. evaluated this method's accuracy against two baselines that weren't related to machine learning. The actual yield from the prior year serves as the first foundation for the forecast. The second baseline is that a human expert manually predicts each plot's goal yield.

[6] The yield of practically every crop grown in India is forecast in this study. By using basic parameters like state, district, season, and area, this script creates something new. The user can also anticipate the crop's yield in the desired year. The study predicts the yield using sophisticated regression techniques such as Kernel Ridge, Lasso, and ENet algorithms. It also applies the idea of stacking regression to improve the algorithms and produce a more accurate prediction.

[7] We may forecast weather conditions in this paper, such as rain, wind, sunshine, storms, floods, temperature changes, etc. These days, as society grows, the weather has a negative effect, leading to significant damage, injuries, and fatalities among farmers. For terrace gardening and agriculture, weather forecasting is crucial. Weather forecasting will improve crop output in remote places. This study discusses a low-cost method for predicting weather forecasts.

[8] In order to improve performance results for each chosen weather parameter, this study employs algorithms including Random Forest, Support Vector Machine, Weather, and K Nearest Neighbor. In order to determine which crops are appropriate for a certain farm or piece of land, we also consider weather and soil data including rainfall, maximum and lowest temperatures, soil fertility, and soil type. Ethiopia compiles agrarian overview data and remotely detected data (RSD) for a significant number of recipients of specially designated imported nutrition assistance.

III. METHODOLOGY

The development of the Agricultural Portal involves a comprehensive and structured methodology to ensure its efficacy, scalability, and usability. The process encompasses several phases, including data collection, machine learning model development, system architecture design, and user interface implementation. Below, the methodology is described in detail:

Sources of Data:

- Historical agricultural data from government records and agricultural research institutions.
- Weather data from meteorological departments and open-source platforms.
- Market prices from local and national agricultural markets.
- Soil health and pest-related data from agricultural extension services.

Data pre-processing: Prior to using the machine learning method, three datasets that were gathered as raw data must be treated. The data gathered is frequently deficient in specific behaviors or patterns, inconsistent, and incomplete. They probably have a lot of mistakes in them as well. Therefore, after being gathered, they undergo pre-processing to create a format that the machine learning algorithm can utilize for the model. utilized Python Pandas to analyze and visualize large amounts of data. Columns that are not needed are eliminated. Mean values are entered into empty columns.

Using an ML algorithm: The following are a few machine learning algorithms in use:

A supervised learning method that can be applied to both classification and regression issues is the decision tree. This classifier is tree-structured, with internal nodes standing in for a dataset's features, branches for the decision rules, and each leaf node for the result.

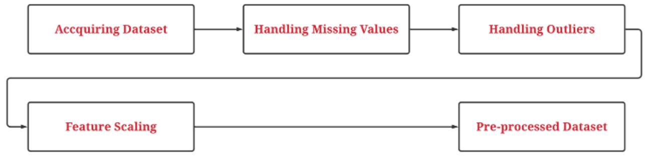
One well-known machine learning method that is a part of the supervised learning approach is called Random Forest. In machine learning, it can be applied to both classification and regression issues. Its foundation is the idea of ensemble learning, which is the process of merging several classifiers to solve a challenging issue and enhance the model's performance.A classifier called Random Forest uses a number of decision trees on different subsets of the dataset and averages them to increase the dataset's predicted accuracy. Rather than depending on a single decision tree, the random forest forecasts the final result by taking the predictions from each tree and calculating the majority vote of predictions. c) XGboost: XGBoost is a Gradient Boosted implementation.

Fig. 2. The web interface of crop yield prediction

TABLE I COMPARISON OF DIFFERENT ML ALGORITHMS ON DATASETS

Data-set used	Algorithm	Accuracy	precision	recall
Crop recommendation data	Support Vector Machine (SVM)	92%	83%	88%
	Random forest classifier	90%	81%	87%
	Gaussian Naive Bayes	88%	90%	89%
	XGBoost	95%	89%	91%
Crop Yield data	Algorithm	Accuracy	Mean Absolute Error	Root Mean Squared Error
	Random forest regression	92.40 %	81.2	367.574
	Ridge regression	70.60 %	645.76	1229.66
	Lasso	75.57 %	642.66	1229.18
	Decision Tree Regressor	86.60 %	105.101	470.38
	Algorithm	Accuracy	Mean Absolute Error	Root Mean Squared Error
Crop price data	Randomforest regression	91.40 %	94.2	468.584
	Ridge regression	72.50 %	685.78	1269.66
	Lasso	70.59 %	678.66	1274.18
	DecisionTreeRegressor	80.60 %	134.101	672.38

Methodology-preprocessing:



1. Purchasing the Dataset:

- Determine which sources of information are relevant to your problem statement.
- Gather the required data from these sources, such as web scraping, databases, and APIs.
- Ensure that the collected data is arranged in a structured way, such as a database table, CSV, or JSON.

2. Dealing with Absent Values:

- Look for any missing values in the dataset; these are represented as null values or NaN (Not a Number).
- Select an appropriate method for handling missing values while keeping in mind the data's properties and missingness trends.

Typical tactics include the following:

- When missing values are rare and the information loss is acceptable, the rows or columns with missing values are removed.
- Imputing the missing values by using a statistical measure such as the mean, median, or mode of the corresponding feature to fill them in.
- estimating missing values through the use of more complex imputation techniques, such as regression or machine learning algorithms.

3. Dealing with Outliers:

- Recognize and handle outliers, which are data points that significantly deviate from the distribution or normal range of the dataset.
- To identify potential outliers in the data, use visualization techniques like box plots or scatter plots.
- Choose the best course of action based on the particular issue and the type of data:
- eliminating outliers that have a significant impact on the study and are caused by mistakes in data entry or measurement issues.
- Using methods such as log transformation or winsorisation to change the data in order to lessen the effect of outliers.
- Applying robust statistical methods or algorithms that exhibit reduced sensitivity to outliers.

4. Feature Scaling:

- Normalize or scale the dataset's features to ensure they are on a similar scale and have comparable ranges.
- Common techniques for feature scaling include:
- Standardization (Z-score normalization): Adjusting the data so that it has a mean of zero and a variance of one.
- Min-max scaling: Adjusting the data to fit within a predetermined range, usually from 0 to 1.
- Robust scaling: Data scaling based on statistical measures that are less affected by outliers.
- It is important to scale features so that those with larger magnitudes do not dominate the learning process and so that algorithms converge more quickly.

5. Preprocessed Dataset:

- Once you finish the steps listed above, you will possess a preprocessed dataset that is ready for model training.
- This dataset requires cleaning, which involves addressing missing values, managing outliers, and scaling features appropriately.
- For future reference and to guarantee consistency in model development, it is recommended to save the preprocessed dataset separately.

Conclusion

This project presents a comprehensive system for predicting crop yield, weather conditions, and fertilizer recommendations using various machine learning algorithms. A web-based platform has been developed to provide these predictions, helping farmers make informed decisions. Among the algorithms explored, the Decision Tree algorithm was identified as providing the most accurate predictions for crop yield, weather, and fertilizer recommendations. The system takes user inputs and delivers precise predictive analysis, ensuring actionable insights for users. The website also offers information on the most suitable crops for specific soil and weather conditions and the fertilizers required for optimal growth. Furthermore, the Random Forest classifier demonstrated the highest accuracy for weather prediction and fertilizer recommendations. These features collectively help farmers maintain proper crop supply, manage costs effectively, and make data-driven decisions regarding crop selection and fertilizer usage. The robustness and reliability of the system have been validated through experiments conducted on a reliable dataset.

Future Enhancements

As a future enhancement, the application can be made more user-friendly by incorporating regional languages into the interface, enabling access to a broader population. Additional features such as a provision for uploading soil test reports, instead of manual input, can improve user experience. Moreover, this research can be extended to develop a recommender system for agricultural production and distribution, empowering farmers to decide which crops to grow during specific seasons for better profitability. The system could also be adapted to include predictions of expected market prices based on factors like location, soil type, and weather conditions, further aiding in financial planning. By continuously improving the interface and extending the system's capabilities, the platform can significantly benefit the agricultural community, fostering sustainable practices and enhanced productivity.

REFERENCES

[1] Thomas van Klompenburga, AyalewKassahuna, CagatayCatalb Crop yield prediction using machine learning: A systematic literature review Computers and Electronics in Agriculture 177, 2020 Elsevier, DOI: 10.1016/j.compag.2020.105709.

[2] D. K, R. M, S. V, P. N, and I. A. Jayaraj, Meta- Learning Based Adaptive Crop Price Prediction for Agriculture Application, in 2021 IEEE, 5th International Conference on Electronics, Communica- tion, and Aerospace Technology (ICECA), 2021, pp. 396-402, DOI: 10.1109/ICECA52323.2021.9675891

[3] P. S. Nishant, P. Sai Venkat, B. L. Avinash, and B. Jabber, Crop Yield Prediction based on Indian Agriculture using Machine Learning, 2020 International Conference for Emerging Technology (INCET), 2020, pp. 1-4, DOI: 10.1109/INCET49848.2020.9154036

[4] Y. J. N. Kumar, V. Spandana, V. S. Vaishnavi, K. Neha, and

V. G.R. R. Devi, Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector, 2020 5th International Conference on Communication and Electronics Systems (ICCES), 2020, pp. 736-741,

DOI: 10.1109/ICCES48766.2020.9137868

[5] Pausanias Charoen-Ung, PraditMittrapiyanuruk, Sugarcane Yield Grade Prediction Using Random Forest with Forward Feature Selection and Hyper-parameter Tuning Recent Advances in Information and Commu- nication Technology 2018 pp 33-42

[6] Potnuru Sai Nishant, Pinapa Sai Venkat, Bollu Lakshmi Avinash3, B. Jabber. “Crop Yield Prediction based on Indian Agriculture using Machine Learning”, International Conference for Emerging Technology (INCET) Belgaum, India. Jun 5-7, 2020

[7] ShivaniTuramari, Pooja Patil, MallappaHallad, ManjunathBilagi, “Weather Forecast Prediction for Agriculture”, (IJERT) ISSN: 2278-0181 ,Special Issue - 2022.

[8] PallaviShankarraoMahore, Dr. Aashish A. Bardekar, “Crop Yield Prediction using DifferentMachine Learning Techniques”, International Journal of Scientific Research in Computer Science, Engineering and Information Technology ISSN : 2456- 3307,Volume 7, Issue 3,May-June- 2021.