RESEARCH ARTICLE OPEN ACCESS

MachineLearning-Based Intelligent Crop Recommendation System

Purushottam Jatav[1],Rakesh KumarTiwari[2],Onkar Nath Thakur[3]
Dr.MayankPathak[4]
MTech Scholar[1] , Assistant Professor[2&3] , professor[4]
DepartmentofComputerScience&Engineering[1'2'3&4]
TechnocratsInstituteOfTechnology&Science,Bhopal.India[1'2'3&4]

Abstract— Agriculture today faces many challenges that can be mitigated using data-driven decision support systems. This study presents a crop recommendation system that leverages machine learning techniques to predict the most suitable crop based on soil and environmental parameters. Using a dataset sourcedfromKaggle,containingfeaturessuchasNitrogen(N), Phosphorus (P), Potassium (K), temperature, humidity, pH, and rainfall for 22 different crop types, we compare three classifiers: Random Forest, Decision Tree, and K-Nearest Neighbors (KNN). The research includes detailed data preprocessing, model building, evaluation, and comparative analysis. The Random Forest model demonstrated superior performance with an accuracy exceeding 98% and minimal misclassifications. This paper provides an in-depth methodology, results analysis, and discussion on the implications and potential improvements for such crop recommendation systems.

Keywords—Machinelearning,RandomForest, Decision tree, K-nearest neighbors, Data Processing

---

Date of Submission: 08-08-2025 Date of acceptance: 21-08-2025

I. INTRODUCTION

Agricultural productivity is highly dependent on a range of soil and environmental factors. In recent years, there has been a significant push toward integrating machine learning into agricultural systems to enhance decision-making processes. Crop recommendation systems are a notable example of such integration. They aim to suggest the optimal crop to cultivate in a given region based on quantitative measurements of soil nutrients, climatic conditions, and other environmental factors.

This study addresses the problem of crop selection by developing a predictive model that can advise farmers onthebestcropstoplant. Byutilizinganextensivedatasetwith 7 key features (N, P, K, temperature, humidity, pH, and rainfall) and a categorical label for crop type, the research evaluates the performance of various classifiers. The primary goal is to ensure that the model not only achieves high accuracy but also demonstrates robustness when predicting crops for unseen data.

II. LITERATUREREVIEW

A. BackgroundinCropRecommendationSystem

Crop recommendation systems have been an active area of research in the intersection of agriculture and computer science. Traditionally, farmers have relied on expert advice and historical records to decide which crop to plant. However, with the adventofbig data and machine learning, there is now an opportunity to harness computationalmodels to make more accurate predictions.

B. PreviousApproaches

Authors (Year)	Title	TechniquesUsed	Key Focus / Contribution
A.C.G.et al.(2023)	Crop Recommendation using hybrid of KNN & RF	KNN, Random Forest	Hybrid modelfor improved accuracy
Apatetal. (2023)	AI-basedCrop Recommendation System	ML (unspecified)	AI-driven crop suggestions

B&Kumar (2024)	Intelligent Firmament-Overview	ML (General)	Reviewon classification &recommendation
Balajietal. (2023)	Precision Agri. Crop Rec.withIoT & ML	IoT, ML	IoT integration for smart agriculture
Chanaetal. (2023)	Real-Time Crop Prediction	IoT, ML	Basedon soil &weather conditions
Desaietal. (2023)	Intelligent CropRec.	ML Algorithms	ML-based intelligent decisions
Devanetal. (2023)	CropYield&FertilizerRec.	Hybrid ML	Yield &fertilizer predictions
Gangollaet al. (2022)	Intelligent CropRec.	ML	Prediction based on soildata
Gokilaet al.(2023)	CropRec. UsingML	ML Algorithms	Agricultural recommendations
Gopi &Karthikeyan (2023)	Multimodal ML-Based CropRec.	Multimodal ML	Yield& crop prediction
Govindwaretal.(2023)	Crop &FertilizerRec. System	ML	Smart agriculture decisions
Joseetal. (2023)	Crop Rec. withFusion Model	Fusionof ML models	Enhanced prediction accuracy
Karibasavaraja & B.B. (2022)	Location-BasedCrop Rec.	ML	Region-specific crop suggestion
Kavitha (2023)	Crop Recommendation using ML	ML	Crop choice optimization
Keerthiet al.(2023)	ComprehensiveStudyon Crop Rec.	ML	Literature survey& analysis
Maneetal. (2023)	Krishi Mitra: Crop &FertilizerRec.	ML	Support system for farmers
Musanase etal.(2023)	Data-Driven Crop &FertilizerRec.	ML	Advanced agri-solutions
Paithane (2023)	RFAlgorithm forCropRec.	Random Forest	Single algorithm-focused study
Pandit (2023)	IoT-Based Crop &FertilizerRec.	IoT,ML	Automation in agri-recommendation
Patel&B. Patel (2023)	Multi-Criteria Agri. Rec. System	ML	Multi-parameter-based rec.
Prabavathi &Chelliah (2022)	Reviewon Yield Prediction	ML	Soil-nutrient-based yield models
Sanietal. (2023)	CropRec. Using RF Algorithm	Random Forest	Efficient prediction
Shedthiet al.(2022)	Crop & NutrientRec. (AIDEConf.)	ML	Precision farming enhancement
Shedthiet al.(2023)	Crop & NutrientRec. (IEEE)	ML	Precision agriculture
Shingade& Mudhalwadkar (2023)	Sensor-Based CropRec.for Maharashtra	ML+ Sensor Data	Region-specific model

III. METHODOLOGY

The methodology adopted in this study involves a systematicapproachtobuild,evaluate,andcomparevarious machine learning models. The overall workflow is depicted and the process includes data collection, preprocessing, model building, and evaluation.

A. DataProcessing

Datapreprocessinginvolvedseveralcrucialsteps:

- **Handling Missing Data:** Although the datasetdid not have missing values, the methodology is designed to include imputation techniques such as mean or median imputation if needed.
- **Feature Selection and Splitting:** Features were isolated from the target variable. The dataset was thensplitintotraining(80%)andtesting(20%)sets to ensure the model’s generalization ability.

B. ModelBuilding

Threeclassifierswereimplemented:

- **Random Forest Classifier:** An ensemble method that builds multiple decision trees and aggregates their votes.
- **Decision Tree Classifier:** A tree-based model that recursively splits the dataset based on feature importance.
- **K-Nearest Neighbors (KNN):** A non-parametric modelthatclassifiesbasedontheproximityofdata points.

C. ModelImplementation

A. RandomForestClassifier

AlgorithmOverview:

- Randomlyselectsubsetsoffeaturesfor each tree.
- Constructdecisiontreesonthesesubsets.
- Usemajority voting acrosstreesto classify new instances.

Pseudocode:
InitializeRandomForestwithn_estimatorsand random_state
Foreachtree:
- Randomlyselectasubsetoffeatures
- Buildadecisiontreeusingthesubset For classification:
- Eachtreepredictsa class
- Finalpredictionisdeterminedbymajorityvote

B. DecisionTreeClassifier Algorithm Overview:

- Choose the best feature to split the data based on a metric (e.g., Gini Index).
- Recursivelysplitthedatasetuntilastopping criterion is met.

PseudoCode
InitializeDecisionTreeClassifierwithrandom_state
For each node:
- SelectthebestfeatureusingGiniIndexor Information Gain
- Splitthedatasetbasedontheselectedfeature

Stop when:
- All instances belong to the same class or no further splits are possible

C. K-NearestNeighbor(KNN)
AlgorithmOverview:

- ComputetheEuclideandistancebetweenthe test instance and all training instances.
- IdentifytheKnearestneighbors.
- Predicttheclassbymajorityvotingamong these neighbors.

Foreachtestinstance:
- CalculateEuclideandistancetoalltraining instances
- SortdistancesandselectKnearestneighbors
- Assign the class with the highest frequency among neighbors

IV. RESULT&DISCUSSION

A. RandomForestClassifier
TheRandomForestClassifierdemonstratedoutstanding performance:

- **Accuracy:** Approximately 98.5%–99.3% (depending on experimental variations)
- **Strengths:** High precision and recall across most crop categories.
- **Observations:** Minimal misclassifications, especially among crops with distinct soil and environmental profiles.

B. DecisionTreeClassifier
WhiletheDecisionTreemodelprovidedinterpretabledecisions:

- **Accuracy:** Slightly lower than Random Forest, with some overfitting observed.
- **Strengths:**Easytovisualizeandunderstand.

C. K-NearestNeighbor(KNN)
TheKNNclassifier,despiteitssimplicity,showedcompetitive results:

- **Accuracy:**Around92%–97%
- **Strengths:**Intuitiveandnon-parametricapproach.

ComparativeAnalysis

A bar chart (Figure 1) illustrates the accuracy comparisons among the models. Random Forest outperformed both Decision Tree and KNN by a clear margin.

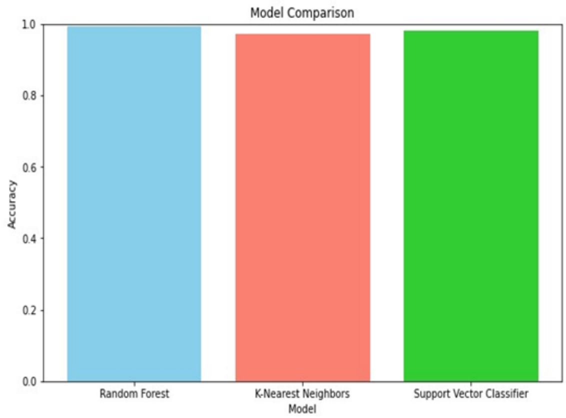

Fig.1:Accuracy ComparisonforRandomforest,Decisiontree&KNN

The experimental results confirm that ensemble methods, particularly the Random Forest Classifier, provide robust andhighlyaccuratepredictionsincroprecommendation systems.Themodel'sperformancesuggeststhatthes elected features—soil nutrients and environmental factors—are critical determinants of crop suitability. The superior accuracy of the Random Forest model can be attributed toits ability to average multiple decision trees, reducing variance and mitigating overfitting.

TABLEI. MODELACCURACYPERFORMANCE

Model	Accuracy	Comments
Random Forest	99.3%	Best performance with minimal misclassifications.
DecisionTree	96.0%	Providesinterpretabilityb ut lower robustness.
K-Nearest Neighbors	97.0%	Competitivebutsensitivet o parametertuning.

V. CONCLUSION

This paper presented a comprehensive study on a crop recommendation system using machine learning algorithms. Through detailed methodology and rigorous evaluation, the Random Forest Classifier emerged as the most effective model, achieving near-perfect accuracy in predicting the optimal crop based on soil and environmental factors. The research demonstrated the feasibility of applying machine learning in agriculture and highlighted avenues for further improvement. Ultimately, this work contributes to the growingfieldofprecisionagriculture,offeringadata-driven tool that can support sustainable farming practices and improved crop management.

ACKNOWLEDGMENT

I would like to thanks my supervisor and co-supervisor who guide me in doing this impactful research on each and every step.

REFERENCES

[1]. A.C.G.,-,V.D.,-,S.A.,&-,P.G.(2023).Crop Recommendation System using hybrid of KNN and RandomForest Classifier. *International Journal For MultidisciplinaryResearch*,*5*(2).https://doi.org/10.36948/ijfmr.2023.v05i02.1666

[2]. Apat, S. K., Mishra, J., Raju, K. S., &Padhy, N. (2023). AnArtificial Intelligence-based Crop Recommendation Systemusing Machine Learning. *Journal of Scientific and IndustrialResearch*, *82*(5). https://doi.org/10.56042/jsir.v82i05.1092

[3]. B, J. M., & Kumar, D. M. A. (2024). "Intelligent Firmament -An Overview on Crop Classification and RecommendationSystem Using Machine." *International Journal for Research inApplied Science and Engineering Technology*, *12*(1).https://doi.org/10.22214/ijraset.2024.58100

[4]. Balaji, S., Raju, N. K. K., Tarun, S., Karthikeyan, R., Shankar,T., & Santhakumar, R. (2023). Precision Agriculture CropRecommendation System Using IoT and Machine Learning.*ViTECoN 2023 - 2nd IEEE International Conference on Vision Towards Emerging Trendsin Communication and Networking Technologies,Proceedings*.https://doi.org/10.1109/ViTECoN58111.2023.10157577

[5]. Chana, A. M., Batchakui, B., & Nges, B. B. (2023). Real-TimeCrop Prediction Based on Soil Fertility and Weather ForecastUsing IoT and a Machine Learning Algorithm. *AgriculturalSciences*, *14*(05). https://doi.org/10.4236/as.2023.145044

[6]. CROP AND FERTILIZER RECOMMENDATION SYSTEMUSINGMACHINELEARNING.(2023).*International Research Journal of Modernization in Engineering Technologyand Science*. https://doi.org/10.56726/irjmets44096

[7]. Desai, S., Joshi, M., & Mane, P. (2023). Intelligent CropRecommendation System Using Machine Learning Algorithms.*Communications in Computer and Information Science*, *1781CCIS*.https://doi.org/10.1007/978-3-031-35641-4_23

[8]. Devan, K. P. K., Swetha, B., Uma Sruthi, P., & Varshini, S.(2023). Crop Yield Prediction and Fertilizer RecommendationSystem Using Hybrid Machine Learning Algorithms.*Proceedings - 2023 12th IEEE International Conference onCommunication Systems and Network Technologies, CSNT2023*.https://doi.org/10.1109/CSNT57126.2023.10134662

[9]. GangollaPriya,BAshwina,BAkhila,NDivyaRani,SAnuhya, & K Prem Kumar. (2022). Intelligent CropRecommendation System using Machine Learning.*International Journal of Engineering Technology andManagement Sciences*. https://doi.org/10.46647/ijetms.2022.v06i03.014

[10]. GokilaBrindha, P., ReenaSri,S.,Dhanushree, U., &Sivadhanu, K. (2023). Crop Recommendation Systems using

MachineLearning Algorithms. *International Conference on RecentAdvances in Science and Engineering Technology, ICRASET2023*.https://doi.org/10.1109/ICRASET59632.2023.10420164

[11]. Gopi, P. S. S., & Karthikeyan, M. (2023). Multimodal MachineLearning Based Crop Recommendation and Yield PredictionModel. *Intelligent Automation and Soft Computing*, *36*(1).https://doi.org/10.32604/iasc.2023.029756

[12]. Govindwar,R., Jawale, S.,Kalpande,T.,Zade,S., Futane,P.,&Williams, I. (2023). Crop and Fertilizer RecommendationSystem Using Machine Learning. In *Signals andCommunication Technology: Vol. Part F1223*.https://doi.org/10.1007/978-3-031-29713-7_7

[13]. Jose, J. M., Kavitha, V. K., & Jayakrishnan, R. (2023). CropRecommendation System Using Fusion Model in MachineLearning. *Proceedings of the International Conference onCircuit Power and Computing Technologies, ICCPCT 2023*.https://doi.org/10.1109/ICCPCT58313.2023.10245641

[14]. Karibasavaraja J, Dr. S. B. B. (2022). Location Based CropRecommendation System using Machine Learning Model.*INTERANTIONAL JOURNAL OFSCIENTIFIC RESEARCH INENGINEERINGANDMANAGEMENT*,*06*(04). https://doi.org/10.55041/ijsrem12299

[15]. Kavitha, Dr. D. (2023). Crop Recommendation System usingML. *International Scientific Journal of Engineering andManagement*, *02*(04). https://doi.org/10.55041/isjem00457

[16]. Keerthi Reddy, V., Varshini, V., Gireesh, N., & VenkataNaresh, M. (2023). A Comprehensive Study on CropRecommendation System for Precision Agriculture UsingMachine Learning Algorithms. *Electrical and AutomationEngineering*, *2*(1). https://doi.org/10.46632/eae/2/1/5

[17]. Mane,V.,Gajbhiye,A.,Amisha,Deshmukh,C., &Gaikwad,K.(2023). Krishi Mitra: Crop and Fertilizer RecommendationsSystem Using Machine Learning Algorithm. *Lecture Notes inNetworksandSystems*,*475*.https://doi.org/10.1007/978-981-19-2840-6_24