Cash Crop Recommendation System using Machine Learning Algorithms

Abhijeet Arun Pawar

Aakif Mohammad Iqbal Shaikh,**

**(Computer Science, CKT ACS College, New Panvel(Empowered Autonomous))*

Email : abhijeetpawar1575@gmail.com

Email : aakif.sh09@gmail.com

How to Cite this Article:

Pawar, A. A. & Shaikh, A. M. I. (2026). Cash Crop Recommendation System using Machine Learning Algorithms. International Journal of Creative and Open Research in Engineering and Management, 02(04). https://doi.org/10.55041/ijcope.v2i4.403

https://doi.org/10.55041/ijcope.v2i4.403

Abstract :

Agricultural productivity is very important for making sure that people have enough food and for helping rural economies. Traditional farming methods often lead to low crop yields because decisions are not based on data and resources are not used properly. This paper presents a Cash Crop Recommendation System (CCRS) that utilises machine learning algorithms to suggest appropriate crops according to soil and environmental conditions. The suggested system combines several datasets, such as historical crop yield records, soil parameters, and climate data, to make predictions more accurate. Random Forest and Support Vector Machine are two examples of supervised learning methods that are used to classify and suggest the best crops. The system also gives fertiliser suggestions based on the levels of Nitrogen (N), Phosphorus (P), and Potassium (K) in the soil to make it more fertile. The system is made to be easy for farmers to use and get to through a simple interface. The proposed model's goals are to boost agricultural output, lower its effect on the environment, and promote sustainable farming.

Keywords: Cash Crop Recommendation System, Machine Learning, Random Forest, Support Vector Machine, Precision Agriculture, Crop Yield Prediction.

I. Introduction

India is a country that relies heavily on farming, and many of its people depend on it for their livelihoods. But traditional farming methods, a lack of technological support, and bad decision-making often keep agricultural productivity low. Farmers often have trouble choosing the right crops and fertilisers, which has a direct effect on the health of the soil and the yield of the crops.

Using too much or the wrong kind of fertiliser can harm the soil, pollute groundwater, and lower the productivity of farms. Also, farmers have a hard time making good choices about what crops to grow and how to use resources when they own small, scattered pieces of land. These problems show that farmers need better technology to help them be more productive and environmentally friendly. Machine learning has become a useful way to look at agricultural data and make accurate predictions. Earlier studies have demonstrated that algorithms like Random Forest and Back Propagation can effectively predict crop yield and optimise fertiliser application [1]. These methods help find patterns in old data and make better decisions in farming.

Also, machine learning-based crop recommendation systems look at things like soil nutrients, temperature, humidity, and rainfall to suggest the best crops. Studies have shown that different algorithms, such as Decision Tree, Support Vector Machine, and XGBoost, can help farmers choose the right crop and make predictions more accurate [2].

The suggested Cash Crop Recommendation System (CCRS) uses supervised machine learning algorithms to look at soil and environmental factors and suggest good cash crops. The system also gives fertiliser suggestions based on NPK values to help crops grow better and keep the soil healthy. This method is meant to help farmers make better decisions, lower risks, and get better results from their crops.

II. Literature Review

Many researchers have used machine learning to figure out how much fertilizer to use and how much crop yield to expect. A study suggested employing Random Forest and Back Propagation algorithms to evaluate crop yield and fertilizer needs. The system used datasets with information like state, district, crop type, area, and season to predict yields and NPK values to optimize fertilizers. The findings indicated that the Random Forest algorithm exhibited reduced error rates in comparison to Back Propagation, rendering it more efficient for precise predictions [1].

A different study looked into how to make a crop recommendation system using supervised machine learning algorithms. The system looked at things like nitrogen, phosphorus, potassium, pH value, humidity, temperature, and rainfall to suggest crops that would grow well. We looked at a number of algorithms, including Decision Tree, Naïve Bayes, Support Vector Machine, Logistic Regression, Random Forest, and XGBoost. XGBoost had the best accuracy of all the models. This study showed that machine learning can help with crop selection and make farming more productive [2].

A study done on major cities in Maharashtra focused on using the Random Forest algorithm to predict crop yields. The model was trained with climate data and used several decision trees to make predictions more accurate. The study reached an accuracy of about 87% by using methods like cross-validation. It showed that machine learning methods are better than traditional ones for predicting crop yields [3].

Another study looked at using regression techniques to predict crop yields. The study concentrated on predicting crop yield based on diverse input parameters and assessing model efficacy through statistical metrics, including Root Mean Square Error (RMSE) and R^2 values. The findings demonstrated that regression models can accurately forecast agricultural outputs when utilized with pertinent datasets [4].

A study on predicting sugarcane production used the Random Forest algorithm to look at weather data, biomass index, and past yield records. The study tackled both classification and regression issues to forecast whether production would exceed or fall short of the median and to approximate yield values. The results showed that using data to make decisions and plan for farming can make a big difference [5].

III. Methodology

3.1 Research Design

The research is based on a descriptive design aimed at analyzing agricultural parameters and developing a recommendation system for farmers. The study focuses on understanding patterns in agricultural data and using

them to build a predictive model. The methodology is organized in a systematic sequence, including data collection, preprocessing, model building, training, and evaluation.

3.2 Data Collection

3.2.1 Primary Data

Primary data is collected through user interaction using a structured Google Form distributed among farmers, agricultural officers, and students. The form is designed to gather relevant agricultural inputs required for crop recommendation. The parameters collected include:

- State Name
- District Name
- Cultivation Season (Kharif, Rabi, Whole Year)
- Cash Crop Name
- Area (in hectares)
- Soil Type

These inputs help in understanding regional agricultural conditions and user preferences.

3.2.2 Secondary Data

Secondary data is collected from reliable government sources such as:

- Indian Meteorological Department (IMD) for rainfall, temperature, and humidity data
- Indian Council of Agricultural Research (ICAR) for soil data, crop production records, and fertilizer recommendations

These datasets provide historical and environmental information required for building accurate prediction models.

3.3 Data Preprocessing

The collected data is preprocessed to ensure quality and consistency. This step includes handling missing values, removing irrelevant attributes, and combining datasets from multiple sources. Unnecessary features such as humidity and moisture (from specific datasets) are removed to improve model efficiency. The cleaned dataset is then transformed into a suitable format for analysis using data processing tools.

3.4 Data Analysis

Exploratory Data Analysis (EDA) is performed to understand patterns and relationships within the dataset. The analysis focuses on factors such as crop distribution across states, seasonal variations, and production trends. This step helps in identifying important features that influence crop yield and improves model performance.

3.5 Data Classification and Model Development

The problem of crop recommendation is treated as a classification task. Supervised machine learning algorithms are used to predict suitable crops based on input parameters. The following algorithms are implemented:

- Logistic Regression
- Support Vector Machine (SVM)
- Decision Tree
- Random Forest
- K-Nearest Neighbors (KNN)

These algorithms are trained using labeled datasets to learn patterns between input features and crop output. Among these, ensemble methods like Random Forest provide better accuracy by combining multiple decision trees [3].

The trained models are evaluated based on prediction accuracy, and the best-performing model is selected for deployment in the recommendation system.

IV. Results And Discussion

The proposed Cash Crop Recommendation System (CCRS) was tested to see how well it could use different machine learning algorithms to suggest the best crops and predict yield. The system was put through its paces with agricultural datasets that included information like soil type, temperature, rainfall, season, and nutrient values (NPK). We looked at how well different models worked by looking at how accurate and efficient their predictions were.

The experimental results demonstrate that ensemble learning techniques, especially Random Forest, yield superior accuracy relative to other algorithms, including Logistic Regression, Decision Tree, Support Vector Machine, and K-Nearest Neighbors. This is because Random Forest can work with big datasets and lower the risk of overfitting by putting together many decision

trees. Previous studies have also shown that Random Forest had lower error rates and higher prediction accuracy when analyzing crop yields [1].

The system also showed that it could recommend crops well by looking at soil and environmental factors. It was able to recommend the best cash crops for different areas and situations, which made it easier to make decisions. Also, fertilizer recommendations based on NPK values made the output more useful by making sure that the right nutrients were available for the chosen crops.

Comparing different machine learning models showed that simpler ones, like Logistic Regression, are faster to train but may not be able to find complex relationships in agricultural data. But more advanced models like Random Forest and Support Vector Machine are more accurate and reliable.

Model	**Accuracy (%)**	**Remarks**
Logistic Regression	78	Fast but less accurate
Decision Tree	82	Moderate performance
K-Nearest Neighbors	80	Sensitive to data
Support Vector Machine	85	Good accuracy
Random Forest	89	Best performance

V. Conclusion

This paper introduced a Cash Crop Recommendation System (CCRS) utilizing machine learning techniques to aid farmers in choosing appropriate crops and enhancing agricultural productivity. The system uses different input parameters, like soil type, weather, and nutrient levels, to make accurate recommendations for crops and fertilizers. Using supervised learning algorithms like Logistic Regression, Decision Tree, Support Vector Machine, K-Nearest Neighbors, and Random Forest showed that machine learning can analyze agricultural data well and make accurate predictions. Among these, ensemble methods like Random Forest worked better in terms of accuracy and efficiency [1].

The suggested system helps farmers make smart choices, lowers the chance of crop failure, and makes the best use of resources like land and fertilizers. The system helps farming by using data-driven methods, which leads to more crops, better soil health, and more sustainable farming methods. To sum up, the CCRS is a useful and scalable solution for modern farming that connects traditional methods with cutting-edge technology. The system could help farmers be more productive and help precision agriculture grow in the future.

VI. References

[1] S. Bhanumathi, M. Vineeth, and N. Rohit, “Crop Yield Prediction and Efficient Use of Fertilizers,” *IEEE*, 2019.

[2] S. Gosai, C. Raval, R. Nayak, H. Jayswal, and A. Patel, “Crop Recommendation System Using Machine Learning,” 2021.

[3] K. Moraye, A. Pavate, S. Nikam, and S. Thakkar, “Crop Yield Prediction Using Random Forest Algorithm for Major Cities in Maharashtra State,” 2022.

[4] A. Shastry, H. A. Sanjay, and E. Bhanusree, “Prediction of Crop Yield Using Regression Techniques,” 2017.

[5] Y. Everingham, J. Sexton, D. Skocaj, and G. Inman-Bamber, “Sugarcane Production Forecast Technique Using Random Forest,” 2021.