



Recommending and Predicting Crop Yield using Smart Machine Learning Algorithm (SMLA)

K. SUTHA*, N. INDUMATHI and S. UMA SHANKARI

Department of MCA, College of Science and Humanities, SRM Institute of Science and Technology, Ramapuram Campus, Chennai, Tamil Nadu, India.

Abstract

Agriculture is always needed by every human and responsible for the economic growth of a country. Developed countries likewise America, Japan, China are leading and making other countries too dependent on their technologies. But developing countries like India are expecting a lot of new technological innovations in the field of agriculture. Innovations may be in the form of smart machines, automation systems, sensor-based instruments, etc. and an advantage for society. In this paper, we have proposed Recommending and Predicting Crop Yield using Smart Machine Learning Algorithm (SMLA). The proposed algorithm namely SMLA is compared with other traditional algorithms to predict crop yield. In comparison to other algorithms the proposed algorithm works efficiently and produces 95% accuracy.



Article History

Received: 05 April 2023

Accepted: 18 September 2023

Keywords

Agriculture; Crop; Decision Support System; Yield; Smart Machine Learning Algorithm (SMLA).

Introduction

Today, the biggest challenges of farmers are determining which crops are suitable for the soil, what type of fertilizer to be use, how to deal with unexpected climatic changes and how to make crop production profitable by using minimum inputs. This research paper primarily focuses on yield recommendations and a prediction that uses a Smart Machine Learning Algorithm (SMLA). The proposed algorithm makes smart decisions and recommends appropriate sets with 95% accuracy. There is a lot of research work initiated by the Government of India and pitch your best idea to farmers.

Every farmer is expecting automation for various agricultural tasks such as crop selection, disease detection in crops, presides spraying of pesticides, prediction of crop yield, estimation of soil quality. There are many algorithms such as SVM, ANN, RF, MLR, and KNN that are being extensively used in the agricultural domain to find exact crops, perfect fertilizers, and disease identification. Out of these algorithms, RF and SVM is the process of analyzing exact crops, correct fertilizers, and disease in the leaf. Using the SMLA algorithm, a trained model is used to detect fertilizers and diseases.

CONTACT K. Sutha ✉ ksutha1986@gmail.com 📍 Department of MCA, College of Science and Humanities, SRM Institute of Science and Technology, Ramapuram Campus, Chennai, Tamil Nadu, India.



© 2023 The Author(s). Published by Enviro Research Publishers.

This is an  Open Access article licensed under a Creative Commons license: Attribution 4.0 International (CC-BY).

Doi: <https://dx.doi.org/10.12944/CARJ.11.2.30>

Related Work

This study inspects the application of ensemble approaches for predicting the crop type based on location-specific characteristics.¹ Machine learning may predict the output from a set of input parameters using supervised or unsupervised learning methods. We must have to create a suitable and satisfying function utilizing a set of variables that will represent the output using the provided input variables or parameters in order to obtain the required output parameter. Using search methods, authors were able to finalize 28242 occurrences and nearly 7 features across many databases. The key factors considered by our search approach are connected to climatic factors including temperature, rainfall and crop type. After conducting study, the conclusion was that an ensemble of Decision Tree and AdaBoost Regressors provided significant accuracy.

For farmers, this research suggested a practical and approachable yield prediction system. Through a Smartphone application, the proposed method connects farmers. GPS assists in locating the user. The user enters the location and soil type.² The most profitable crop list can be picked using machine learning algorithms, and forecasted crops yield for user-selected crops. Selected Machine Learning Algorithms, including SVM, ANN, RF, MLR, and KNN, are used to estimate crop production. The Random Forest among them demonstrated the greatest outcomes with a 95 percent accuracy rate. The algorithm also makes recommendations on when to apply fertilizers to increase yield.

Authors identified the best crop for the farmer, identified pests that may impact on crop yield and suggest the ideal pest management approaches. They used the SVM classification model, the Decision Tree Algorithm, and the Logistic Regression method in their study.³ They also discovered that the SVM classification model provides greater accuracy in comparison to the other algorithms. They concluded accounts' accuracy is to 89.66 percent. Moreover, they successfully predicted the crop and the pest control strategy to a quite satisfactory level after training a model with a large amount of data values, and the rules are inferred from algorithms like SVM and Decision Tree.

The proposed technique to forecast agricultural yield using historical data.⁴ By using Machine Learning

Algorithms like SVM and RF on agricultural data, this is accomplished and fertilizer recommendations are made that are appropriate for each individual crop. The research focuses on developing a prediction model that could be applied to crop yield forecasting in the future. It gives a succinct description on how agricultural yield can be predicted using machine learning methods. According to the research, RF algorithm has a better accuracy rate 86.35% when compared to SVM for soil categorization. In comparison to the RF algorithm, the SVM is good at predicting crop yields with an accuracy of 99.47%.

The use of machine learning techniques in agriculture to forecast crop yield and prescribe fertilizers and pesticides may assist to increase crop production quality and output, which will be beneficial for the economy of the nation.⁵ This study examines the value of machine learning in agriculture using SVM, KNN, RF and ANN. These algorithms are to predict the best crops for a given plot of land based on the seasons and soil characteristics, which can assist the farmers for crop selection decisions. In order to get the conclusion that the K-Nearest Neighbor Algorithm outperforms other methods, authors have investigated and evaluated the effect of these algorithms on crop prediction using the standard deviation and absolute error method.

The ability to anticipate yield is a critical issue in agriculture. Any farmer is curious to know how much of a crop he can anticipate.⁶ To analyze the soil's alkalinity, authors analyzed different factors that are associated, such as the location and the pH level. Location is used in conjunction with the use of third-party applications like APIs for type of soil, soil composition, weather, nutrient value of the soil in that region, temperature, and amount of rainfall in that region. Finally, authors focused on percentages of nutrients like Nitrogen (N), Phosphorous (P) and Potassium (K).

In order to boost productivity and profit from suggested techniques, farmers would benefit from proposed work by choosing the proper crop based on parameters like cost of cultivation, cost of production, and yield. In order to understand the various machine learning algorithms, the authors proposed new strategies in their research work.⁷ They evaluate their metrics for a particular dataset and propose an approach that outperforms competing learners.

In order to obtain a more effective result, the paper suggests a methodology in which the data set is processed using decision trees, k-th nearest neighbors, logistic regression, random forests, and gradient boosting classifiers.

Remote Sensing helps to differentiate land cover and vegetation much easier with the assistance of Normalized Difference Vegetative Index (NDVI).⁸ These findings recommend a recommendation framework to predict the crop yield with the help of various algorithms. These algorithms are closely associated with Machine Learning and statistical models like XGBoost, Gradient Boosting. In addition, the authors also extract trends by evaluating various parameters such as NDVI Rainfall, Temperature, Air Quality Index (AQI) etc.

The suggested technique for crop recommendations on the predictions of factors such as soil type, meteorological data, and the geographical conditions. Additionally, they have created a platform where farmers can utilize fertile unoccupied land for cultivation and get agricultural land by way of lease for the same.⁹ Users of this programme can learn about the latest developments in agriculture through the news feed interface and a location-specific Agri assistance option.

The project's foundation is the creation of an application that will enable farmers to forecast the output of crops in each region based on environmental factors like temperature and rainfall. The recommended model is used to estimate crop yield using a dataset of different crops from different regions of India, as well as rainfall and temperature data for the same locations. The proposed model's goal is to provide a tool that will deliver predictions based on an individual's crop.¹⁰ A list of recommendations helps to choose crops that are best for their farm based on details like location, farm size, weather, rainfall, and other crop datasets.

In order to boost agricultural yields, cut costs, and automate processing inputs, precision agriculture is suggested in this research. Data mining approaches for crop suggestion based on meteorological parameters, notably rainfall, temperature, and location, are used to modernize the process of farming crops.¹¹ For farming, fertilizers and seeds

are essential inputs, thus purchasing them in bulk and paying fair prices is essential. It might be tedious to buy something on an e-commerce website with the best price and quantity, and many times one misses out on things when exploring. ShopBot has demonstrated its value for online consumer search as a result of the rise in online shopping. A ShopBot compares products from several websites and offers a low-cost search option.

In recent years, machine learning has demonstrated its effectiveness in resolving agricultural issues like crop recommendation, crop production prediction, and many more. Numerous issues in the agricultural sector are minutely resolved thanks to developments in the machine learning subfield known as deep learning.¹² This study uses ensembling, correlation analysis, distribution analysis, and majority voting to select 22 different crop kinds. The crop recommendation problem is implemented using a three-tiered system. It contains modules for data pre-processing, classification, and performance assessment. Using ensembling methods, the following feature analysis is performed by utilizing correlation plots, density distribution and classification. Therefore, the majority voting method is used to evaluate performance. Additionally, this essay employs ensembles with fundamental learners, such as decision trees and random.

Crop cultivation is possible on soil that has vital nutrients. However, the use of additional fertilizers has led to a decline in some soil nutrients. The crop productivity is declining as a result. Therefore, to boost crop production, the suggested methodology uses all soil micro- and macronutrients to forecast the suitability of a crop for a given region.¹³ In order to increase forecast accuracy, beyond the rough set, fuzzy logic is utilized during data categorization to manage the boundary values of the numerical structures. The rules are created using a rough set-based rule induction method, and the fuzzy rules are utilized to predict the appropriateness of the crop. Benchmark methods used to compare the findings include CN2, LEM2, AQ, and Indiscernibility. The datasets that have been discretized and fuzzy are taken into consideration for experiments. The effectiveness of the algorithms is assessed using a variety of evaluation criteria, including precision, recall, f1 score, and accuracy.

This study introduces a decision support system that assists farmers in making informed decisions and aids them in determining soil quality, choosing the best crop, predicting rice disease, and preparing for disasters in Sri Lanka's wet zone. In order to help farmers or investors maximize yield, this initiative has integrated technologies including Image Processing, Machine Learning, Deep Learning, and Internet of Things.¹⁴ The four parts of "Agromate" are: determining the best crop to plant, predicting rice disease, and predicting natural disasters. Additionally, the application offers advice whenever rice diseases or natural disasters are discovered and proposes fertilizer whenever the soil is poor in quality. Users "can access" and use the system through an android mobile application that has been built. 'Agromate' is more likely to increase the productivity of crops.

Authors discussed IoT technology and how it supports precision farming through smart agriculture. To improve agriculture by improving production, the following properties are for measuring precise soil such as moisture, temperature, humidity, PH, and nutritional content/fertility, soil monitoring and Internet of Things (IoT). Proper data processes allowed us to optimize farming techniques effectively. All the processes are managed by the cloud itself and helps to analyze future data.¹⁵ Thus, authors concise about optimising yield by allocating resources and directing farming practises. They suggested the Internet of Things (IoT) system is made up of Wi-Fi-enabled microcontrollers and microprocessors, PH sensors, humidity, temperature sensors, soil moisture sensors, and soil nutrient sensors (NPK) probes. The sensors are used to measure the associated properties and send time-stamped real-time data to the cloud server. These sensors cooperate with one another to give the analyst accurate data. The SVM and Decision Tree method is suggested for the recommendation system to get the suitable crop for the provided soil data and aids to increase growth employing an optimum farming procedure.

Crop yield prediction using machine learning and deep learning algorithms discussed with real time data¹⁶. The systematic survey and decision support system is analyzed based on the features of temperature, rainfall, and soil type. Most of these features works with the help of Artificial Neural

Networks and Convolutional Neural Networks (CNN) model. Crop recommendation and yield prediction using ML algorithms discussed elaborately.¹⁷ Predicting the proper crop and production will support for better decisions making, reducing losses, and dealing the risk of price variations. Authors were applied classification and regression algorithms to get performance. ML techniques to predict crop yields based on weather data and link the data about production.¹⁸ Irish potatoes and maize data collected from various sources and analyzed using Random Forest, Polynomial Regression, and Support Vector Regressor.

Machine Learning Algorithms Support Vector Machine (SVM)

The Support Vector Machine (SVM) algorithm is a supervised learning algorithm found in the subfield of classification methods. It is a binary classification method that extrapolates an ideal hyperplane in n-dimensional space from the training dataset. New data sets can be sorted with the help of this hyperplane. Since it is a binary classifier, the hyperplane uses the training data set to create two categories. Data can be classified on a multidimensional hyperplane or a two-dimensional plane using SVM algorithms. In order to classify the data in the multidimensional hyperspace, "Kernels" are used.

Artificial Neural Network (ANN)

The human neuron serves as the basis for a machine learning method known as a neural network. The number of neurons in the human brain is in the millions. Electrical and chemical signals are used to transmit and process information. Synapses are the specific structures that link these neurons to one another. Signals are transmitted between neurons via synapse. Neural networks are built from networks of many virtual neurons. Information can be processed using an artificial neural network. It operates similarly to how the human brain does. The information in an ANN is processed by a network of several interconnected processors. They also produce useful outcomes as a result.

Random Forest (RF)

Regression and classification issues can be tackled with the use of a random forest, a machine learning technique. Ensemble learning, a method that uses

several classifiers to solve a problem, is used. Several decision trees make up a random forest algorithm. The random forest algorithm produces a "forest" that is "trained" using bagging or bootstrap aggregation. Machine learning algorithms can be made more precise by employing bagging, an ensemble meta-algorithm.

Multivariate Linear Regression (MLR)

A statistical method that combines many explanatory factors to predict the result of a response variable is known as multiple linear regression (MLR) or simply multiple regression. The purpose of multiple linear regression is to simulate the linear connection between the independent variables and the dependent ones. Simply put, when more than one explanatory variable is involved, ordinary least-squares (OLS) regression expands into multiple regression.

K-Nearest Neighbour (KNN)

- One of the simplest Machine Learning algorithms, K-Nearest Neighbour relies on the tried-and-true Supervised Learning method.
- The K-Nearest Neighbours (K-NN) algorithm classifies new instances or data by their presumed resemblance to previously classified data.
- The K-Nearest Neighbours algorithm remembers everything and then uses that information to provide a label to a new data point. This means that the K- NN algorithm can be used to effectively categorize new

data as it becomes available.

- Although the K-NN technique has applications in Regression, it is more commonly utilized for Classification.
- K-NN does not presuppose anything about the underlying data, as it is a non-parametric algorithm.
- It is also referred to as a "lazy learner" algorithm since it does not immediately use what it has learned from the training set but rather saves the information and applies it during classification.
- Throughout the training phase, the KNN algorithm simply saves the dataset, and when it receives new data, it assigns it to a category that best fits it.
- Let us say we have a picture of something that could be either a cat or a dog, but we really need to know which it is. Since the KNN algorithm is based on a similarity measure, it can be used for this identification. Our KNN model will compare the new set of data to the cat and dog photos, and classify the data based on the similarities between the two sets.

Results and Discussion

Fig 1 shows the home page of crop recommendation and prediction. The work was split into three major parts and displayed via user interface. The entire work mainly focuses on crop recommendation, fertilizer selection and disease affected in the leaves. This application window is easy to access and predict the most accurate result.

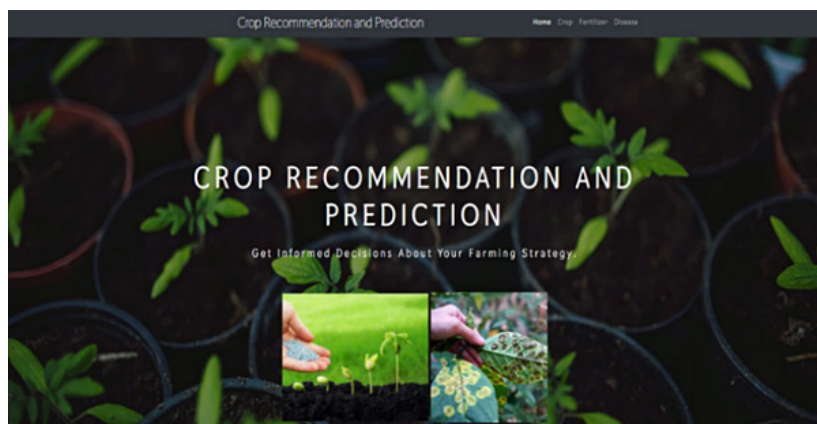


Fig. 1: Home page of Crop Recommendation and Prediction

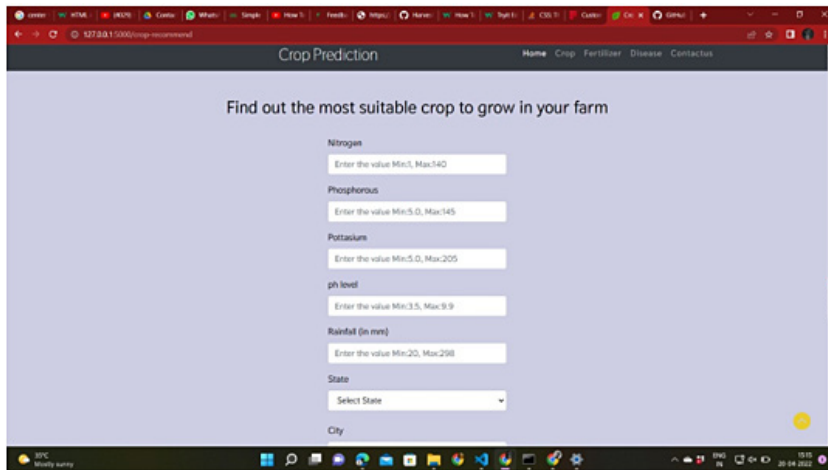


Fig. 2: Crop Selection

Fig 2 describes crop selection. During crop selection, we must enter Nitrogen, Potassium, Phosphorous, PH, Rainfall value and geo-location. Each value is trained by an intelligent algorithm and finalizing min

and max limits. Based on the values, an intelligent system trains the model and predicts the exact crop to grow in the soil. Fig 3 displays suitable crops predicted by inputs entered by the farmers.

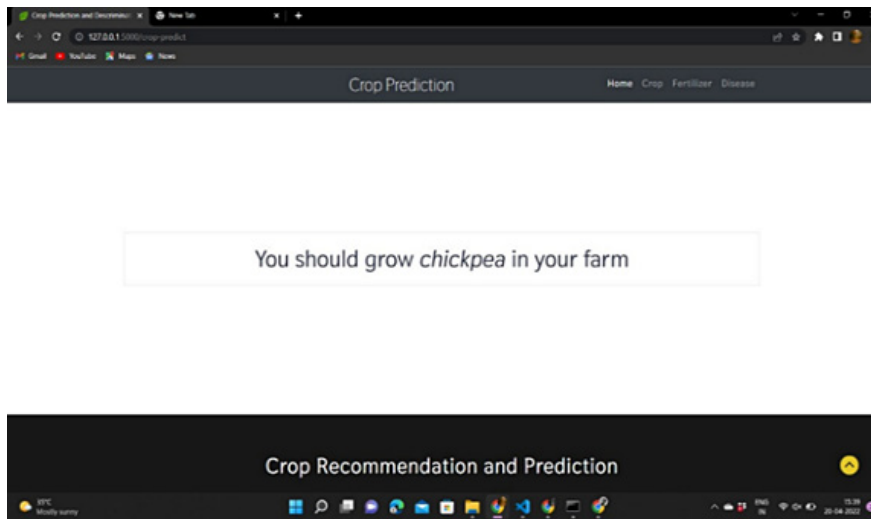


Fig. 3: Predicted Crop

Fig 4 illustrates fertilizer selection. Here we must enter Nitrogen, Potassium, Phosphorus and type of crop. Each input contains a minimum and maximum limit to select exact fertilizer. Fig 5 shows fertilizer recommendation. The fertilizer is predicted based on the soil and type of soil. Once fertilizer recommendation is finalized, we can easily increase the profit and avoid unnecessary expenses using intelligent machine learning algorithms.

Fig 6 shows plant disease. Trained dataset helps to analyze the diseases and gives better solutions to the problem. There are a lot of methods available to detect disease in a plant. Proposed SMLA algorithm works efficiently and detects disease affected in the plant.

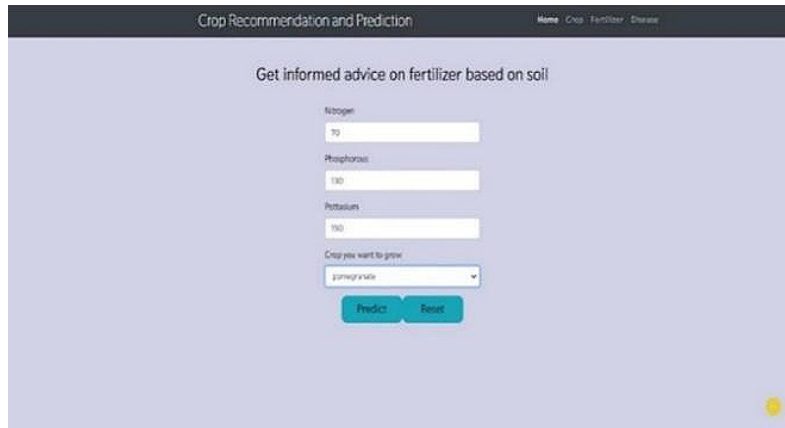


Fig. 4: Fertilizer Selection

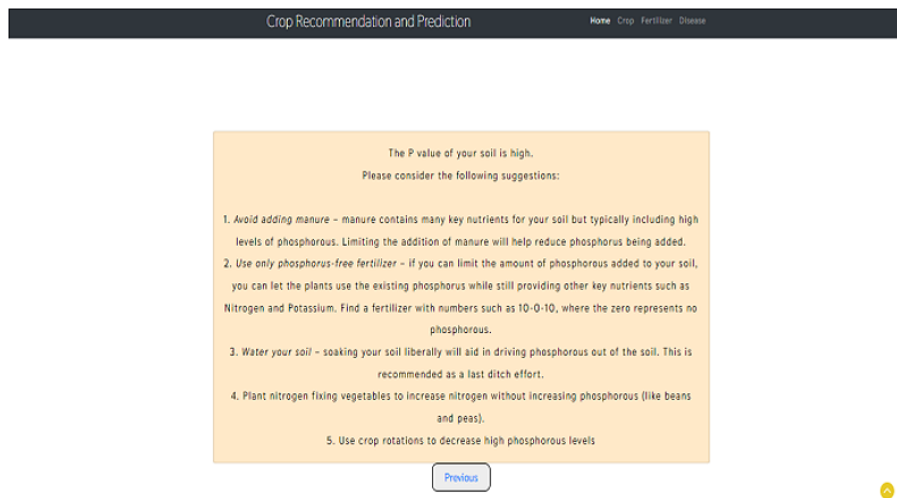


Fig. 5: Fertilizer Recommendation

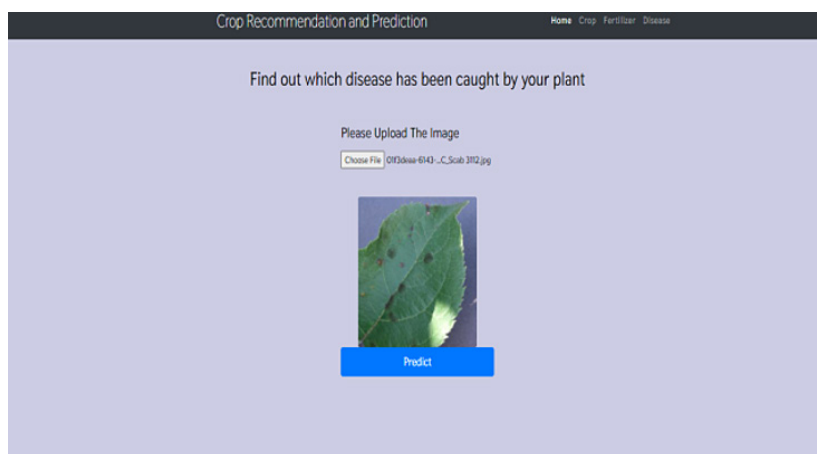


Fig. 6: Plant Disease

Conclusion and Future Work

Technological development has aided agricultural research, particularly by combining them within a sustainable agriculture production system. Technology has made farming a profitable enterprise by electrifying every farming procedure. In doing so, the middleman who typically buys low from farmers and sells expensive to end consumers is cut out, saving the farmer money. Site-specific decision modeling issues in agricultural systems are addressed by recent uses of computational intelligence techniques. This paper proposed the best Crop Yield Recommendation and Prediction Mechanism using Smart Machine Learning Algorithm. The proposed algorithms work efficiently and produce 95% accuracy. Many Internets of Things (IoT) devices provide automation methods. In future, farmers will be able to use intelligent IoT sensors to increase their use of smart farming practices and anticipate and recommend crop yield.

Acknowledgment

Authors thank the management for providing such a great platform to do this research work. Our sincere thanks to Dean(S&H), SRM Institute of Science and Technology, Ramapuram Campus, Chennai, India for his valuable guidance and constant support. The working model of this project is demonstrated in Research Day 2022 conducted by SRM Institute of Science and Technology, Ramapuram Campus and awarded first prize.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. M. Keerthana, K. J. M. Meghana, S. Pravallika and M. Kavitha, "An Ensemble Algorithm for Crop Yield Prediction," 2021 *Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, 2021, pp. 963-970, doi: 10.1109/ICICV50876.2021.9388479.
2. S. M. PANDE, P. K. RAMESH, A. ANMOL, B. R. AISHWARYA, K. ROHILLA and K. SHAURYA, "Crop Recommender System Using Machine Learning Approach," 2021 *5th International Conference on Computing Methodologies and Communication (ICCMC)*, 2021, pp. 1066-1071, doi: 10.1109/ICCMC51019.2021.9418351.
3. A. Kumar, S. Sarkar and C. Pradhan, "Recommendation System for Crop Identification and Pest Control Technique in Agriculture," 2019 *International Conference on Communication and Signal Processing (ICCSP)*, 2019, pp. 0185-0189, doi: 10.1109/ICCSP.2019.8698099.
4. Devdatta A. Bondre and Mr. Santosh Mahagaonkar (2019), Prediction of Crop Yield and Fertilizer Recommendation Using Machine Learning Algorithms, *International Journal of Engineering Applied Sciences and Technology*, 2019 Vol. 4, Issue 5, ISSN No. 2455-2143, Pages 371-376
5. K. Patel and H. B. Patel, "A Comparative Analysis of Supervised Machine Learning Algorithm for Agriculture Crop Prediction," 2021 *Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, 2021, pp. 1-5, doi: 10.1109/ICECCT52121.2021.9616731.
6. S. Bhanumathi, M. Vineeth and N. Rohit, "Crop Yield Prediction and Efficient use of Fertilizers," 2019 *International Conference on Communication and Signal Processing (ICCSP)*, 2019, pp. 0769-0773, doi: 10.1109/ICCSP.2019.8698087.
7. G. Dhabal, J. Lachure and R. Doriya, "Crop Recommendation System with Cloud Computing," 2021 *Third International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2021, pp. 1404-1411, doi: 10.1109/ICIRCA51532.2021.9544524.
8. A. Shah, R. Agarwal and B. Baranidharan, "Crop Yield Prediction Using Remote Sensing and Meteorological Data," 2021 *International*

- Conference on Artificial Intelligence and Smart Systems (ICAIS)*, 2021, pp. 952-960, doi: 10.1109/ICAIS50930.2021.9395849.
9. N. Valarmathi, N. Srija, S. Ananthi, B. P, E. T and N. N, "Amenity Agriculture App using Random Forest Algorithm," *2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, 2022, pp. 522-527, doi: 10.1109/ICSCDS53736.2022.9760901.
 10. Meeradevi and H. Salpekar, "Design and Implementation of Mobile Application for Crop Yield Prediction using Machine Learning," *2019 Global Conference for Advancement in Technology (GCAT)*, 2019, pp. 1-6, doi: 10.1109/GCAT47503.2019.8978315.
 11. A. Chaudhari, M. Beldar, R. Dichwalkar and S. Dholay, "Crop Recommendation and its Optimal Pricing using ShopBot," *2020 International Conference on Smart Electronics and Communication (ICOSEC)*, 2020, pp. 36-41, doi: 10.1109/ICOSEC49089.2020.9215411.
 12. R. K. Ray, S. K. Das and S. Chakravarty, "Smart Crop Recommender System-A Machine Learning Approach," *2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 2022, pp. 494-499, doi: 10.1109/Confluence52989.2022.9734173.
 13. A. M. Rajeswari, A. S. Anushiya, K. S. A. Fathima, S. S. Priya and N. Mathumithaa, "Fuzzy Decision Support System for Recommendation of Crop Cultivation based on Soil Type," *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*, 2020, pp. 768-773, doi: 10.1109/ICOEI48184.2020.9142899.
 14. D. S, S. M, N. P. R. N, R. W. A. I. M. N. Weerakkodi, A. Jayakody and N. Gamage, "Agro-Mate: A Virtual Assister to Maximize Crop Yield in Agriculture Sector," *TENCON 2021 - 2021 IEEE Region 10 Conference (TENCON)*, 2021, pp. 387-392, doi: 10.1109/TENCON54134.2021.9707199.
 15. R. Reshma, V. Sathiyavathi, T. Sindhu, K. Selvakumar and L. SaiRamesh, "IoT based Classification Techniques for Soil Content Analysis and Crop Yield Prediction," *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, 2020, pp. 156-160, doi: 10.1109/I-SMAC49090.2020.9243600.
 16. Thomas van Klompenburg, Ayalew Kassahun, Cagatay Catal, Crop yield prediction using machine learning: A systematic literature review, *Computers and Electronics in Agriculture*, Volume 177, 2020, 105709, ISSN 0168-1699, <https://doi.org/10.1016/j.compag.2020.105709>.
 17. Sundari V, Anusree M, Swetha U and Divya Lakshmi R (2022) Crop recommendation and yield prediction using machine learning algorithms, *World Journal of Advanced Research and Reviews*, 14(03), 452-459.
 18. Kuradusenge, M., Hitimana, E., Hanyurwimfura, D., Rukundo, P., Mtonga, K., Mukasine, A., Uwitonze, C., et al. (2023). Crop Yield Prediction Using Machine Learning Models: Case of Irish Potato and Maize. *Agriculture*, 13(1), 225. MDPI AG.