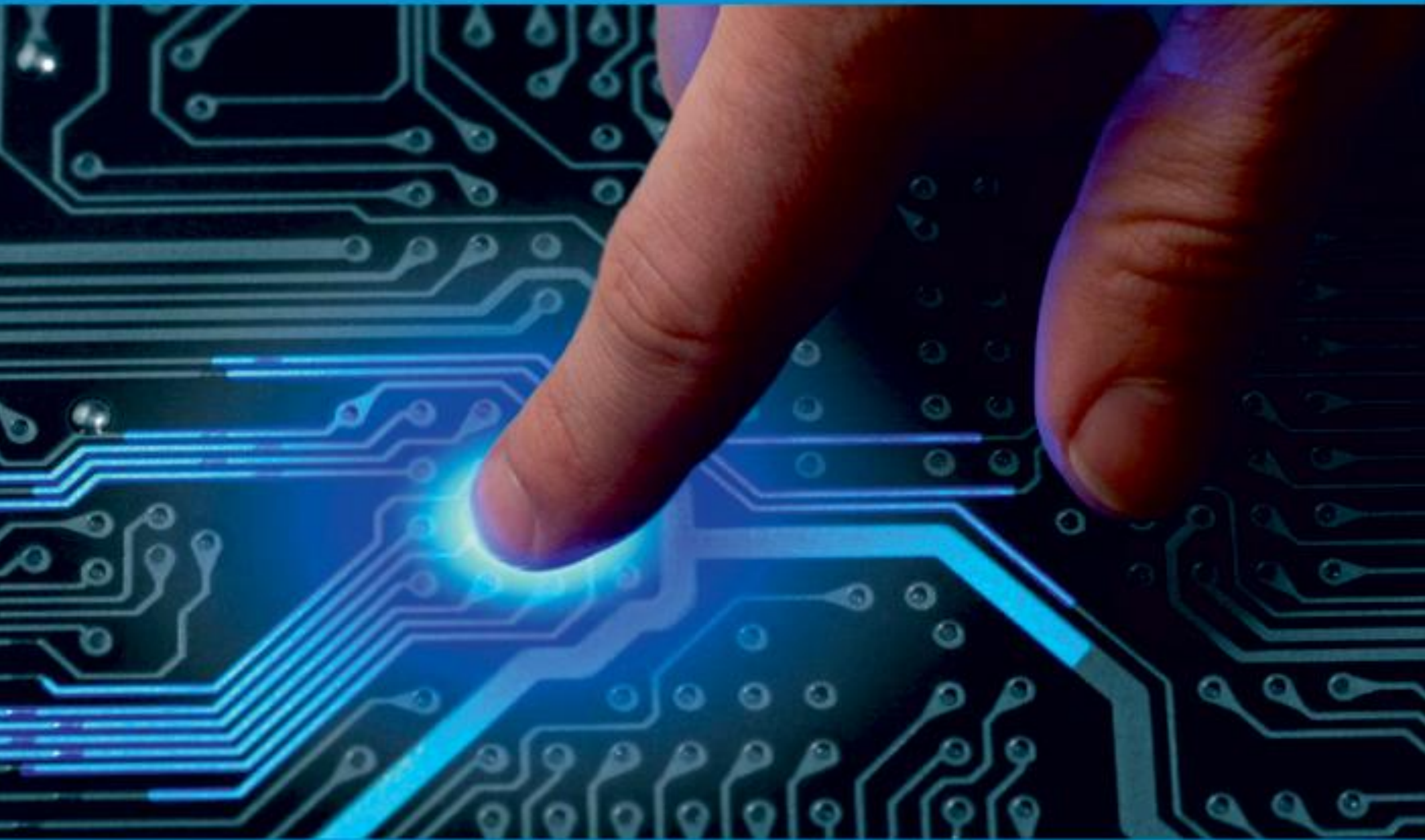
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205293 |

E-Krishi: Enhanced Crop Production Through Machine Learning Using Water, Soil and Weather as Features

Gurmeet Jandoo, Pranjal Kumar, Weona Lazarus, Dr. Ganesh Pathak

U.G. Student, Department of Computer Science and Engineering, MIT School of Engineering, MIT ADT University, Loni Kalbhor, Pune, India

Faculty, Department of Computer Science and Engineering, MIT School of Engineering, MIT ADT University, Loni Kalbhor, Pune, India

ABSTRACT: By utilising machine learning, the E-Krishi application is an agricultural initiative with the goal of increasing the productivity of India's agricultural sec-tor. Indian farmers face challenges with inefficient resource use and limited awareness of agro-technology. The E-Krishi app uses machine learning to optimise agricultural output and enhance farmers' lives in order to solve these concerns. The three main elements that need to be improved for efficient agricultural production optimization are soil, water, and crops. Modern sensor technology can acquire crucial scientific data that traditional agricultural methods frequently ignore, resulting in subpar crop yields. This study offers a thorough analysis of the literature on the use of machine learning in agricultural production systems.

By utilising machine learning and big data technologies, farming management systems are evolving into real-time artificial intelligence-powered programmes that offer beneficial recommendations and insights for enhancing farmers' decision-making processes. A more effective agricultural practice may be created because of the fusion of machine learning and the Inter-net of Things. Users can use a mobile application to track soil fertility and decide whether to grow particular crops for financial gain.

The integration of machine learning and IoT in agriculture has the capability to transform agricultural practices by offering real-time data-driven in-sights and allowing farmers to optimise resource utilisation, increase crop yields, and make educated decisions. The E-Krishi initiative is an important step towards making India's agriculture industry more sustainable and productive. Using the data about the crops grown, the soil health and moisture in the soil, the farmers can be intelligently informed about some ways in which their produce can be optimised.

KEYWORDS: Crop optimization; Machine Learning; Artificial Intelligence; Mobile application, Precision agriculture; Sustainable agriculture; Smart farming

I. INTRODUCTION

India is a global agricultural powerhouse. For millennia, agriculture has been the backbone of India's economy. It is the world's largest producer of milk, pulses, and spices, as well as the largest cow herd (buffaloes) and the largest area under wheat, rice, and cotton cultivation. It ranks second in the production of rice, wheat, cotton, sugarcane, farmed fish, sheep and goat meat, fruit, vegetables, and tea. The country has around 195 million ha under cultivation, with approximately 63 percent being rainfed (nearly 125 million ha) and 37 percent being irrigated (70 million ha).

The burden on farmers has intensified as a result of the nation's ever-expanding population and increased food consumption. Nevertheless, Indian farmers have not been able to maximise their crop production to meet the in-creased demand due to a lack of knowledge and resources. According to a study, farmers should use current inputs to offset potential losses due to climate change. Indian agriculture sector needs to increase its production efficiency to fulfil these demands, and contemporary technologies like machine learning and IoT can be quite helpful in the area. By utilising machine learning and the Internet of Things, the agricultural project E-Krishi seeks to increase the productivity of the industry.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205293 |

The main issues that Indian farmers face are the ineffective use of resources and their lack of awareness of agro-technology. The three crucial elements that must be improved to maximise agricultural yield are soil, water, and crops. Farmers do not take into account the scientific information that can be gathered using modern sensor technology in their traditional farming techniques. Farmers suffer a loss as a result of not having an optimal crop yield. The E-Krishi project, which utilises machine learning to maximise crop output and enhance farmers' quality of life, tackles this problem.

The main objectives of the E-Krishi project are as follows:

1. To create an AI-based farm management system that can provide real-time in-sights and recommendations to farmers.
2. To develop a mobile application that can be used by farmers to monitor soil fertility, crop growth, and water usage.
3. To optimise crop production by identifying the optimal amount of water and nutrients required for different crops.
4. To reduce resource wastage by enabling farmers to make informed decisions based on real-time data.
5. To improve the livelihoods of farmers by increasing their productivity and profitability.
6. To facilitate farmers with quick access to reports thereby, allowing them to improve farming techniques dynamically.

India's agriculture sector desperately requires tools and techniques that ease the burden of decision-making when deciding what fertiliser to use and how much water to feed the crops. The weather in India's agricultural areas has be-come incredibly unpredictable as a result of the growth in global warming, which has farmers concerned about their livelihoods and the health of their crops. Ma-chine learning integration into precision agriculture is of utmost relevance since it can be used to predict weather changes, sample soil to identify its nutrients, and predict which crops would thrive under such weather and soil circumstances.

II. METHODOLOGY

The E-Krishi project aims to optimise crop production through Machine learning and Internet of Things. To accomplish this goal, the following steps have been taken:

a) Data Collection: The project collects data on soil moisture, temperature, humidity, and other relevant parameters using IoT sensors. This information is sent to a central server, where it is processed and analysed by machine learning algorithms.
b) Machine Learning: Machine learning algorithms are used to analyse collected data and provide insights into the optimal amount of water and nutrients needed by various crops. Based on the nutrient composition of the soil, the algorithms also recommend the type and quantity of fertiliser to be used.
c) Mobile App: The project's mobile app gives farmers real-time information on soil fertility, crop growth, and water usage. The application also makes recommendations on the ideal amount of water and nutrients to use for various crops.
d) Farm Management System: The AI-powered farm management system provides farmers with real-time insights and recommendations for crop production optimization. The system analyses the collected data and makes crop management recommendations using machine learning algorithms

III. COMPARISON OF VARIOUS MACHINE LEARNING MODELS

Predictions can be made using a variety of machine learning methods. The criteria for using data and producing outputs vary between machine learning models. Here are some machine learning algorithms that have been compared to examine their advantages and disadvantages.

Table 1. Comparison of various machine learning models

Algorithm	*Complexity Characteristics*	*Limitations*	*Current Applications*
Artificial Neural Networks (ANN)	Very complex, use more power, are inclined to overfitting, require	High power consumption, overfitting, large dataset	Classification of patterns, attribute mapping, crop estimation,

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205293 |

	large datasets, and require feature adjustment.	requirements, and time-consuming parameter tuning	investigation of soil attributes, and estimation
Decision Trees (DT)	Low complexity, sensitivity to inconsequential changes in training data, sporadic instability, and propensity for overfitting	Instability occasionally and sensitivity to small variations in training data	Using soil variables to categorise and estimate crop yields
Support Vector Machines (SVM)	Low complexity, quick and accurate, simple to implement, small training sample size, robust to noise in training data, and less prone to overfitting	Mostly used for classification tasks	Disease classification, soil mapping, vegetation attribute retrieval, and estimates
Random Forests (RF)	Greater efficiency, high prediction performance, short training time, simple parametrization, but prone to overfitting	Overfitting problem	Attribute mapping, classification, regression, and yield estimation
Deep Learning (DL)	High computational cost, huge dataset requirements, highly promising for agriculture applications	Require large datasets and high computational costs.	Measure features, yield estimation, applications for classification and regression, etc.

As seen from the above table, Machine learning algorithms have various advantages and disadvantages. The ideal algorithm for a given application will vary depending on its individual needs. An ANN could be the best option, for instance, if the model needs to categorise a lot of photos. An RF would be a preferable choice if the model needs to anticipate a continuous variable, like agricultural production. Additionally, a DT may be the ideal option if the project aims to generate predictions using a smaller dataset. The E- Krishi project uses the Decision Learning algorithm

coupled with the Random Forest algorithm to make the best predictions of which crop will be best suited for the nutrients present in a particular soil sample.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205293 |

IV. PROPOSED SYSTEM

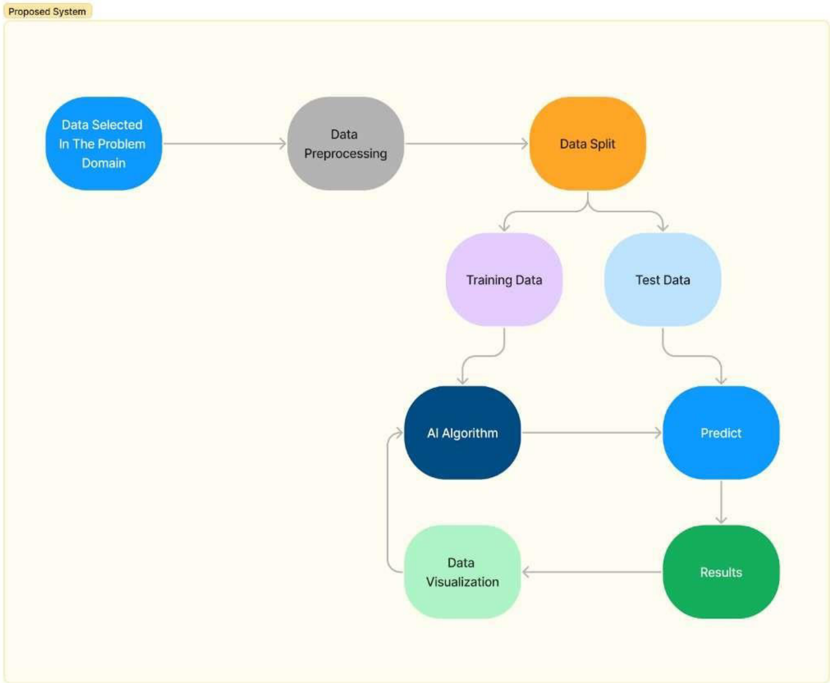


Fig. 1. This Figure Shows the Proposed System and The Flow Of Data Through The Model

E-KRISHI MODEL ELUCIDATION AND MOBILE APPLICATION USER INTERFACE

Exploratory data analysis (EDA) is a critical stage in any data science endeavour (EDA). The usage of EDA helps the project participants better understand the data they are working with and identify trends, connections, and potential issues that could be affecting model performance. The following is a list of possible steps one might take when performing EDA for a regression machine learning problem:

1. Data gathering and examination: The dataset must be gathered and examined first. The quality of the data can be determined by knowing the data's source and format, which are both crucial. To avoid having our model perform poorly, we look for duplicates, outliers, and missing values.
For this project, we trained our machine learning model on a dataset ac-quired from Kaggle as the test dataset. The dataset's specifics are as follows:

i. Context
Precision farming is becoming popular. This is due to the fact that it empowers farmers to decide on their farming strategies with confidence. This dataset can be used to build a prediction model that, depending on a number of variables, will recommend the best crops to plant on a specific farm. The rainfall, climate, and fertiliser data sets that were previously available for India were improved to create this dataset.

ii. Parameters in the dataset
N: ratio of nitrogen content in the soil
P: ratio of phosphorus content in the soil
K: ratio of potassium content in the soil
Temperature: temperature in degrees celsius
Humidity: relative humidity in percent
pH: pH value of soil
Rainfall: rainfall in mm

2. Visualisation before training the model: Visualising the data is the first stage once it has been cleansed. Plots like scatter plots, which allow us to examine the relationship between two variables at once, have been extensively

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205293 |

used. Histograms are used to comprehend how the data are distributed. With five numbers: minimum, lower quartile (25th percentile), median (50th percentile), higher quartile (75th percentile), and maximum data values, box plots are used to show outliers in the data as well as the statistical overview of the data.

The project uses various plot kinds to spot trends and patterns among all the features. To visualise the relationship between all the features pairwise, we also created a correlation plot. Any non-linear relationships that might not be obvious from the data might be found with this assistance. Some deductions from visualising the data are as follows:
During the rainy season, the average yearly rainfall is high (120 mm), and the temperature is low (less than 30°C). The pH of the soil is impacted by soil moisture, which is impacted by rain. The following list includes the crops that are most likely to be sown at this time of year.

Rice requires both heavy rainfall (>200 mm) and high humidity levels (above 80%). It makes sense that most of India's rice is grown along its east coast, which receives 220 mm of rainfall yearly on average.
The considerable exports from coastal regions around the country are due to the fact that coconut is a tropical crop that demands high humidity. The nutritional value of food is directly impacted by these elements of the soil. Fruits rich in nutrients frequently contain stable potassium levels.

3. Feature selection: EDA aids in choosing the features that are most pertinent to the output variable. We decided to train our machine learning models with all the features because the correlation plots indicated that most features are not overly connected, with the exception of nitrogen and phosphorus.
i. Outlier detection: Outliers can impair the performance of a model. Any outliers in the dataset can be found with the aid of EDA. To find outliers, we used box plots and made the appropriate corrections to the dataset.
ii. Normalisation of data is the process of scaling the data to a predetermined range. Data normalisation is crucial for regression machine learning models. Which normalisation technique to employ can be determined by EDA. /Train test split

4. Model selection: EDA assists in selecting the optimal model to use with the data. Use scatter plots or correlation plots to look for any non-linear relation-ships. In order to choose the best model for the task at hand, the problem, data, candidate models, and evaluation criteria are all carefully considered throughout the model selection process.
The models that were developed, tested, and cross-validated include the following:

i. Multinomial Logistic Regression
The multinomial logistic regression technique extends the logistic regression model by changing the loss function to a cross-entropy loss and the predicted probability distribution to a multinomial probability distribution in order to natively handle multi-class classification issues. The various classes here represent the various types of crops that can be grown successfully. Cross validation with five folds using the scikit-learn logistic regression implementation resulted in an average cross validation accuracy of 95.5%.

ii. Decision Tree
Decision Tree is the next model under consideration. A non-parametric super-vised learning approach called a decision tree is used for both classification and regression applications. A root node, branches, internal nodes, and leaf nodes make up its hierarchical tree structure. Decision tree learning employs a greedy search method, a divide-and- conquer strategy, to locate the best split points in-side a tree.
Using CART 5.0 (Classification and Regression Trees) and k-fold cross validation with 5 folds, an accuracy of 91.40% was attained.

iii. Random Forest
A random forest is a meta estimator that fits multiple decision tree classifiers to various subsamples of the dataset, increasing predicting accuracy and decreasing overfitting.

The ensemble model with 20 trees and fivefold cross validation had an accuracy of 99.45%.

Hence, for our application, we went with this random forest model. Farmers can learn about the best crops to grow based on the findings of the nutrient analysis of the soil.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205293 |

Model Deployment: The application comes equipped with this model. The application advises the farmer of the crop alternatives after receiving the farmer's soil test results from the soil testing laboratory.

5. Visualising the prediction: The mobile application will plot the most recent soil data and the soil statistics that are best for the predicted crop after the model produces a prediction in order to show the farmer which nutrients are low in the soil and which nutrients are abundant. When they are balanced for the best crop output, the software will then suggest plans and techniques.

V. FUTURE SCOPE

The E-Krishi project has incredible potential. Following are a few points for the prospective enhancements in the project:

1. IoT Integration: The E-Krishi project will be even more tailored to each farmer's land and the agricultural practices will be even more exact by integrating sensor networks to gather real-time soil data and meteorological conditions.
2. Pest and Disease Prediction: Machine learning algorithms may be trained to anticipate and identify pest infestations and illnesses by using historical data, meteorological information, and crop health information. Early diagnosis can assist farmers in taking preventive actions, such as crop rotation or targeted pesticide use, to avoid crop loss and boost total production.
3. Market Analysis and Price Forecasting: By incorporating market data and trends into the E-Krishi initiative, farmers will be better able to assess market demands, choose the best crops, and develop effective pricing strategies. In order to forecast price changes, machine learning algorithms can analyse previous market data, giving farmers the opportunity to choose the best crops and make the most money. Farmers may be thus further empowered to make informed decisions about their crop selection and pricing strategies by combining market data and price forecasting.
4. Community and Knowledge Sharing: By providing a forum for farmers to discuss their failures, successes, and farming methods, we may help to build a feeling of community among farmers and promote knowledge sharing.
5. Collaborative Data Sharing: Creating a platform for collaborative data sharing among farmers, researchers, and agricultural experts can enhance the effectiveness of the E-Krishi project. By pooling together data from various sources, the system can gain more insights and improve its predictive capabilities.
6. Real-time Market Data: Integrating real-time market data into the mobile application can enable farmers to access up-to-date information on crop prices, demand, and supply, empowering them to make timely decisions and maximise their profits.

VI. CONCLUSION

The E-Krishi project makes use of machine learning and intends to make use of IoT as the initial stage in future technological advancements to boost agricultural output and better the quality of life for farmers. The effort collects data on soil moisture, temperature, humidity, and other relevant factors and analyses it using machine learning algorithms to provide farmers with real-time insights and suggestions. Using the mobile app developed as part of the project, farmers may keep an eye on crop growth, soil fertility, and water usage and then make their decisions on the proposed measures. increasing industry productivity by utilising IoT and machine intelligence. Farmers may get real- time data on soil fertility, crop growth, and water use using the project's smartphone application. The initiative offers a practical response to the problems facing the Indian agriculture industry.

ACKNOWLEDGEMENTS

We would like to express our sincere gratitude to our director, professor Rajneeshkaur Sachdeo, as well as to our mentor, professor Ganesh Pathak, for providing us with the chance to work on this worthwhile project, "E-Krishi: Enhanced Crop Production through Machine Learning Using Water, Soil, and Weather as Features." This project gave us in-depth understanding of machine learning models, database management, and the method for developing user- friendly applications employing UI/UX design concepts.

We are appreciative to Krishi Vigyaan Kendra, Baramati, Maharashtra, for let-ting us tour their campus, see their soil laboratory, and learn from the experts there about significant agricultural practices. We would like to thank Prof. Dr. Swati Shirke, who made this field trip possible.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205293 |

Through our study, the project offers insightful data that aspiring data analysts and app developers may use to carefully plan the layout of their future applications. Every single person in the team gave thoughtful contributions to make this project a success.

REFERENCES

1. Liakos, Konstantinos G., et al. “Sensors | Free Full-Text | Machine Learning in Agriculture: A Review.” MDPI, 14 Aug. 2018, www.mdpi.com/1424-8220/18/8/2674.
2. Mandal, S. K., et al. “Assessment of River Water Quality for Agricultural Irrigation - International Journal of Environmental Science and Technology.” SpringerLink, 21 Feb 2018, link.springer.com/article/10.1007/s13762- 018-1657-3.
3. Hindawi, et al. “Machine Learning Algorithm for Soil Analysis and Classification of Micronutrients in IoT-Enabled Automated Farms.” Machine Learning Algorithm for Soil Analysis and Classification of Micronutrients in IoT-Enabled Automated Farms, 8 June 2022, www.hindawi.com/journals/jnm/2022/5343965.
4. “Implementation of Artificial Intelligence in Agriculture for Optimization of Irrigation and Application of Pesticides and Herbicides - ScienceDirect.” Implementation of Artificial Intelligence in Agriculture for Optimization of Irrigation and Application of Pesticides and Herbicides - ScienceDirect, 22 Apr. 2020, www.sciencedirect.com/science/article/pii/S258972172030012X.
5. “Implementation of Artificial Intelligence in Agriculture for Optimization of Irrigation and Application of Pesticides and Herbicides - ScienceDirect.” Implementation of Artificial Intelligence in Agriculture for Optimisation of Irrigation and Application of Pesticides and Herbicides - ScienceDirect, 22 Apr. 2020, www.sciencedirect.com/science/article/pii/S258972172030012X.
6. S. A. Bhat and N. -F. Huang, "Big Data and AI Revolution in Precision Agriculture: Survey and Challenges," in IEEE Access, vol. 9, pp. 110209-110222, 2021, doi: 10.1109/ACCESS.2021.3102227.
7. R. Sivakumar et al., “Internet of Things and Machine Learning Applications for Smart Precision Agriculture | IntechOpen,” Internet of Things and Machine Learning Applications for Smart Precision Agriculture | IntechOpen, May 07, 2021. https://www.intechopen.com/chapters/76652 (accessed Oct. 12, 2022).
8. Gebregiorgis, M.J., Savage, M.F., 2006. Soil-plant-atmosphere continuum research unit, School of Environmental Sciences, University of KwaZulu-Natal. South African Journal of Plant and Soil 23 (3), 145–151.
9. N. Zhang, M. Wang, and N. Wang, ‘‘Precision agriculture—A worldwide overview,’’ Comput. Electron. Agricult., vol. 36, nos. 2–3, pp. 113–132, 2002.
10. R. Gebbers and V. I. Adamchuk, ‘‘Precision agriculture and food security,’’ Science, vol. 327, no. 5967, pp. 828–831, Feb. 2010.