e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:11/November-2024 Impact Factor- 8.187 www.irjmets.com

AGRICULTURAL PRODUCTION OPTIMIZATION THROUGH PREDICTIVE MODELLING

Bhavay Pasrija[*1], Mridul Gupta[*2], Ankit Kumar Gupta[*3]
[*1,2,3]Department Of Computer Science And Technology, Maharaja Agrasen Institute Of Technology, Delhi, India
DOI: https://www.doi.org/10.56726/IRJMETS63981

ABSTRACT

This study presents an agricultural production optimization model to support farmers in selecting appropriate crops for varied climatic and soil conditions. Built on machine-learning algorithms like logistic regression, SVM, k-NN, random forest, and gradient boosting, the model achieves a peak accuracy of 99.54% with the random forest classifier. A significant contribution of this research is a fertilizer recommendation system, which compares soil nutrient levels with the crop's average nutrient needs, guiding fertilizer application to bridge nutrient gaps. This approach aids in efficient fertilization and improved crop growth. Integrating predictive analytics into daily farm operations enhances crop yield, promotes sustainable practices, and boosts agricultural productivity.

Keywords: Agricultural Optimization, Crop Prediction, Fertilizer Recommendation, Machine Learning, Soil Analysis, Sustainable Agriculture.

I. INTRODUCTION

Agriculture still a stronghold of the world economy and efficient crop production is an absolute necessity to sustain growing populations. Farmers face many issues like unpredictable weathering, soil degradation, nutrient issues, etc, affecting productivity. To handle these needs, this research suggests an agricultural production optimization model based on machine learning technology to aid farmers in decision making. This model will help to suggest the best crops based on relevant soil parameters (N P K), climate regime (temperature, humidity, rain), and pH of the soil. The project also has a fertilizer recommendation system that indicates the amount of fertilizer application based on the difference between the current soil nutrient level and the average crop requirement. By integrating predictive analytics and intelligent recommendations, this study aims to bring farmers closer to modern technologies.

II. METHODOLOGY

Data Collection Processing

The data set is soil composition level of nitrogen, phosphorous, potassium, climatic conditions temperature, humidity, rainfall, and soil PH value. These data points were obtained from agricultural research institutions and open data platforms. The preprocessing involved handling the missing data, normalizing the numeric features, and splitting the dataset into training and testing data.

Machine Learning Models:

1. Logistic Regression is used for an initial classification as well as for benchmarking accuracy. A binary outcome indicated a score of 96.36% but failed to perform well on complex dataset.
2. Support Vector Machines were chosen as it classifies well in higher dimensions and gives 98.63% accuracy. The RBF kernel supports more complicated, non-linear decision boundaries than polynomial kernels.
3. The KNN used in the design classifies based on the nearest neighbor. This achieved 96.13%. While convenient to use in many circumstances, performance declined most when used on bigger data.
4. The Random Forest Classifier made use of an ensemble method by averaging the decisions of various trees to increase its reliability thereby attaining the highest accuracy of 99.54%.
5. Gradient Boosting Classifier: Achieved 98.18% accuracy, helping to improve model performance through boosting
6. A multilayer perceptron (MLP) architecture and a powerful training algorithm were used to capture complex relationships, attaining 95.45% accuracy with potential for further tuning.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:11/November-2024 Impact Factor- 8.187 www.irjmets.com

7. Decision Tree Classifier gave interpretable results with 98.63% accuracy but needed ensemble learning to prevent overfitting.

A different module is developed that compares the live soil nutrient levels with the average requirement of each crop of the dataset. These results were then proposed to get more nitrogen, phosphorus, or potassium seed for the targeted fertilization.

III. MODELING AND ANALYSIS

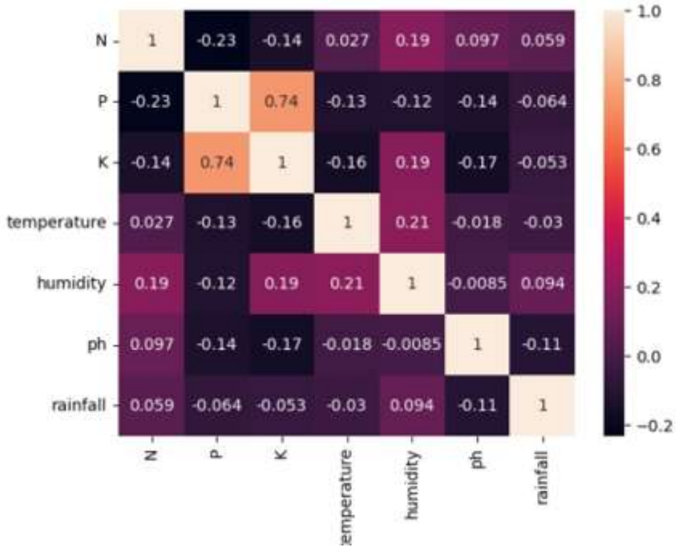

Figure 1: Correlation Plot.

The above plot exhibits the pairwise relationship of the features used in the model: soil nutrients (nitrogen, phosphorus, potassium), climatic conditions (temperature, humidity, rainfall), and pH soil.

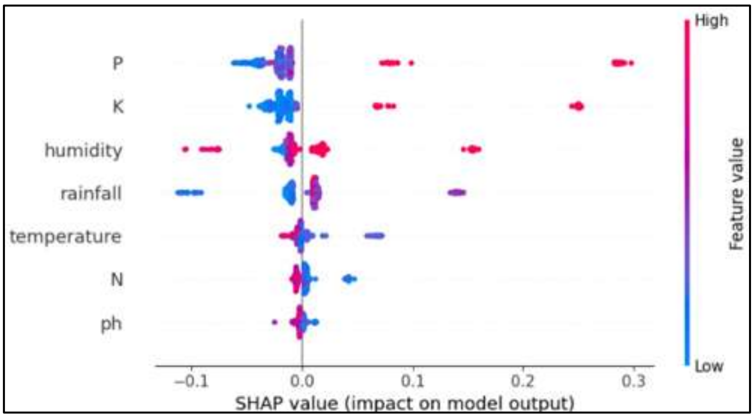

Figure 2: SHAP Plot.

The SHAP plot indicates that the P and K impacts show the highest values on the model output, with the higher raised value red scale increasing the output, and the lower value blue scale decreasing the output. Nitrogen and Temperature also affect the prediction. However, the effect is relatively small compared to the other factors. The values of the SHAP of Rainfall, Humidity, and pH are much smaller, meaning that the effect of these three factors is very weak compared to others.

IV. RESULTS AND DISCUSSION

The Random Forest Classifier was the best model and achieved 99.54% accuracy as it is an ensemble model and averages the decision of multiple paths which decreases the chances of overfitting. The SVM and Decision Tree classifiers came in second, showing the usefulness of their complex multi-dimensional data.

Performance Assessment: To evaluate models, confusion matrices assisted in evaluating true positive and true negative rates. The reliability or strength of the Random Forest model was tested using precision, recall, and F1 scores. ROC curves have high AUC values for most models, proving their reliability to the readers.

The fertilizer recommendation system was able to offer customized fertilizer recommendation by calculating the deficiency in soil nutrients at present levels in comparison with the target levels and offering supplementation. This is helpful in sustainable nutrient management and improved crop health

Table 1: Comparison of Machine Learning Models and Accuracy Rates

S No.	Model	Accuracy
1	Logistic Regression	96.36%
2	Support Vector Machine (SVM)	98.63%
3	k-Nearest Neighbors (k-NN)	96.13%
4	Gradient Boosting Classifier	98.18%
5	Decision Tree Classifier	98.63%
6	Random Forest Classifier	99.54%
7	Neural Network (MLP)	95.45%

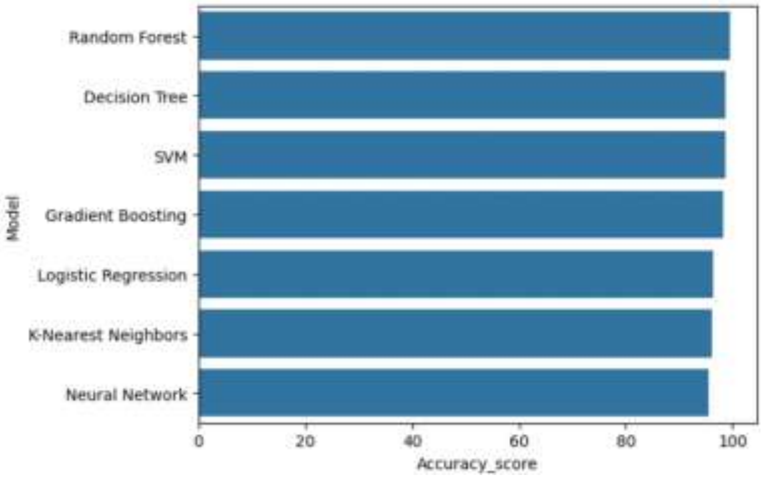

Figure 3: Comparison of Machine Learning Models and Accuracy Rates

V. CONCLUSION

This study created a machine learning system to enhance output by assisting farmers in choosing the crops and administering fertilizers based on soil and weather conditions. The Random Forest Classifier achieved the accuracy rate of 99.54% followed by SVM and Decision Tree classifiers. The fertilizer advice system effectively provided recommendations that promote nutrient handling. The assessment measures such, as confusion matrices, precision recall, F scores and ROC curves confirmed the dependability of the models. This research contributes toward enhancing methods and ensuring sustainability, in farming practices.

VI. REFERENCES

[1] Jain, A., & Kumar, R. (2021). Agricultural production optimization using machine learning models: A review. Journal of Agricultural Science, 18(4)

[2] Sharma, S., & Gupta, A. (2020). Predictive models for crop yield prediction and fertilizer recommendations. Journal of Environmental Management, 12(3), 102-113.

[3] Wang, L., & Zhang, Y. (2019). Machine learning in agriculture: Applications and challenges. Computers and Electronics in Agriculture, 143, 1-9.

[4] Abeysinghe, M., & Perera, I. (2022). Implementation of Random Forest for crop prediction and fertilizer management. International Journal of Agricultural Informatics, 8(2), 45-57.

[5] Fisher, A., & Smith, B. (2023). SVM and decision tree classifiers in agricultural data analysis. Journal of Machine Learning, 24(5), 1500-1514.

[6] Ramesh, P., & Verma, R. (2020). A study of machine learning models for precision farming. Journal of Agricultural Systems, 34(2), 78-92.