IIIT ALLAHABAD

Crop Recommendation System

Course Instructor: Dr.Muneendra Ojha

Presented by Anshul Bhardwaj(IIT2021057)

IIT Allahabad

Overview

Presented by Anshul Bhardwaj

IIT Allahabad

Introduction

Machine Learning is highly proficient at examining data related to soil conditions, encompassing factors such as **moisture levels, temperature, and chemical composition,** all of which influence the growth of crops and the well-being of livestock.

In contemporary agriculture, this capability enables a **higher degree of precision in crop cultivation**, allowing farmers to address the needs of plants and animals on an almost individual basis. As a result, it greatly **enhances the efficiency of farmers' decision-making processes**.

Presented by Anshul Bhardwaj

IIT Allahabad

Through this exploration, we seek to address questions such as:

- Which crops are best suited for a particular region based on its soil content and temperature patterns?
- How can we optimize crop production by aligning planting decisions with environmental conditions?
- Can we improve crop yield and quality by understanding the interplay between soil nutrients and temperature?

Presented by Anshul Bhardwaj

IIT Allahabad

Significance of Crop Prediction:

01

Enhances Farmer's decision making process.

04

Improved Crop Yield and Quality: By leveraging data-driven insights

·05

Sustainability: By optimising crop choice and cultivation practices. This project contributes to more eco friendly farming

02

Resource Efficiency: Understanding the relationship between soil content, temperature, and crop selection allows for the efficient allocation of resources such as water, fertilizers, and pesticides. This can reduce wastage and environmental impact.

03

Risk Mitigation: Farmers often face risks due to unpredictable weather conditions. This project can help identify crops that are more resilient to specific climate conditions, reducing the impact of adverse weather events.

IIIT ALLAHABAD

Problem Statement

The primary goal is to develop a predictive model that can forecast the most suitable crops
for a given region based on the soil composition and temperature conditions. By analyzing
the data at hand, we aim to provide valuable insights into crop selection and yield estimation.

Presented by Anshul Bhardwaj

Literature Review

Several studies emphasize the significance of utilizing machine learning and data-driven approaches for crop forecasting. Machine learning techniques, including decision trees, support vector machines, and ensemble methods, have been applied to analyze complex relationships between soil properties, temperature, and crop performance.Research by Central Queensland University, Melbourne 3000, Australia.

Additionally, research has delved into the practical implications of deploying predictive models in the agricultural domain. Decision support systems that integrate crop prediction models have been proposed to aid farmers in making informed decisions about crop selection, resource allocation, and risk management. The study by [Author et al., Year] exemplifies the successful implementation of such systems, demonstrating their potential to enhance agricultural productivity and sustainability. Research by Ch. Rakesh D; Vishnu Vardhan; Babu Bhavani Vasantha; G. Sai Krishna

IIIT Allahabad

Literature Review

The integration of remote sensing technologies and satellite imagery has emerged as a key trend in crop prediction literature. These technologies enable researchers to capture real-time data on soil moisture, nutrient levels, and temperature variations, contributing to a more comprehensive understanding of the agro-ecosystem.Crop monitoring by multimodal remote sensing: A review Priyabrata Karmakar a b, Shyh Wei Teng b, Manzur Murshed a b, Shaoning Pang b, Yanyu Li c, Hao Lin d showcase the potential of remote sensing in enhancing the spatial resolution of predictive models, thereby improving their precision.

In conclusion, the literature review underscores the significance of developing predictive models that integrate soil composition and temperature conditions for crop forecasting. Existing research demonstrates the efficacy of machine learning techniques, the integration of remote sensing technologies, and the adaptation of models to climate change challenges.

IIIT Allahabad

Literature Review

01

Crop Yield Prediction with Deep Learning and Remote Sensing mdpi research

Santoso Wibowo, Srimannarayana Grandhi, Priyanga Muruganantham

Central Queensland University, Melbourne 3000, Australia

02

Crop Prediction using Machine Learning Approaches

International Journal of Engineering Research and

04

Soil vs Weather monitoring system.

Velmurugan Sathya Narayanan, Kavin N Raj, Kishore Kumar and Manoj Kumar

03

Classification of crop based on macronutrients and weather data using machine learning techniques
Author links open overlay panel

Ritesh Dash a , Dillip Ku Dash , G.C. Biswal

IIIT ALLAHABAD

Dataset Discussion

Context: Precision agriculture is in trend nowadays. It helps the farmers to
get informed decision about the farming strategy

Source:
This dataset was build by augmenting datasets of rainfall, climate and fertilizer data
available for India. Gathered over the period by ICFA, India.

Presented by Anshul Bhardwaj

IIIT ALLAHABAD

Data fields

N - ratio of Nitrogen content in soil

P - ratio of Phosphorous content in soil

K - ratio of Potassium content in soil

temperature - temperature in degree Celsius

humidity - relative humidity in %

ph - ph value of the soil

rainfall - rainfall in mm

Crop_recommendation

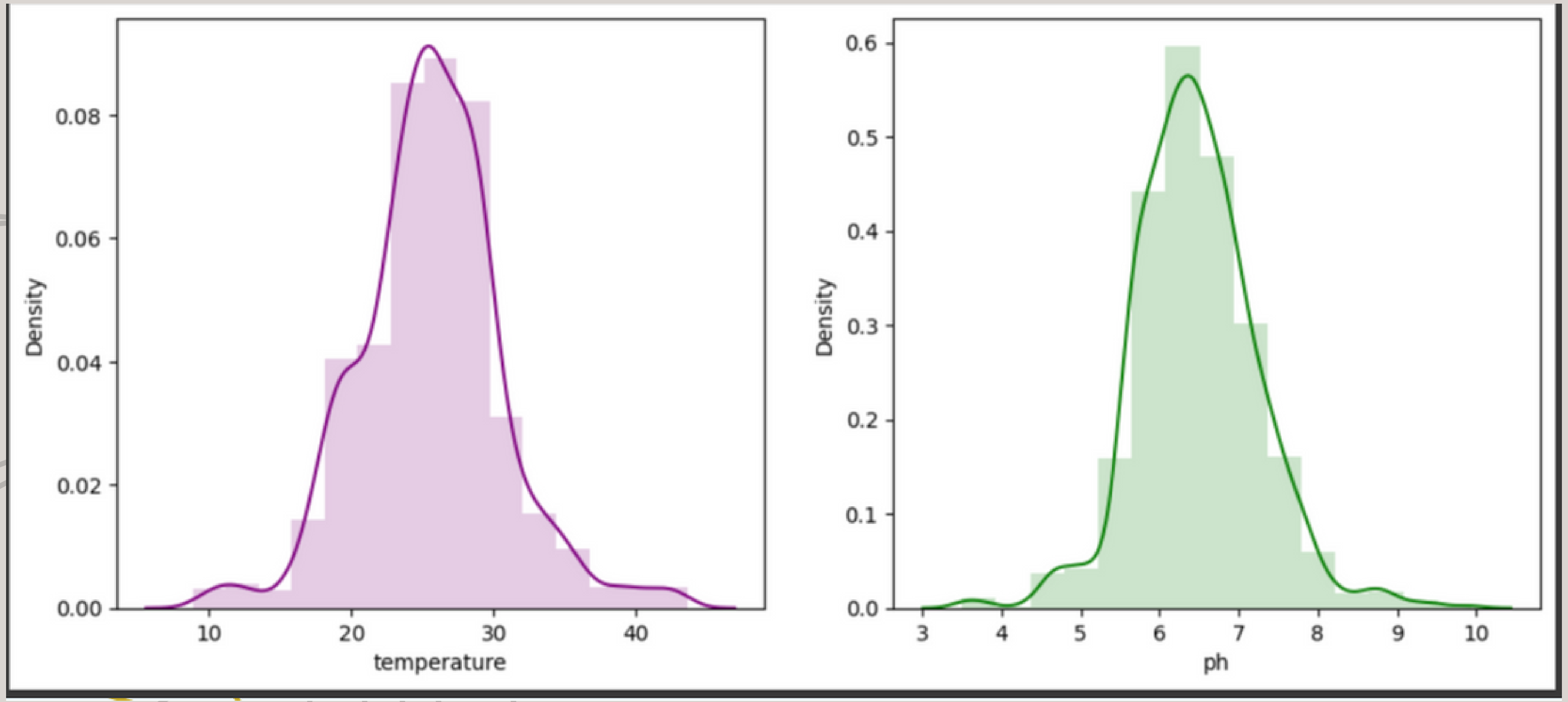
N	P	K	temperature	humidity	ph	rainfall	label
90	42	43	20.87974371	82.00274423	6.502985292000000	202.9355362	rice
85	58	41	21.77046169	80.31964408	7.038096361	226.6555374	rice
60	55	44	23.00445915	82.3207629	7.840207144	263.9642476	rice
74	35	40	26.49109635	80.15836264	6.980400905	242.8640342	rice
78	42	42	20.13017482	81.60487287	7.628472891	262.7173405	rice
69	37	42	23.05804872	83.37011772	7.073453503	251.0549998	rice
69	55	38	22.70883798	82.63941394	5.70080568	271.3248604	rice

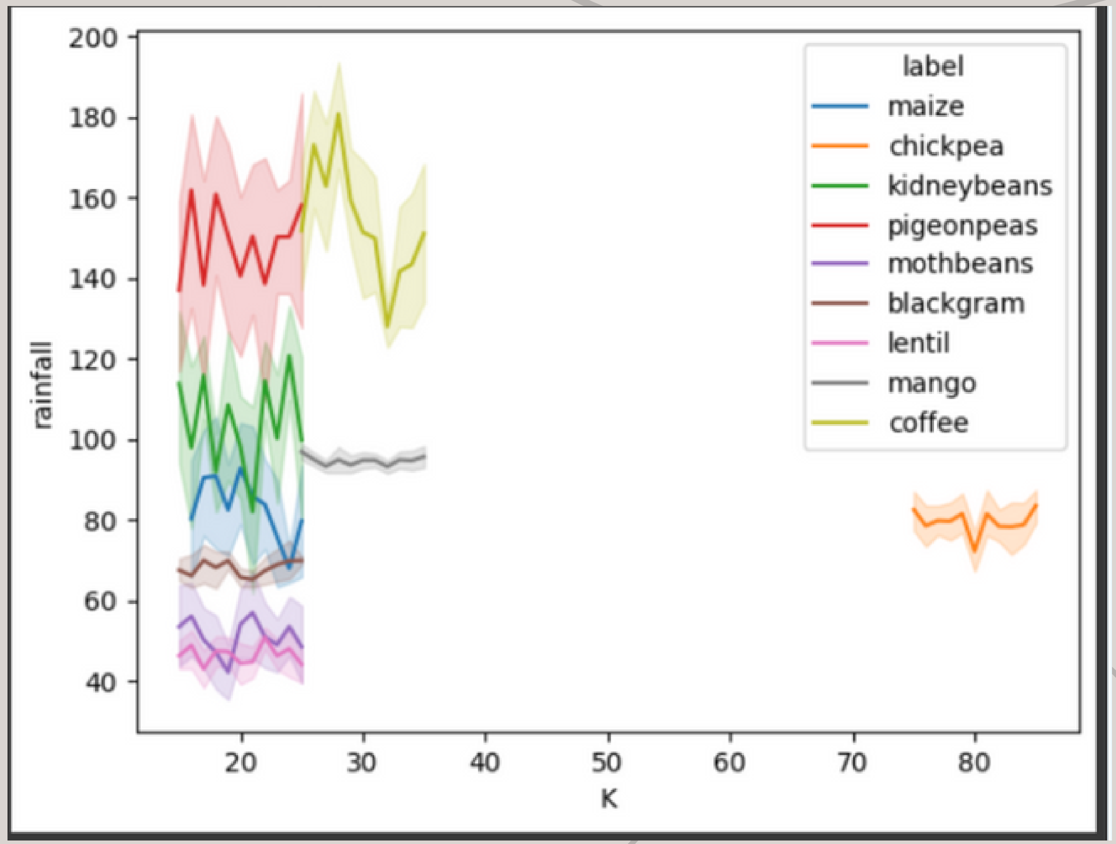
Presented by Anshul Bhardwaj

IIIT ALLAHABAD

DATA Parameters visualization

Identify Here, is some dataset parameters visualizaion in graphical manner



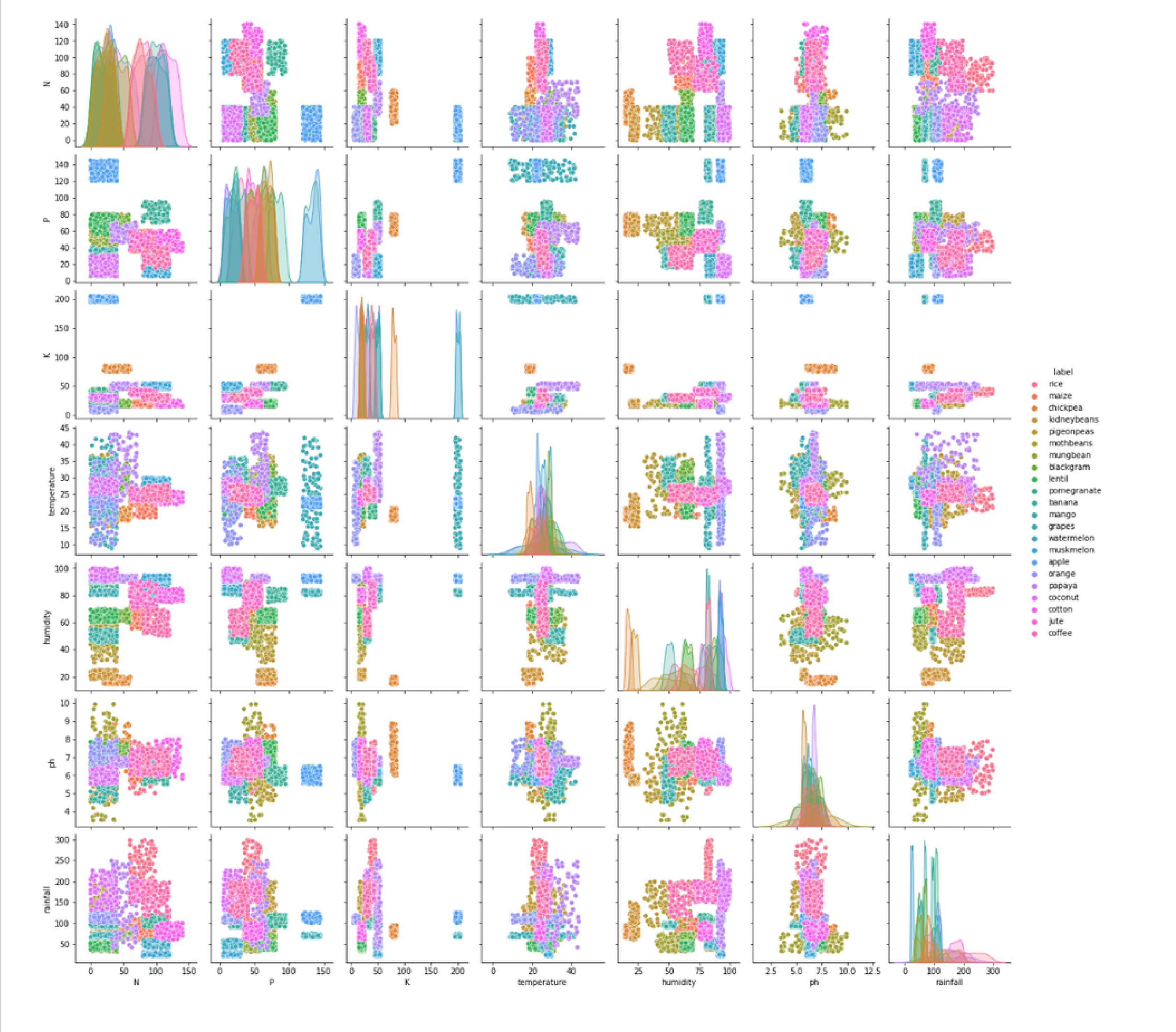


temprature and ph have a Gaussian Bell shaped Distribution

Presented by Anshul Bhardwaj

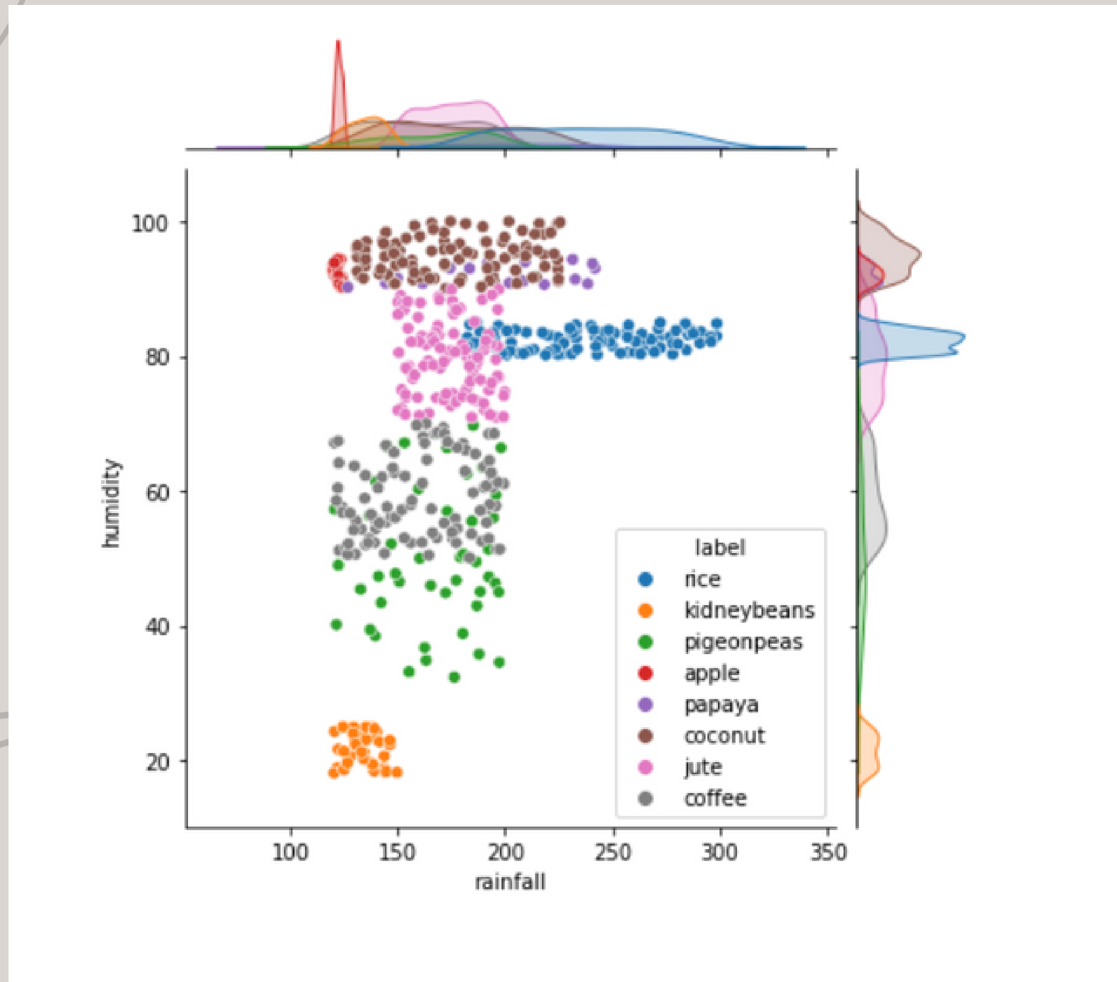
IIIT ALLAHABAD

Digonal distribution b/w 2 features using pairplot

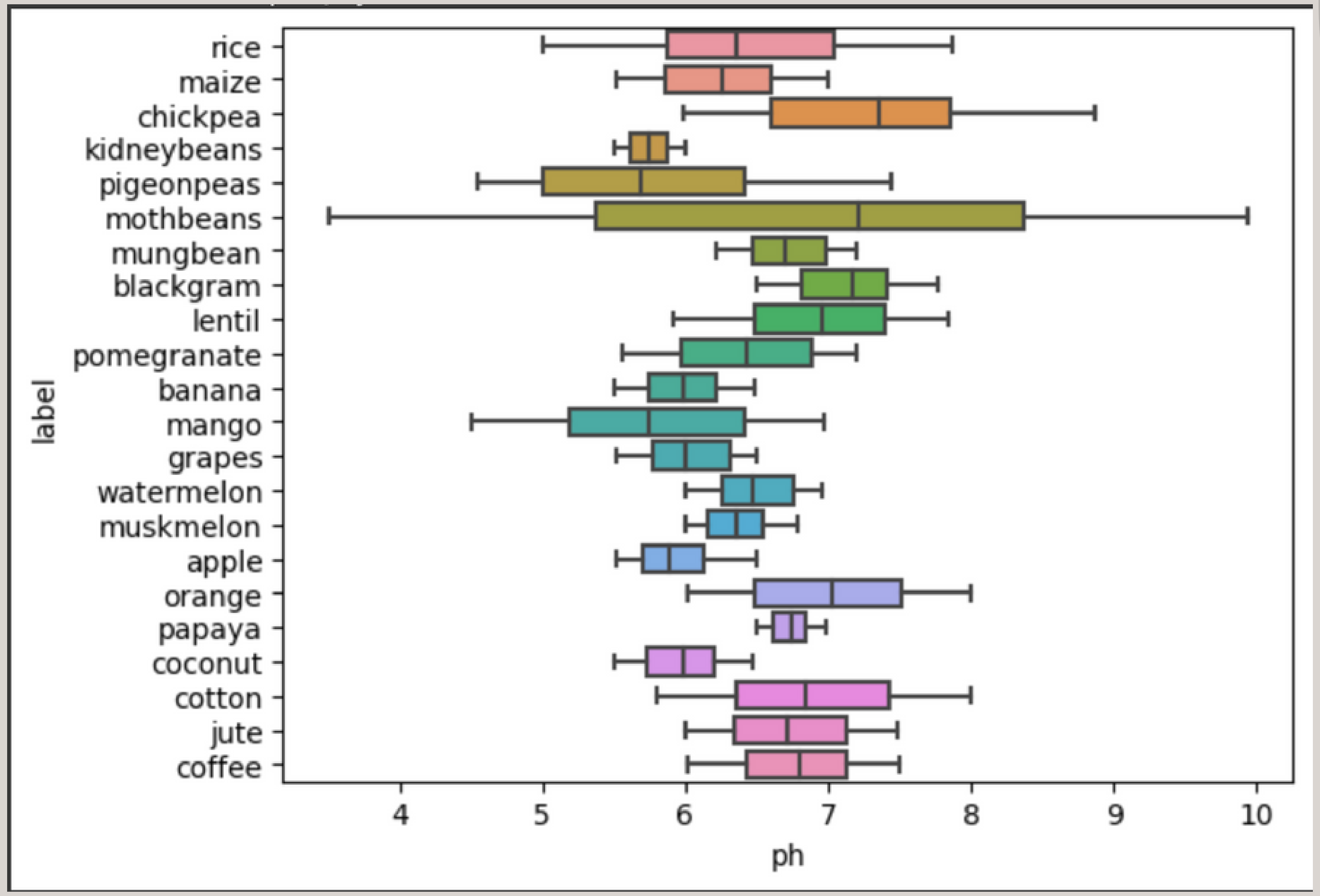


Presented by Anshul Bhardwaj

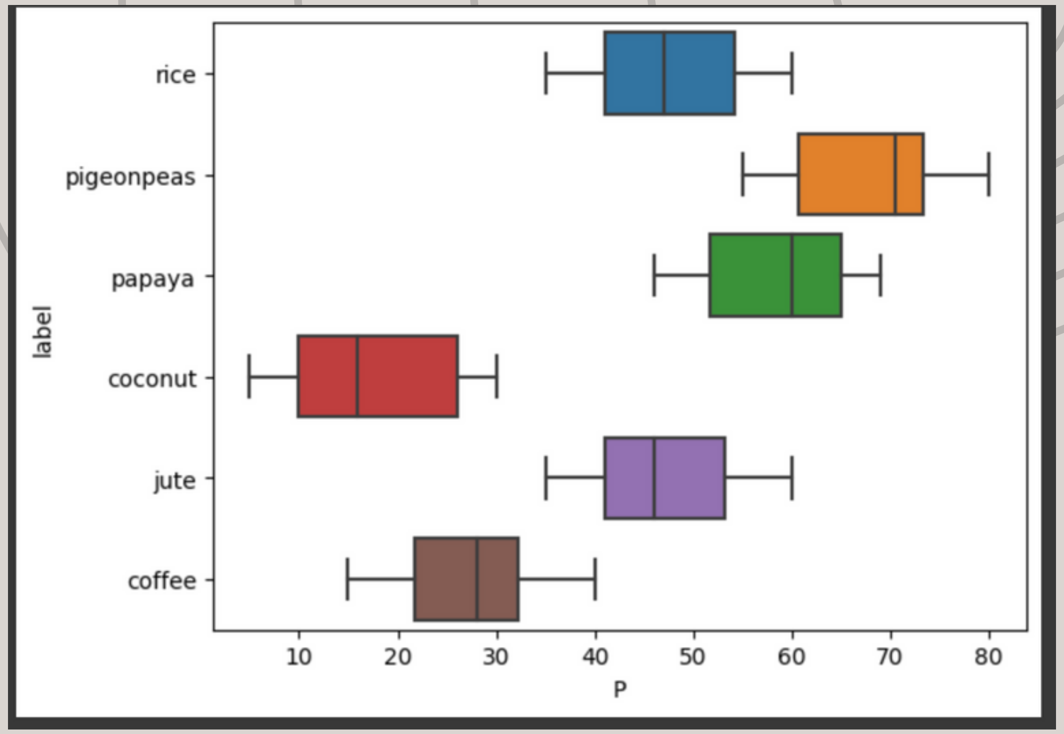
IIIT ALLAHABAD



This graph correlates with average potassium (K) and average nitrogen (N) value (both>50



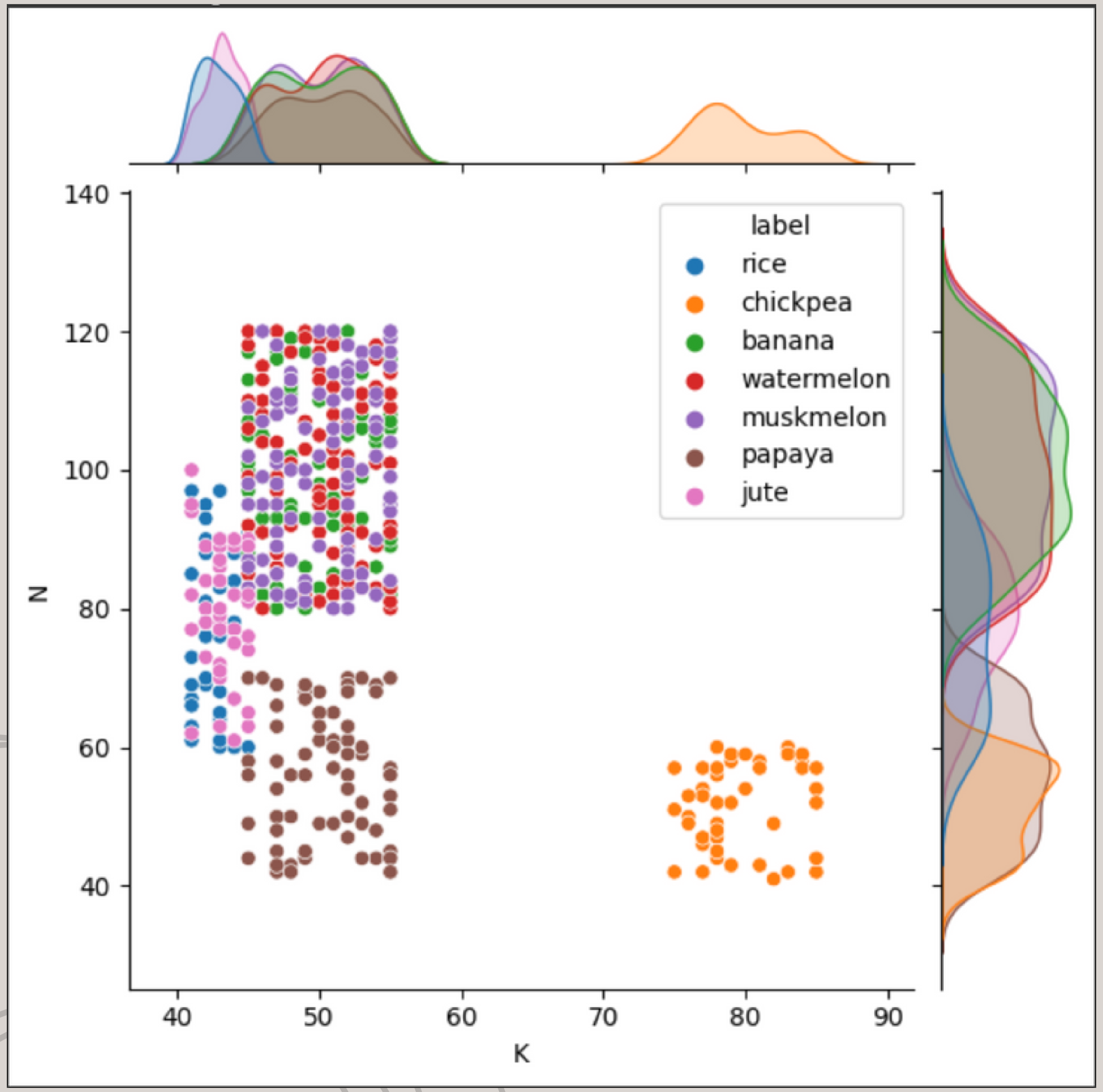
Ph values are critical. preferable values are between 6 and 7.



Phosphorous levels are quite differentiable when it rains heavily (above 150 mm).

Presented by Anshul Bhardwaj

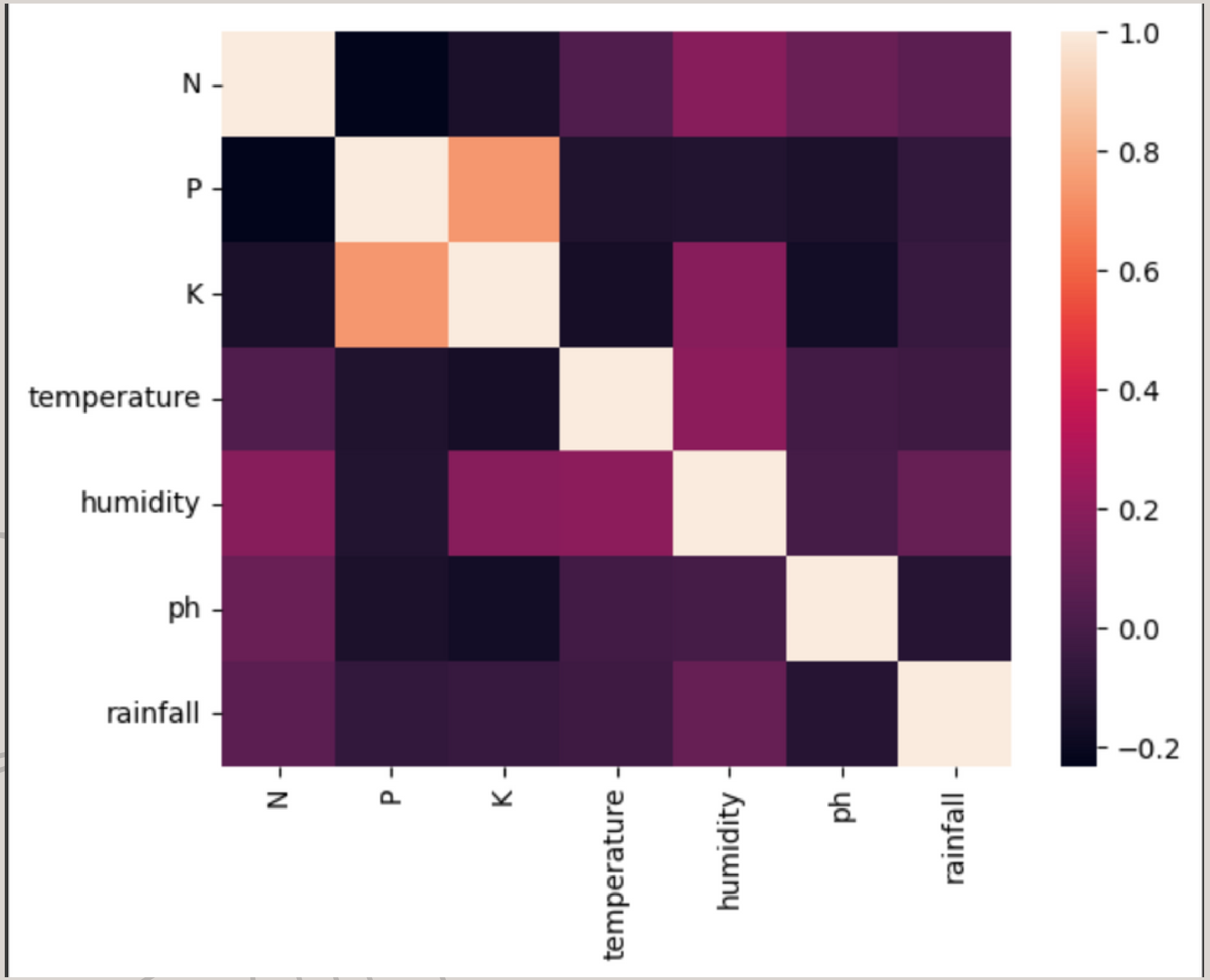
IIIT ALLAHABAD



This chart shows a connection between the mean levels of potassium (K) and nitrogen (N) in the soil, with both values being greater than 50. These soil components have a direct impact on the nutritional content of the food.

Presented by Anshul Bhardwaj

Data Preperation



As mentioned earlier, two of our features (temperature and pH) follow a Gaussian distribution. To scale these features within a range of 0 to 1, we can use the MinMaxScaler.

Correlation visualization between features. We can see how Phosphorous levels and Potassium levels are highly correlated.

IIIT ALLAHABAD

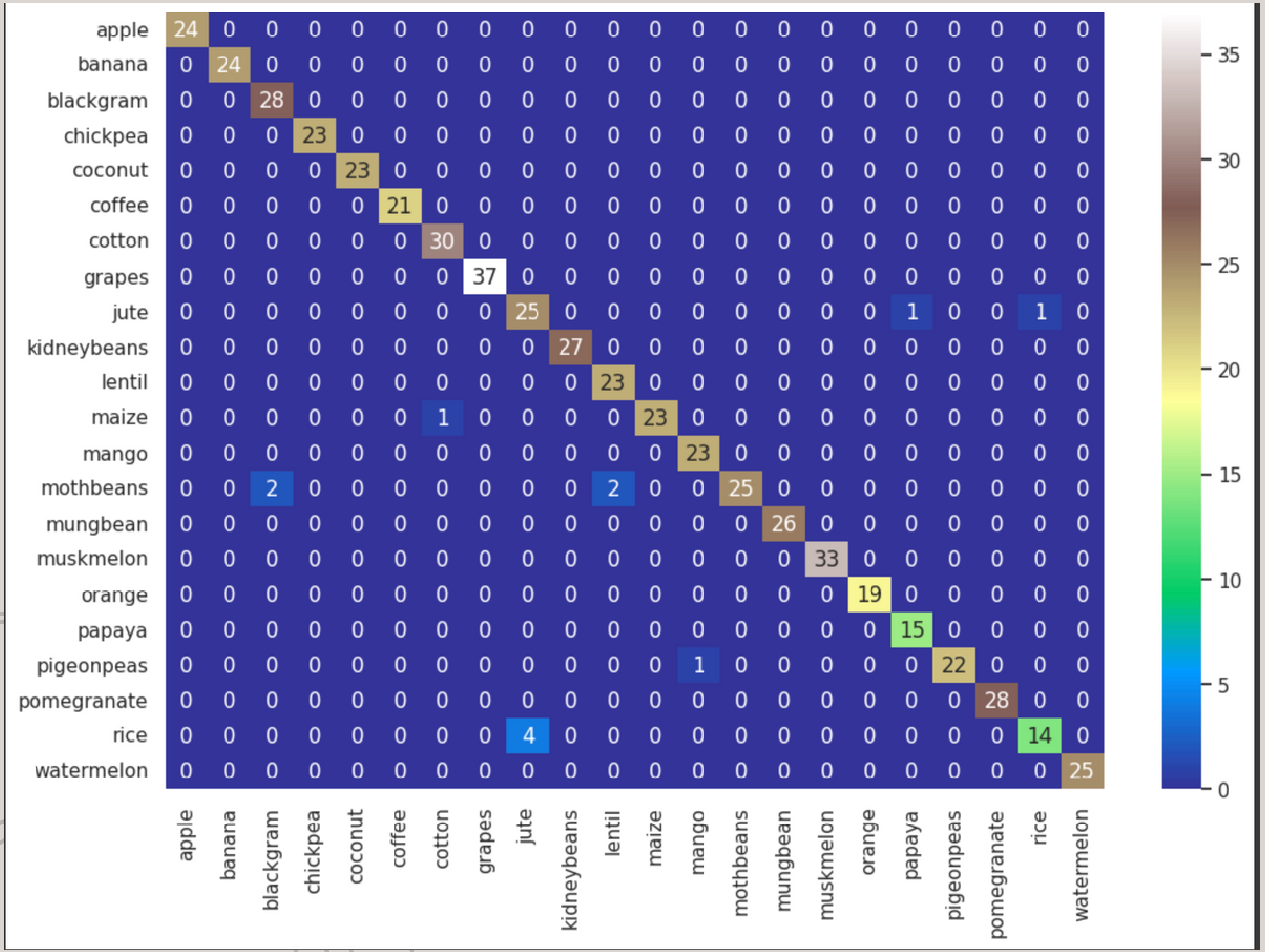
KNN Classifier

KNN: K-Nearest Neighbors (KNN) is a popular machine learning algorithm used for
both classification and regression tasks. It's a simple and intuitive algorithm that

- makes predictions based on the similarity between data points.
- Simple to understand and implement.
- Can be used for both classification and regression tasks.
- Works well for multi-class classification.
- Non-parametric, meaning it doesn't make strong assumptions about the underlying
  data distribution.

Presented by Anshul Bhardwaj

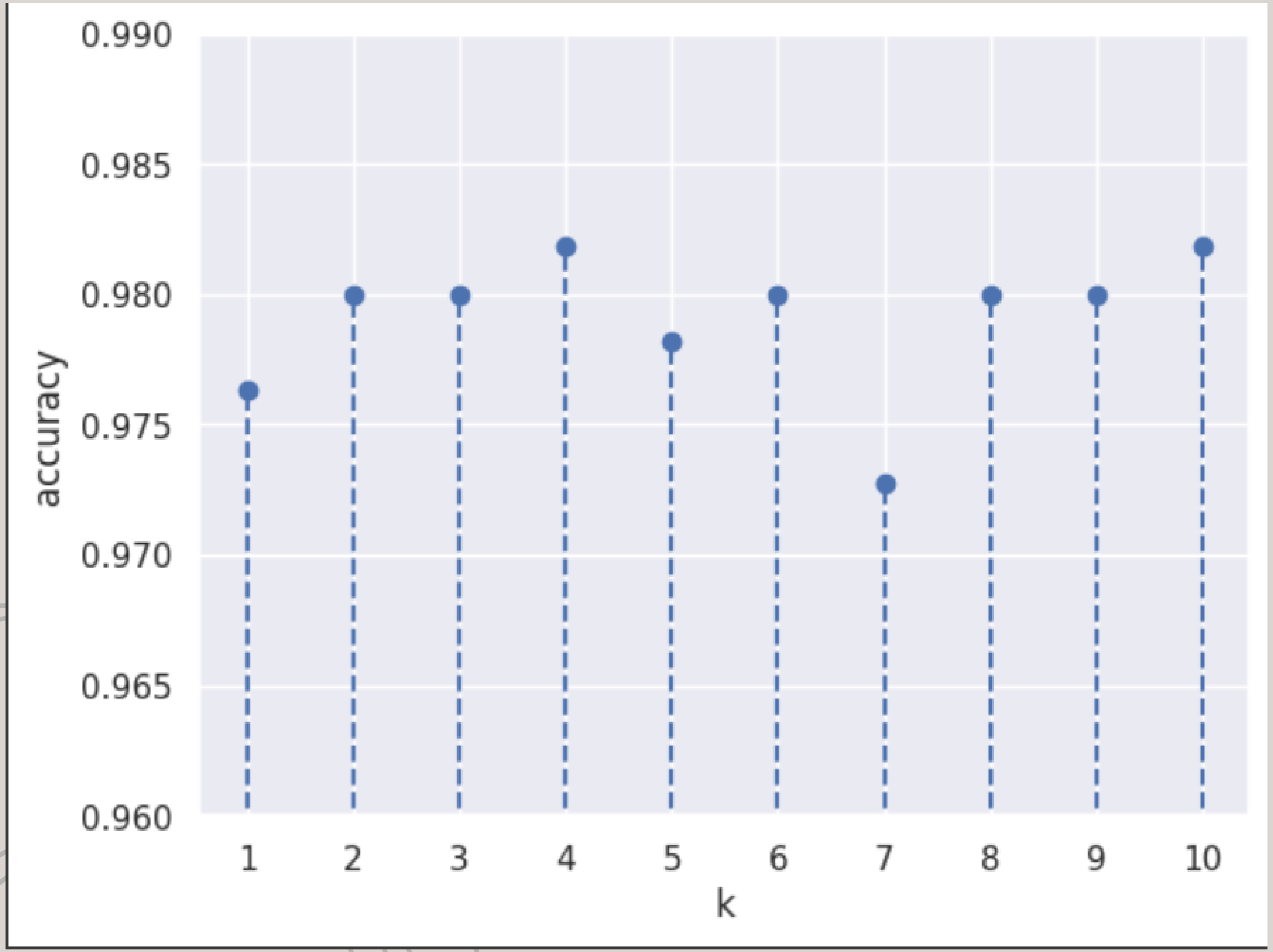
Confusion matrix



KNN Accuracy :97.8%

Presented by Anshul Bhardwaj

KNN values over different n selection



Here 4 seems to be the best value of K in terms of maintaining balance between performance ,computational efficiency and accuracy of model.

Presented by Anshul Bhardwaj

IIIT ALLAHABAD

CLASSIFICATION USING SUPPORT VECTOR CLASSIFIER (SVC)

Support Vector Classifier (SVC) is employed for crop prediction using different kernel functions. Three instances of the SVC are trained and evaluated on a crop prediction dataset. The first model utilizes a linear kernel, the second employs a radial basis function (RBF) kernel, and the third employs a polynomial kernel. Each SVC is fitted to the scaled training data, and the accuracy of each model is assessed on the scaled testing data.

```
Linear Kernel Accuracy:  0.9745454545454545
Rbf Kernel Accuracy:  0.9872727272727273
Poly Kernel Accuracy:  0.9890909090909091
```

Presented by Anshul Bhardwaj

IIIT ALLAHABAD

Hyperparameter tuning of SVM

Grid Search: Grid search is a common method for hyperparameter tuning. It involves defining a grid of hyperparameter values and systematically searching through the grid to find the combination that results in the best model performance.

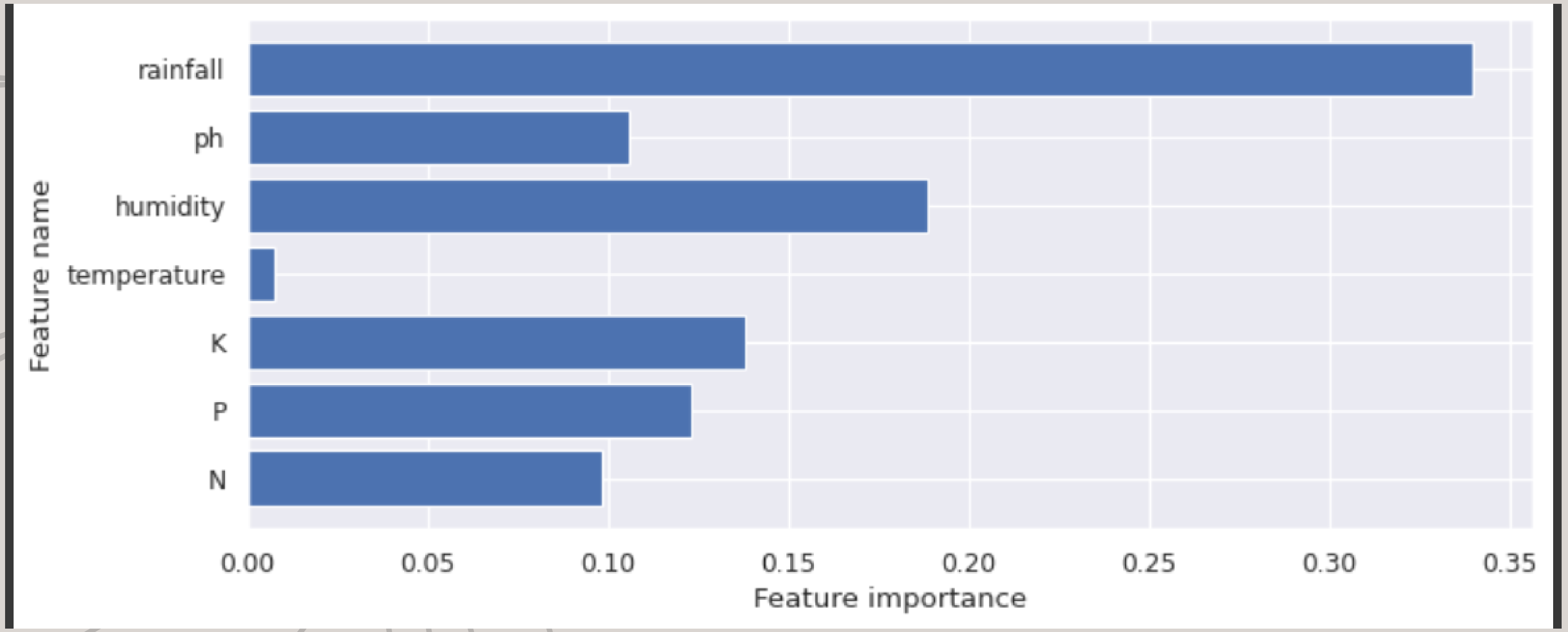
Linear Kernel accuracy after grid search tuning:98.86%

Presented by Anshul Bhardwaj

IIIT ALLAHABAD

CLASSIFICATION USING DECISION TREE

Classification using a decision tree involves constructing a hierarchical tree-like model that systematically partitions the dataset based on input features, ultimately leading to a decision or prediction.



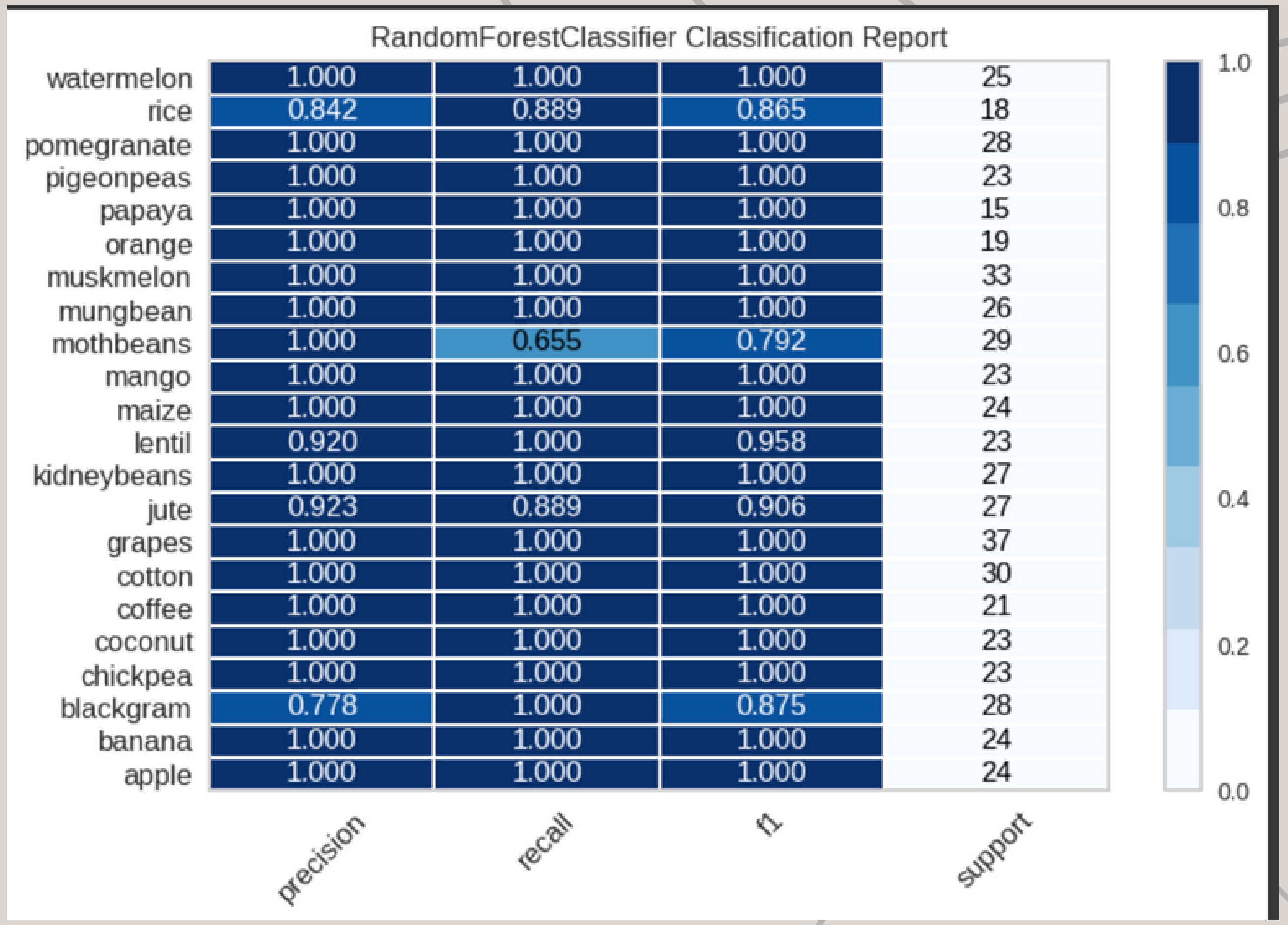
Decision Tree Accuracy:98.87%

IIIT ALLAHABAD

CLASSIFICATION USING RANDOM FOREST

RandomForestClassifier is instantiated with specific hyperparameters, including a maximum depth of 4 for each tree (max_depth=4), 100 estimators (trees) in the forest (n_estimators=100), and a fixed random state for reproducibility (random_state=42)

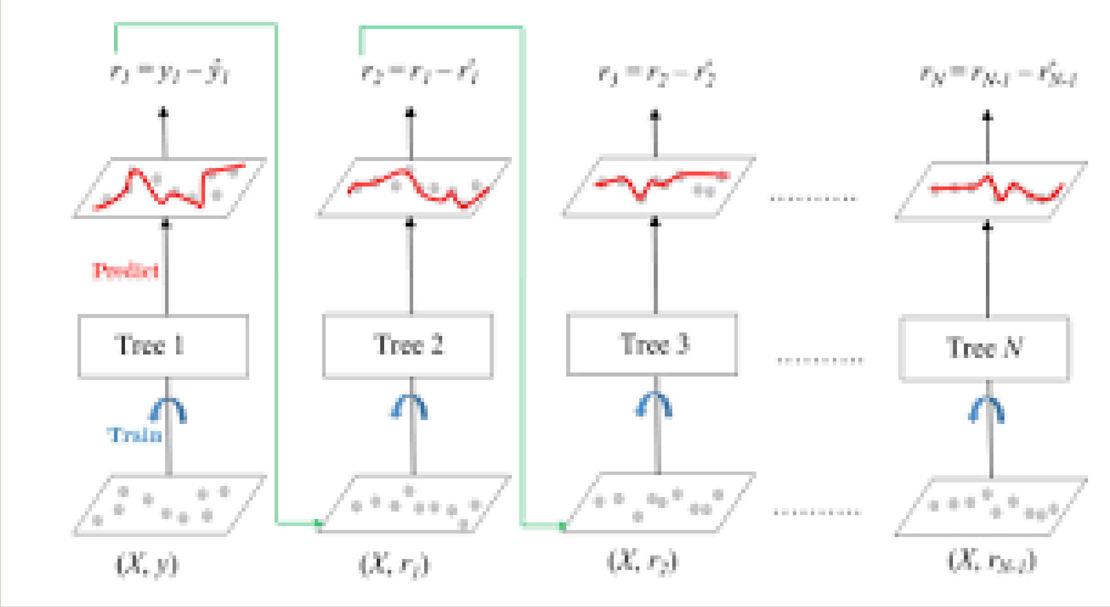
Random Forest Accuracy:97%



IIIT ALLAHABAD

CLASSIFICATION USING Gradient Boosting

Gradient Boosting is an ensemble learning technique that builds a predictive model in a stage-wise fashion by combining the predictions of multiple weak learners, often decision trees. Unlike traditional boosting methods, gradient boosting optimizes a loss function by fitting each new tree to the residual errors of the combined model, gradually reducing the overall error.

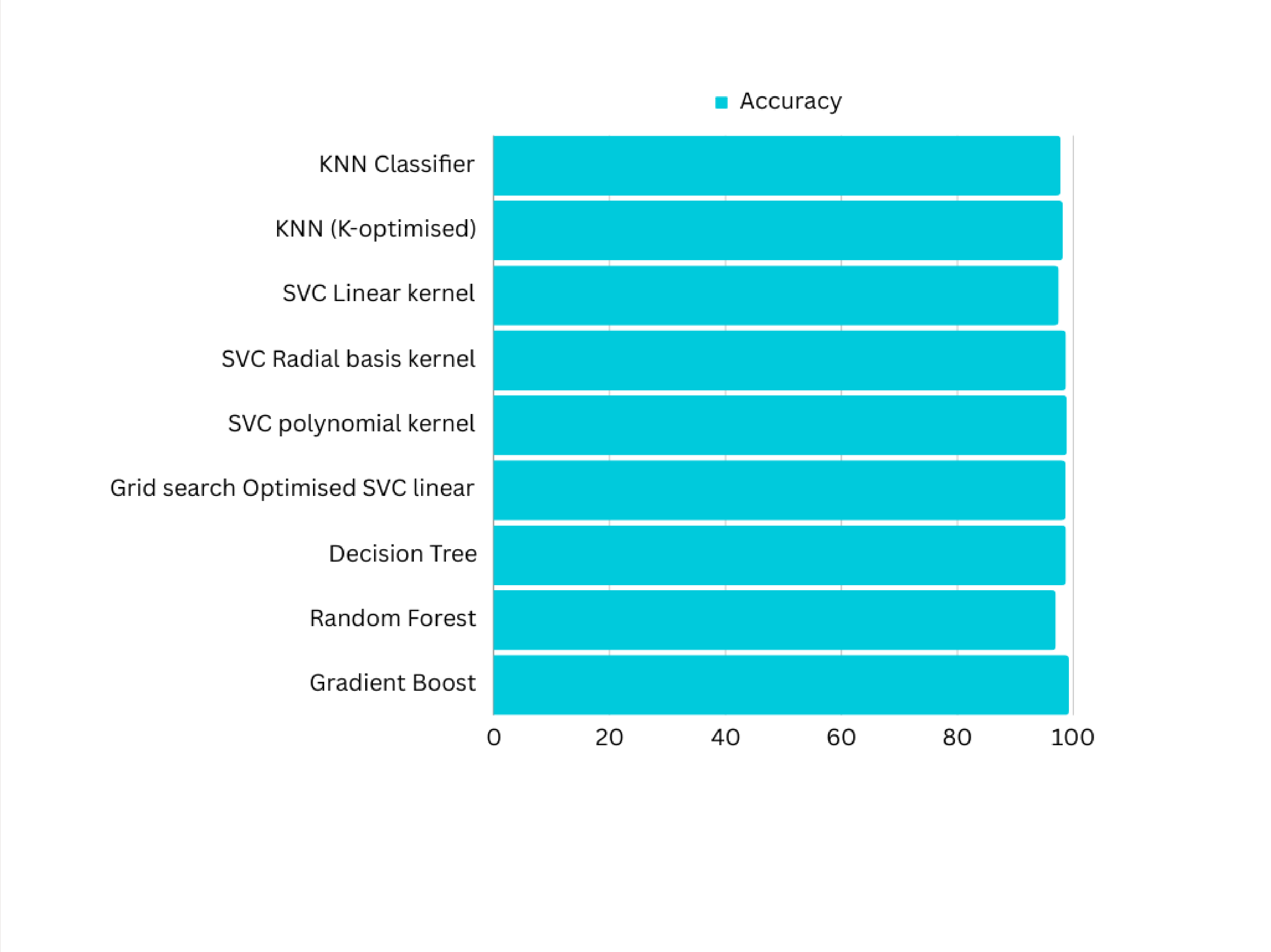


Accuracy:99.27%

Performance Comparison of Models

Gradient Boosting comes out to be the most accurate model followed by polynimial kernel Support vector classifier.

Polynomial kernel for SVC is particularly efficient as we do not need to perform hyperparameter optimistation.



IIT ALLAHABAD

Conclusion

In conclusion, the crop prediction and analysis project represents a significant stride towards leveraging data-driven solutions for optimizing agricultural practices. By integrating soil chemical properties, weather conditions, and other environmental factors, the project has sought to enhance the accuracy and efficiency of crop prediction models. The utilization of machine learning techniques has proven instrumental in unraveling the intricate relationships among diverse variables, allowing for more informed decisions in crop selection and resource management.

As we move forward, the project underscores the importance of continued research in refining models, promoting data standardization and sharing, and developing user-friendly decision support systems for farmers. Ultimately, the intersection of technology and agriculture holds immense potential in fostering sustainable practices, adapting to climate change, and ensuring global food security.

Presented by Anshul Bhardwaj

IIIT ALLAHABAD

Future Work

Looking ahead, there are several avenues for future work that can further enhance the impact and applicability of the crop prediction and analysis project. A few more applications and integrations as we move forward in our learning and expanding our domain beyond this course

- The integration of advanced data sources, such as high-resolution satellite imagery and drone technology, can provide real-time insights into crop health, growth stages, and potential stressors.
- exploring the application of emerging technologies like **Internet of Things (IoT)** devices within the agricultural domain holds promise. **Deploying sensors in fields to monitor soil conditions, temperature, and humidity in real-time** can offer a wealth of data for analysis, enabling more precise and timely predictions.
- As climate change continues to impact agriculture, future research could investigate **how prediction models can be adapted to account for evolving climate patterns** and identify resilient crop varieties.

Presented by Anshul Bhardwaj

IIIT ALLAHABAD

References

- Dr. Muneendra Ojha IML Course Slides

  Ojha, Dr. Muneendra - Department of IT, IIIT-Allahabad

  Indian Institute of Information Technology, Allahabad
  https://it.iiita.ac.in › uid=muneendra

- S. Pudumalar, Ramanujam Elangovan, R. Rajashree, C. Kavya, T. Kiruthika and J Nisha, Crop recommendation system for precision agriculture, pp. 32-36, 2017.
  Show in Context Google Scholar

- Yogesh Gandge and Sandhya, A study on various data mining techniques for crop yield prediction, pp. 420-423, 2017.
  Show in Context Google Scholar

- Simon Tavasoli, "10 Machine Learning Algorithms You Need to Know", Simplilearn Solutions.
  Show in Context Google Scholar

Presented by Anshul Bhardwaj

- Avinash Kumar, Sobhangi Sarkar and Chittaranjan Pradhan, "Recommendation system for crop identification and pest control technique in agriculture", 2019 International Conference on Communication and Signal Processing (ICCSP), pp. 0185-0189, 2019.

Show in Context View Article

Google Scholar

- Stanford Machine Learning specialization by Andrew Ng
Stanford Machine Learning Specialization [3 courses] (Stanford)

- A review Priyabrata Karmakar a b, Shyh Wei Teng b, Manzur Murshed a b, Shaoning Pang b, Yanyu Li c, Hao Lin d

- Performance.Research by Central Queensland University, Melbourne 3000, Australia.
- Research by Ch. Rakesh D; Vishnu Vardhan; Babu Bhavani Vasantha; G. Sai Krishna

THANK YOU