iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XII **Month of publication:** December 2023

DOI: https://doi.org/10.22214/ijraset.2023.57337

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XII Dec 2023- Available at www.ijraset.com

Soil Classification and Crop Suggestion Using Machine Learning

Madhuri Sagiraju[1], Sandhya Kumari. P[2], Madhuri. M[4], Madhu S, Mahaboob Vali SK[3], Mahathi .M[5], Dr. K. Rajeshwar Rao[6]
[1, 2, 3, 4, 5]*B.Tech, School of Engineering, Hyderabad, India*
[6]*Professor School of Engineering, Mallareddy University*

Abstract: Agriculture, a cornerstone of our society, hinges on the critical factor of soil for its success. Soil composition varies, influencing the growth of crops through its chemical features. Selecting the right crops for specific soil types is pivotal, requiring proper soil classification and informed crop selection for optimizing agricultural productivity. In this project, we strive to empower farmers by developing a system that seamlessly integrates soil classification techniques with crop suggestion algorithms. Leveraging advanced Machine Learning techniques, specifically Random Forest and K- Nearest Neighbors (KNN), we classify soil series data. These classifications are then harmonized with a comprehensive crop dataset to predict suitable crops for specific soil seriesin a given region, considering its unique climatic conditions. Our datasets encompass chemical and geographical attributes of both soil and crops, providing a holistic understanding. In the ever-evolving landscape of agriculture, Machine Learning emerges as a budding technology, promising to enhance productivity and elevate the quality of crops in our vital agricultural sector.
Keywords: Soil classification, Crop selection, Climatic conditions, Agricultural productivity, Soil series data.

I. INTRODUCTION

Agriculture, a cornerstone of human civilization, plays a pivotal role in sustaining life and ensuring food security. In this context, soil, a fundamental component of the agricultural ecosystem, holds paramount importance. The chemical composition and characteristics of soil vary significantly, directly influencing crop growth and agricultural success. The need for a comprehensive understanding of soil attributes andtheir correlation with suitable crops is vital for optimizing agricultural productivity.

Soil Classification and Crop Suggestion using Machine Learning is a groundbreaking project that addresses this critical aspect of agriculture. The project recognizes the intricate relationship between soil properties and crop performance, aiming to empower farmers with informed decision-making tools. By integrating advanced machine learning techniques, the project seeks to revolutionize the waywe approach soil classification and crop selection.

II. PROBLEM STATEMENT

The aim of this project is to develop a machine learning solution for soil classification and crop suggestion. The dataset includes various soil attributes such as Nitrogen (N),Phosphorous (P), Potassium (K), temperature, humidity, pH, rainfall, and corresponding crop labels. The goal is to build a robust model that accurately classifies soil types and provides crop recommendations based on these soil attributes.

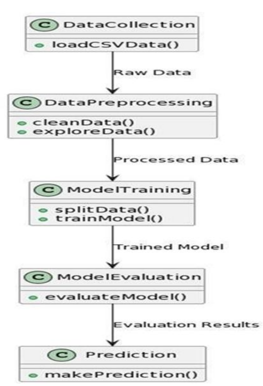


Fig: Dataflow diagram

III. LITERATURE REVIEW

The application of machine learning in agriculture, specifically in soil classification and crop recommendation, has gained considerable attention in recent literature. Several studies have explored the use of diverse machine learning algorithms for similar objectives.

The literature surrounding this domain underscores the pivotal role of data-driven approaches in optimizing crop yield and resource utilization. Researchers have increasingly employed machine learning algorithms, such as K Nearest Neighbours (KNN), Decision Trees, Random Forest, Naive Bayes, and XGBoost, to analyze soil characteristics and recommend suitable crops.

The dataset used in this project, derived from the 'Crop_recommendation.csv' file, is reminiscent of similar agricultural datasets utilized in previous studies. Scholars have emphasized the importance of features like temperature, humidity, pH, and rainfall in determining soil health and, consequently, crop selection.

The exploratory data analysis, visualization techniques, and model evaluation methods showcased in the code align with best practices in the field. Furthermore, the comparative analysis of multiple machine learning algorithms, including a hyperparameter tuning step, mirrors a comprehensive approach found in literature, ensuring robust model selection. This project contributes to the ongoing discourse on leveraging artificial intelligence in agriculture, offering a practical solution for farmers to make informed decisions based on soilcharacteristics and environmental factors.

IV. METHODOLOGY

A. Data Preprocessing

1) *Data Cleaning:* Address any missing or inconsistent values inthe dataset.
2) *Exploratory Data Analysis (EDA):* Analyze the distributionof soil attributes, identify outliers, and understand the characteristics of the dataset.
3) *Normalization/Scaling:* Standardize numerical features tobring them to a common scale.
4) *Encoding:* Convert categorical variables, such as soil types, into numerical representations.

B. Feature Selection

Identify and select relevant features for soil classification andcrop recommendation. This step helps in reducing dimensionality and improving model efficiency.

C. Model Architecture

1) *K-Nearest Neighbors (KNN):* Utilize the KNN algorithm forsoil classification. Adjust the number of neighbors based on cross-validation results.
2) *Decision Tree:* Implement a Decision Tree classifier with entropy as the criterion and a specified maximum depth to prevent overfitting.
3) *Random Forest:* Employ an ensemble of decision trees to enhance predictive accuracy and handle variability in thedataset.
4) *Naive Bayes:* Apply Gaussian Naive Bayes for probabilistic classification based on the assumption of independence between features.
5) *XGBoost:* Utilize XGBoost, a gradient boosting algorithm, tobuild an ensemble of weak learners for improved accuracy.

D. Model Training

Split the dataset into training and testing sets to evaluate modelperformance.

Train each machine learning model on the training set usingthe selected features.

E. Model Evaluation

Assess the performance of each model on the testing set usingmetrics such as accuracy, precision, recall, and F1 score.

Utilize cross-validation to validate model robustness and identify potential overfitting.

F. Hyperparameter Tuning

Conduct grid search or random search to find optimal hyperparameters for selected models, improving overall performance.

G. Crop Recommendation

Once the soil classification model is validated, use it to predictsoil types for new data.

Develop a crop recommendation system based on the predicted soil type, considering the characteristics of crops thatthrive in specific soil conditions.

V. EXPERIMENT RESULTS

```
crop = pd.read_csv('../input/crop-recommendation-dataset/Crop_recommendation.csv')
crop.head(5)
```

	N	P	K	temperature	humidity	ph	rainfall	label
0	90	42	43	20.879744	82.002744	6.502985	202.935536	rice
1	85	58	41	21.770462	80.319644	7.038096	226.655537	rice
2	60	55	44	23.004459	82.320763	7.840207	263.964248	rice
3	74	35	40	26.491096	80.158363	6.980401	242.864034	rice
4	78	42	42	20.130175	81.604873	7.628473	262.717340	rice

Fig : (a) Data description

```
crop['label'].unique()
```

```
array(['rice', 'maize', 'chickpea', 'kidneybeans', 'pigeonpeas',
       'mothbeans', 'mungbean', 'blackgram', 'lentil', 'pomegranate',
       'banana', 'mango', 'grapes', 'watermelon', 'muskmelon', 'apple',
       'orange', 'papaya', 'coconut', 'cotton', 'jute', 'coffee'],
      dtype=object)
```

```
crop['label'].nunique()
```

```
22
```

```
crop['label'].value_counts()
```

```
rice           100
maize          100
jute           100
cotton         100
coconut        100
papaya         100
orange         100
apple          100
muskmelon      100
watermelon     100
grapes         100
mango          100
banana         100
pomegranate    100
lentil         100
blackgram      100
mungbean       100
mothbeans      100
pigeonpeas     100
kidneybeans    100
chickpea       100
coffee         100
Name: label, dtype: int64
```

Fig: (b) Data description

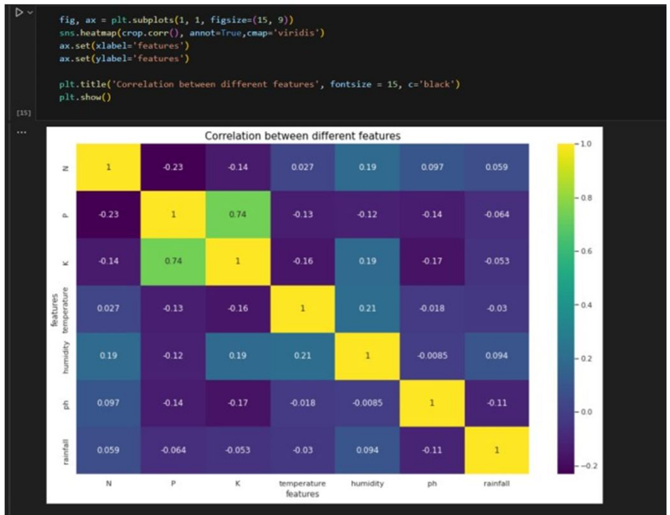


Fig: Correlation between features

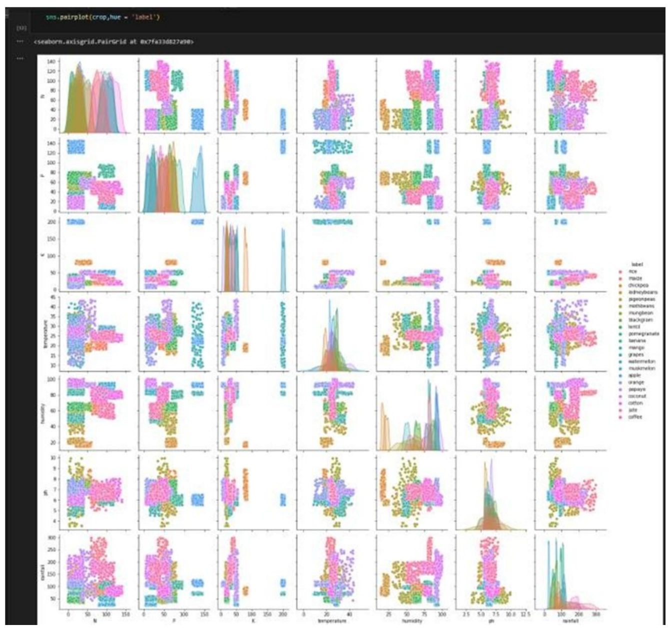


Fig: Pair Plot

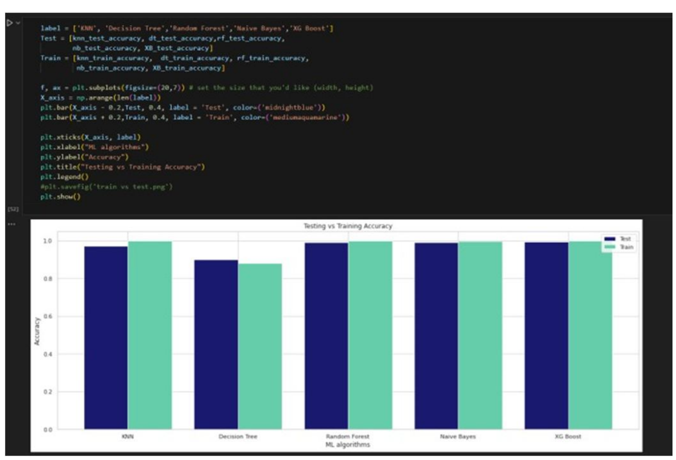


Fig: Training vs Testing accuracy

VI. CONCLUSION

This project has successfully implemented machine learning models for soil classification and crop recommendation, demonstrating the feasibility of leveraging data-driven approaches in agriculture. The developed models showcase varying performances, and the crop recommendation system offers practical insights for farmers, aiding in informed decision-making. While the project contributes to improving crop selection processes, acknowledging limitations and continuous model refinement are crucial for ensuring the system's reliability and relevance in real-world agricultural scenarios.

VII. FUTURE ENHANCEMENT

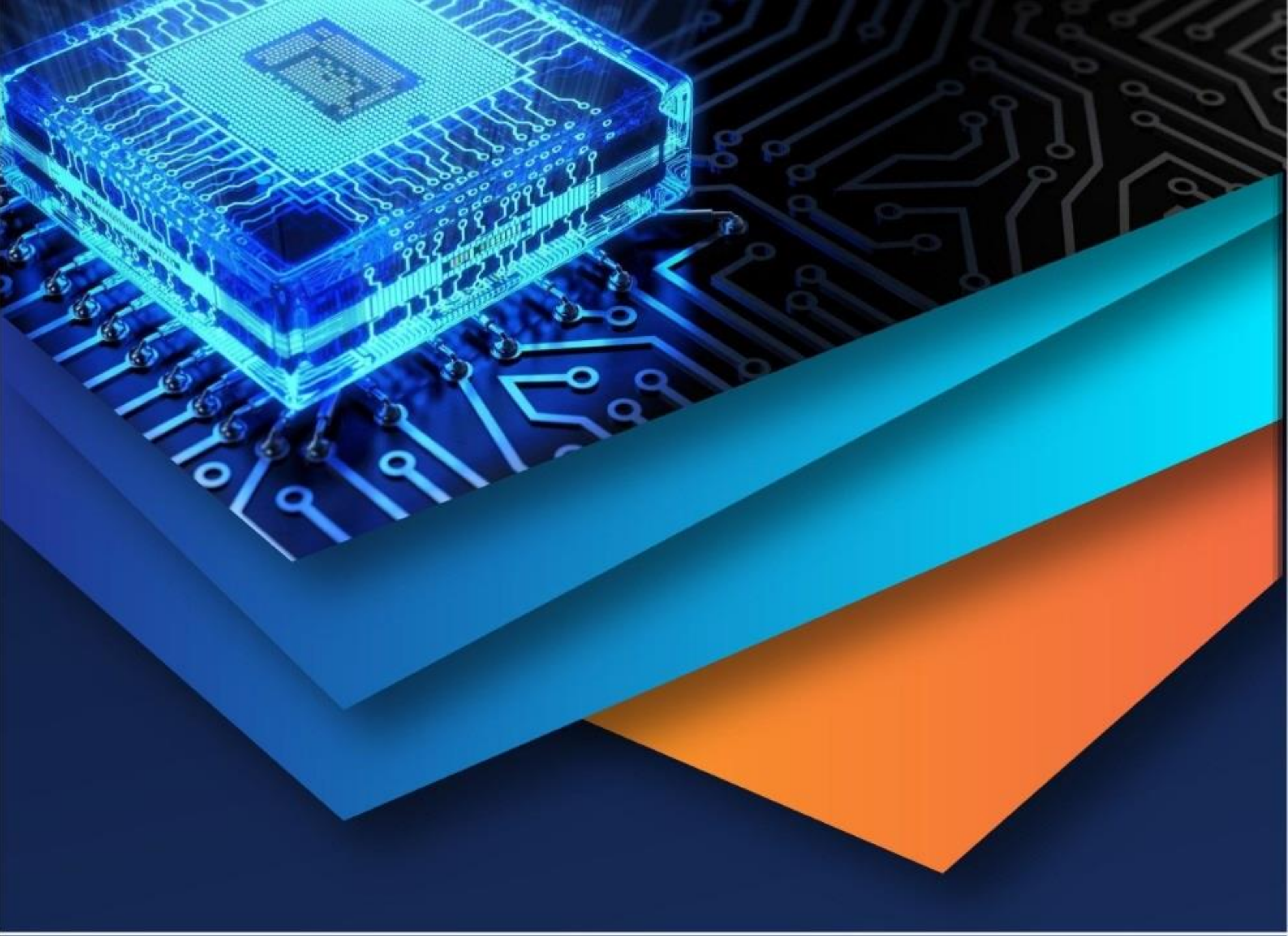
There is a scope for further development in our project to a great extent. In future suitable fertilizers are suggested for the well growth of the crop cultivated. The present models deals with available old data whereas the future model contain the real time a data that is directly received from agricultural land that is placed with sensors .The sensors senses the soil fertility and other minerals contained in the soil.

VIII. ACKNOWLEDGEMENT

I would like to express my heartfelt gratitude to my guide, Dr. K. Rajeshwar Rao, and head of department, Dr . Thayyaba Khatoon, for their invaluable guidance and unwavering support throughout the development of this project. Their insightful feedback helped me to refine my ideasand develop a comprehensive understanding of the subject matter.
Their mentorship was instrumental in shaping my approach towards the project, and I am grateful for the knowledge and experience they shared with me. Without their encouragementand support, this project would not have been possible. Once again, I extend my sincere thanks to my guides and head of department for their unwavering support and guidance.
Our sincere thanks to all the teaching and non-teaching staff ofDepartment of Computer Science and Engineering (AI&ML) for their support throughout our project work.

REFERENCES

[1] Gholap, J., Ingole, A., Gohil, J., Gargade, S. and Attar, V., 2012. Soil data analysis using classification techniquesand soil attribute prediction. arXiv preprint arXiv:1206.1557.
[2] H. Eswaran, R. Ahrens, T. J. Rice and B. A.Stewart, Soil classification: a global desk reference, CRC Press, 2002
[3] M. P. K. Devi, U. Anthiyur and M. S. Shenbagavadivu, "Enhanced Crop Yield Predictionand Soil Data Analysis Using Data Mining", International Journal of Modern Computer Science,vol. 4, no. 6, 2016.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY