International Journal on Advanced Computer Theory and Engineering
ISSN: 2319 - 2526
Volume 15 Issue 2 (2026) | Pages: 136-141

Analysing Bias and Reliability in Soil – Based Crop Recommendation Systems

Khushboo Jain[1], Kanchan Dewangan[2], Anjali Sahu[3], Yogesh Kumar Rathore[4]

[1,2,3,4]Department of Computer Science and Engineering
Shri Shankaracharya Institute of Professional Management and Technology
Raipur, India

Peer Review Information

***Type**: Article*
***Received**:23 April 2026*
***Revised**: 24 May 2026*
***Accepted**: 22 June 2026*
***Published**: 20 July 2026*

Abstract

Machine Learning (ML) is now being used for the development of soil-based crop recommendations that support agricultural decision-making. Many of these systems claim to have very high predictive accuracy, however, very little attention has been paid to the effect of data bias and the quality of recommendations. This research study investigates the influence of data bias and how it relates to the reliability of the recommendations when evaluating the performance of ML models. The research utilizes a soil health data collection consisting of nutrient levels, moisture, humidity, temperature, and crop labels to conduct controlled experiments on the relationship between the features and accuracy/stability of predictions by multiple ML models. The experiments were performed by conducting several controlled experiments using the soil health data collection and tracking the effects of different variations of crops selected for training/testing the models, as well as the number of features being used to create the models. The results show that accuracy ratings can vary considerably between datasets as well as dominant features, suggesting potential bias in how models behave. Therefore, it is necessary to conduct comprehensive dataset analysis and create and use evaluation methods that consider the accuracy of models when making recommendations.

Keywords: AI/ML; Crop Recommendations; Soil Health; Data Bias; Model Reliability; Ensemble Model

How to Cite This Article

Jain, K., Dewangan, K., Sahu, A., & Rathore, Y. K. (2026). Analysing bias and reliability in soil-based crop recommendation systems. *International Journal on Advanced Computer Theory and Engineering,* 15(2), 136–141.

Introduction

With the increase in world population, agriculture plays a fundamental role as a pillar for global food security and economic stability, supporting the livelihoods of many people. However, agricultural sustainability is increasingly threatened by soil degradation, climate variability and inefficient resource management practices [1]. In crop productivity, soil health plays an important role as parameters such as nutrient availability, which directly influence plant growth and yield [2]. In agricultural productivity declination some practices have contributed significantly, such as poor soil management practices, including inappropriate crop selection, excessive fertiliser application and improper irrigation across many regions.

In agricultural sector AL/ML techniques have been widely used and adopted to analyse complex, multidimensional datasets and supportive data-driven decision making and support data-driven decision making [4]. To improve the agricultural planning, a machine learning based soil and recommendation system is used, and that system utilises soil and environmental parameters such as nitrogen. Phosphorus, potassium, moisture, temperature and humidity to predict suitable crops [5],[6].

There are several studies that illustrate that an ML-driven approach can outperform traditional advisory methods by capturing non-linear relationships among soil attributes and climatic variables [5],[7]. These advancements highlight the growth of intelligent decision support systems in precision and smart farming.

Even after all these developments, a critical limitation of existing soil-based crop recommendation systems is the limited examination of dataset bias and recommendation reliability. The main focus of all those studies is on predictive accuracy while overlooking the characteristics of the dataset, such as feature dominance, constrained value ranges, and selective crop inclusion, these are those which affect the behaviour of the model and evolution outcome [8], [6].

As a result, the high accuracy that was reported under controlled experimental conditions may not reliably translate to robust performance in real-world agricultural settings, which raises concern regarding the generalizability and trustworthiness of such systems [8], [9].

Addressing this existing gap, this paper investigates bias and reliability in soil-based crop recommendation systems using machine learning through an experimental analysis framework. Rather than focusing on producing a new model, this study focuses on understanding how dataset composition and feature influence model accuracy and prediction stability. There are various ML models that are evaluated under consistent experimental settings using soil health data, and in the dataset configuration, there are variations that are analysed to reveal their effect on performance. This method highlights the limitations of accuracy evaluation and highlights the need for reliability-aware assessments in agricultural decision support systems.

The remainder of the paper is structured as follows. In Section II, we survey recent literature in the area of soil-based crop recommendation systems with specific emphasis placed upon feature impact and dataset-related issues. Section III introduces the experimental methodology with specific regard to the dataset description/preprocessing and the model training process. Section IV discusses the experimental results with specific regard to the impact of bias upon the model's performance/reliability. Section V concludes the paper with specific regard to the directions taken in the future.

Related Works

The majority of the research on crop recommendation based on soil has focused on the application of machine learning in recognizing the crop based on the nutrients in the soil and other environmental factors. This body of research has proven the capability of supervised machine learning in effectively recognizing the complex relationship between the soil and the crop. For instance, in the research by Prity et al. [5], the authors proposed a crop recommendation system based on the application of different machine learning classifiers. Their research proved the effectiveness of the system over other crop recommendation methods. Afzal et al. [6] also proposed a crop recognition system using the features of the soil and machine learning, proving the effectiveness of nonlinear methods in crop recognition.

In recent years, there has been a progressive trend towards data-driven soil analysis and smart farming applications. In order to enhance agricultural decision-making, Huang et al. [7] developed a framework for evaluating soil and stressed the importance of meticulous preprocessing and consistent feature representation in obtaining trustworthy findings. Cheema et al. [3] produced crop and nutrient recommendations by combining machine learning models with Internet of Things-based soil monitoring technologies. Their research brought to light a number of real-world deployment-related practical issues, such as data noise, sensor variability, and inconsistencies among soil samples.

Few studies have attempted to delve into the impact of dataset composition on model reliability, while results in this area are highly positive. Dey et al [8]. focused on evaluating the performance of crop recommendation based on varying soil nutrient levels and climate conditions. Their conclusion is that accuracy is highly dependent on what can or cannot be shown within a dataset and on glaring aspects. In other words, while excellent results are reported within controlled experiments, these may not be applicable in other agricultural fields. According to

Kamilaris and Prenafeta-Boldú [9], data quality, as well as how relevant features are and the dataset itself, play a significant role in how agricultural machine learning models perform and how reliable predictions are. However, if feature representation is poor, then generalization and prediction may be compromised.

Throughout the literature, there has been a lot of concern with enhancing the accuracy of predictions, but the bias in the dataset, features, and reliability have been neglected. Most research has been conducted by using a fixed setup for the dataset and has not sufficiently investigated the impact of the change in the class balance, features, and crop representation on the system. This research aims to fill the gap by investigating the bias and reliability in the crop recommendation system based on the soil using experiment.

Methodology

This research adopts an experiment design method for investigating biases and reliability of crop recommendations based on machine learning in soil health. The process follows a sequential pipeline starting from preprocessing of soil data through an identical process. Afterwards, different data sets are created through restricting the classes of crops and features. This is done under similar conditions in terms of learning algorithm and experiments. Then, several datasets are generated by imposing different restrictions on the crops and features while maintaining the same learning methods and experimentation settings. Machine learning models are trained with identical training and testing samples in all generated datasets to ensure fair comparison. Model evaluations are performed based on different classification metrics, and differences in model results obtained from various datasets are studied to assess the sensitivity and stability of the models. By isolating the impact of dataset composition and feature constraints, the proposed methodology emphasizes reliability aware evaluation rather than accuracy optimisation, enabling a systematic examination of bias effects in soil based crop recommendation systems.

Dataset Descripton and Characterisation

The experiments conducted in this research work are derived from an open-source the soil crop database containing 8,000 entries, with each entry being defined by nine distinct features. Each entry represents a unique combination of soil characteristics and environmental factors associated with a specific crop. In this dataset, the key inputs include temperature and humidity, soil moisture content, soil type, and nutrient content, namely nitrogen (N), phosphorus (P), and potassium (K). These parameters have been known as the most significant features for measuring soil fertility and suitability for crops.

In this case, the crop type is the target variable used in the study. While there is another variable available, related to fertiliser recommendation, it is purposely left out of the data set. This approach guarantees that the test will only concentrate on the functioning and accuracy of the crop prediction model without interference from any other variables. The data set serves as the basis for the experiment. Several different modifications to it are made in order to evaluate how variations in crop selection and input variables affect model performance.

In this research paper, we have done a summary of all possible configurations of our datasets, involving the restrictions on crop categories, applied constraints on features, and sample size, as mentioned in Table 1. For the evaluation of biases and the reliability of data in our models, we followed this configuration after preprocessing the data.

Table 1. *Dataset Configurations Used for Bias and Reliability Analysis*

Configuration ID	Crop Categories Included	Feature Constraints Applied	Number of Samples
C1	All available crops	None (original dataset)	8000
C2	Paddy, Cotton	None	1428
C3	Paddy, Cotton	Moisture ≥ mean (scaled ≥ 0)	635
C4	Paddy, Cotton	Moisture ≥ mean, Humidity ≤ threshold	422
C5	Paddy, Cotton	Relaxed moisture and humidity limits	613

Data Preprocessing

Preprocessing is conducted consistently on the data set before configuring the data set and training the models. Missing values and duplicate records are checked in the data set, and there were no missing values in the dataset, while duplicate records were dropped. Categorical variables such as soil type are converted into numeric variables, and all numeric variables, including temperature, humidity, soil moisture content, and macronutrients N, P, and K, are normalized to standardize feature scaling. Preprocessing settings are kept consistent for all datasets, and any differences in results will be due only to differences in the data sets themselves.

Dataset Configuration & Feature Constraints

In order to examine the effect of data set construction on the accuracy and robustness of the predictive models, various data sets are generated using the pre-processed data set. Instead of adjusting the learning algorithm or the pre-processing technique, the experiment manipulates the dataset by selecting the crops and restricting the features within certain bounds.

The first data set is the entire data set with all the available crops and their features, serving as a baseline for comparison. The following models limit the size of the dataset by choosing some selected crops like paddy and cotton and also impose certain constraints on important environmental factors like soil moisture and humidity. These constraints are selected according to their importance in agriculture and simulate real-life data limitations, like regional crop cultivation and limited crop diversity.

Each and every set described in Table I is formulated after pre-processing and is equally applied to all sets of models. Since the same pre-processing, training-testing split, and model parameter specifications will be used for each set, this particular strategy will ensure that differences in performance between models are due to changes in datasets only. Such a structure becomes the cornerstone for investigating the impact of dominance of features, class limitations, and class imbalance on model performance

Model Training

Three supervised machine learning algorithms, namely Random Forest, Decision Tree, and Support Vector Machine (SVM), are used to assess the effect of dataset configurations on crop predictions accuracy. The choice of algorithms is made because each algorithm has different learning methods and is commonly applied in the field of agriculture to perform classification tasks. All three models are independently trained for all datasets (C1-C5) through an experiment that uses the same parameters and settings.

In all experiments, the dataset will be divided in the same manner for training and testing. Furthermore, no changes in model parameters will be done throughout the experiment to avoid optimizing its performance. Soil and environmental parameters will be used for training only, while crops will be the class labels.

With consistency in pre-processing, data splitting, and training settings, this experiment guarantees that any variance in performance between different models and settings is due to the nature of the data sets themselves.

Model Evaluation

The performances of the models are assessed using conventional metrics for classification, including accuracy, precision, recall, and F1 score. The same evaluation criteria are uniformly applied to all models, and their performances can be fairly compared using a standardized split ratio (train vs. test) for all datasets C1-C5 because they are based on identical evaluation setups.

Accuracy measures the overall correctness of predictions, while precision reflects the reliability of positive predictions by quantifying false positives. Recall indicates the model's ability to correctly identify relevant crop classes, particularly under constrained dataset conditions, and the F1-score provides a balanced measure by combining precision and recall. Instead of focusing on optimizing the model performance, this approach helps us to analyse the sensitivity of performance over different datasets. The metric-driven approach will help us understand the impact of the data itself on the predictive capacity of the model by making sure that all the changes are due to the datasets themselves.

Bias and Reliability Analysis

The analysis of bias and reliability is based on the analysis of the variations of model's performance depending on the data sets' configuration when other parameters, such as the method of processing and training of data sets, remain constant. The objective here is not to determine the performance of the model depending on its capabilities but rather on dataset biases due to crop choice and restrictions on features, among others.

Reliability can be determined based on whether the trends in performance show stability in going from the base setting to more and more restricted datasets. If the changes to the performance of a model are very sensitive to even small changes to the data set, it means that there is less reliability and that the model is heavily influenced by the data set rather than the interaction between the soil and crop types.

The above discussed framework facilitates distinguishing between observed performance improvement due to dataset limitations and actual prediction reliability. The significance of reliability-driven accuracy evaluation is revealed in the paper through the analysis of accuracy metrics within different datasets. The study reveals that reliability-driven accuracy evaluation plays an essential role in the development of crop recommendation models based on soil characteristics.

Results and Discussion

Performance Across Dataset Configurations

The importance of the way datasets (C1-C5) are constructed has been highlighted. The first configuration (C1) featuring eleven crop types sees models performing at random (about 9-10%). Selecting multiple crops based on soil properties is particularly difficult in the given case. The dataset does not demonstrate any type of discrimination. More information can be found in Table II.

Once the dataset size has been decreased to two crops (configuration C2), model performance increases by about 49-52%. The increase has occurred solely due to a change in the structure of the datasets rather than any other factors. In brief, selection of crops affects model performance.

Effect of Feature-Based Constraints

Soil moisture and humidity limitations are imposed on C3-C5 configurations, making it possible to reduce the number of observations even further and stabilize variance. Despite the substantial differences in model behavior, accuracy levels remain confined within a narrow range (approximately 45-59%) across all configurations.

Stricter limitations do not contribute to a gradual increase in accuracy. In fact, slight modifications in the maximum permissible value can lead to significant deviations in results. The dependence of accuracy upon restrictions is illustrated by Fig. 1, which presented accuracy for each dataset configuration and demonstrated its vulnerability to preprocessing methods applied to data.

Reliability and Bias Implications

It would be good to know that there may exist models that are highly accurate despite their difference in robustness and generalizability. If you use data representing only a small part of reality and train and test your model using such data, then the accuracy will be almost at the level of a random guess, which will provide erroneous conclusions during recommendations. In other words, relying on accuracy alone when evaluating the soil crop recommendation system is insufficient; reliability is just as important.

In addition, the structure of data impacts not only the model performance but also its robustness. In the case of agriculture, creating a proper dataset is vital.

Table 1. *Performance Comparison of Machine Learning Models Across Dataset Configurations*

S. No.	Configuration	model	Accuracy	Precision	Recall	F1 - score
0	C1	Random Forest	0.08812	0.08875	0.08812	0.08642
1	C1	Decision Tree	0.09187	0.08611	0.09187	0.08063
2	C1	SVM	0.09625	0.09702	0.09625	0.08900
3	C2	Random Forest	0.52097	0.52063	0.52097	0.52029
4	C2	Decision Tree	0.51049	0.50946	0.51049	0.50429
5	C2	SVM	0.49650	0.49588	0.49650	0.49531
6	C3	Random Forest	0.47244	0.46099	0.47244	0.46023
7	C3	Decision Tree	0.42519	0.40682	0.42519	0.40875
8	C3	SVM	0.52755	0.51481	0.52755	0.49439
9	C4	Random Forest	0.51764	0.49787	0.51764	0.48126
10	C4	Decision Tree	0.43529	0.41419	0.43529	0.41653
11	C4	SVM	0.58823	0.63145	0.58823	0.51573
12	C5	Random Forest	0.55284	0.55171	0.55284	0.55195

13	C5	Decision Tree	0.47967	0.47405	0.47967	0.47371
14	C5	SVM	0.55284	0.55250	0.55284	0.51909

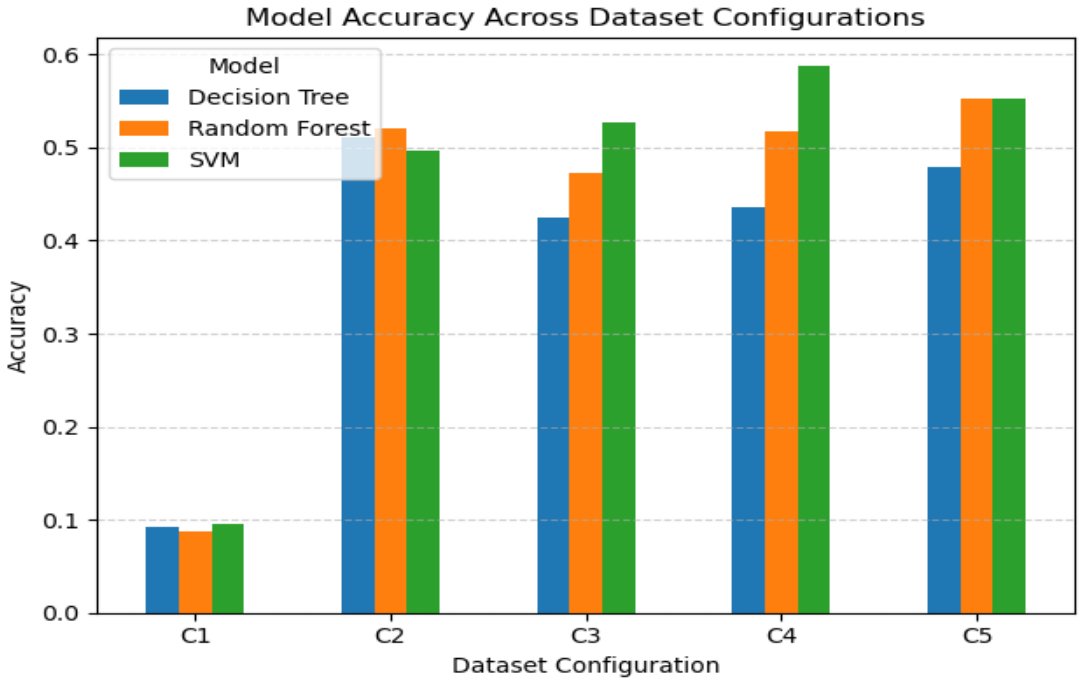

Fig. 1. *Accuracy comparison of Random Forest, Decision Tree, and Support Vector Machine models across dataset configurations (C1–C5)*

References

1. Food and Agriculture Organization of the United Nations, *The State of Food and Agriculture 2022*. Rome, Italy: FAO, 2022. [Online]. Available: https://www.fao.org/publications/sofa/2022/en
2. E. K. Bünemann, G. Bongiorno, Z. Bai, R. E. Creamer, G. B. de Deyn, R. G. M. de Goede, L. Fleskens, V. Geissen, T. W. M. Kuyper, P. Mäder, M. M. Pulleman, W. Sukkel, J. W. van Groenigen, and L. Brussaard, "Soil quality—A critical review," *Soil Biology and Biochemistry*, vol. 120, pp. 105–125, 2018, doi: 10.1016/j.soilbio.2018.01.030.
3. S. M. Cheema and I. M. Pires, "AIoT based soil nutrient analysis and recommendation system for crops using machine learning," *AI Technology*, art. no. 100924, 2025, doi: 10.1016/j.atech.2025.100924.
4. K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, Machine learning in agriculture: A review," *Sensors*, vol. 18, no. 8, art. no. 2674, 2018, doi: 10.3390/s18082674.
5. F. S. Prity, M. M. Hasan, S. H. Saif, M. M. Hossain, S. H. Bhuiyan, M. A. Islam, and M. T. H. Lavlu, "Enhancing agricultural productivity: A machine learning approach to crop recommendations," *Human-Centered Intelligent Systems*, vol. 4, no. 4, pp. 497–510, 2024, doi: 10.1007/s44230-024-00081-3.
6. H. Afzal, M. Amjad, A. Raza, K. Munir, S. G. Villar, L. A. D. Lopez, and I. Ashraf, "Incorporating soil information with machine learning for crop recommendation to improve agricultural output," *Scientific Reports*, vol. 15, art. no. 88676, 2025, doi: 10.1038/s41598-025-88676-z.
7. Y. Huang, R. Srivastava, C. Ngo, J. Gao, J. Wu, and S. Chiao, "Data-driven soil analysis and evaluation for smart farming using machine learning approaches," *Agriculture*, vol. 13, no. 9, art. no. 1777, 2023, doi: 10.3390/agriculture13091777.
8. B. Dey, J. Ferdous, and R. Ahmed, "Machine learning based recommendation of agricultural and horticultural crop farming under NPK, soil pH and climatic variables," *Heliyon*, vol. 10, no. 1, e25112, 2024, doi: 10.1016/j.heliyon.2024.e25112.
9. A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018, doi: 10.1016/j.compag.2018.02.016