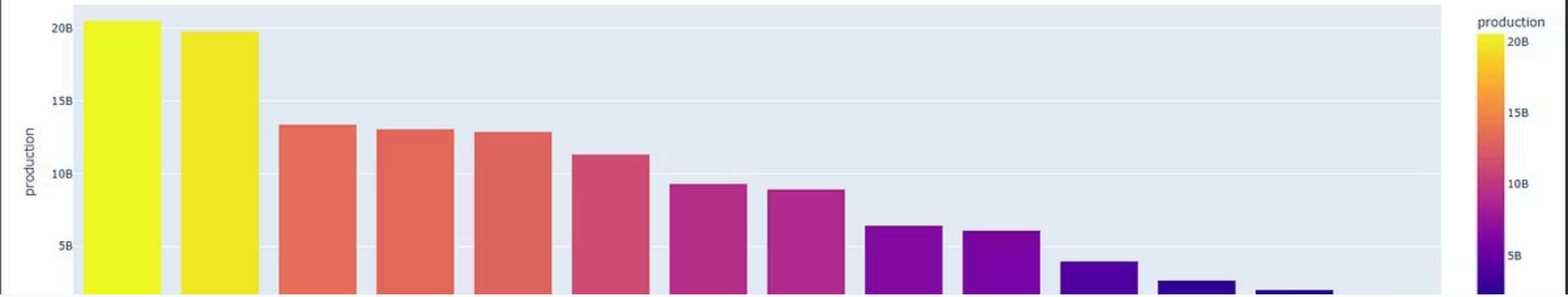
Why Do Some Districts in Kerala Produce More Than Others?

Student Name: Sujal S Acharya
Guide: Dr. Piyush Chauhan
Course: Data Science
Symbiosis Institute of Technology, Nagpur



Unravelling Kerala's Agricultural Puzzle

Agriculture forms the backbone of Kerala's economy, yet a striking disparity in productivity exists across its districts.

Economic Pillar

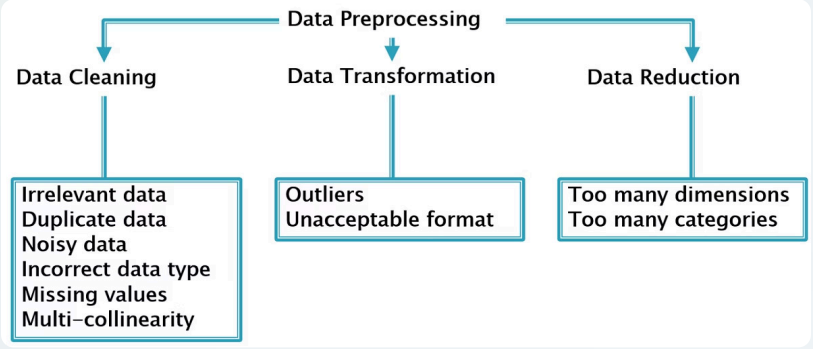
Agriculture is a key sector in Kerala's economy, supporting livelihoods and food security.

Productivity Gaps

We observe significant differences in agricultural output across various districts.

Our Objective

To understand the underlying factors contributing to these productivity disparities.



Data & Sources: Foundations for Analysis

Our investigation is built upon comprehensive agricultural data, capturing the intricate details of Kerala's farming landscape.

District-wise Data

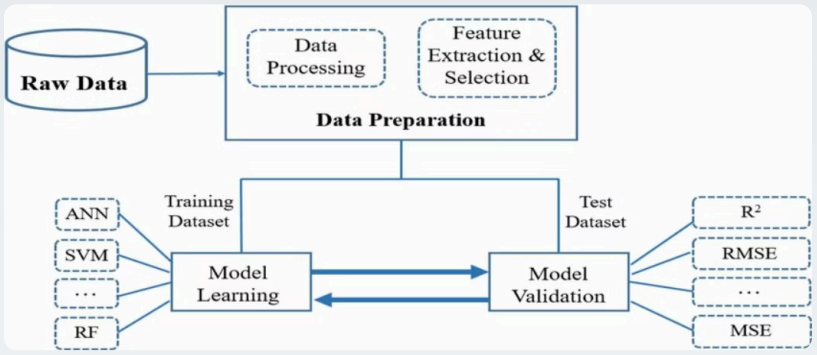
Crop area, production, and yield by district.

Key Agricultural Attributes

Crop type, soil type, and seasonal variations.

Data Preprocessing

Handling missing values, encoding categorical data, and normalization for consistency.



Tools & Methodology: Our Analytical Framework

A robust toolkit and a systematic workflow guided our data-driven exploration.

Tools Utilised

Python with Pandas, NumPy, Matplotlib, Seaborn, and Scikit-Learn.

Workflow

From initial data cleaning to advanced machine learning predictions.

01

Data Cleaning

Ensuring data quality and integrity.

02

EDA

Exploratory Data Analysis to uncover initial insights.

03

Clustering

Grouping districts based on similar characteristics.

04

ML Prediction

Developing models to forecast yield levels.

Exploratory Data Analysis: Key Insights Emerged

Our initial deep dive into the data revealed critical patterns and relationships governing agricultural output.

1

Varied Production

Significant differences in crop production across various districts were observed.

2

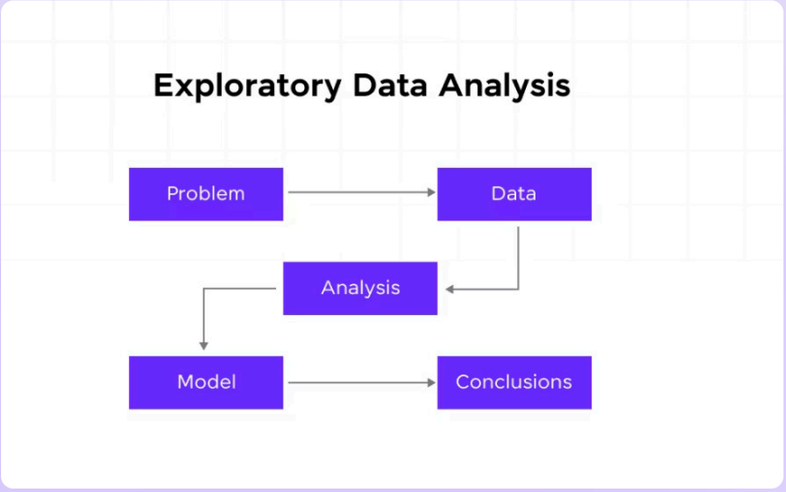
Area-Production Correlation

Generally, larger cultivation areas correlated with higher production, though exceptions exist.

3

Crop Diversity Benefits

Districts with greater crop diversity showed improved yield stability and resilience.

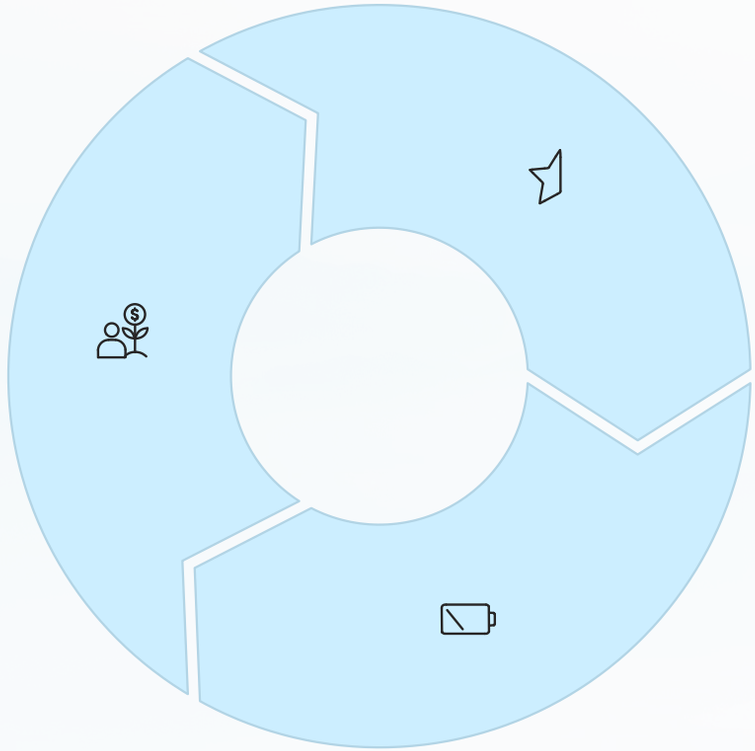


Clustering Results: Grouping Districts by Performance

Using K-Means clustering, we categorised districts into distinct performance groups, offering a clearer picture of regional strengths and weaknesses.

High-Performing

Districts with consistently high yields and efficient practices.



Moderate Performers

Districts with average yields, indicating potential for improvement.

Low Performers

Districts struggling with lower yields, requiring targeted interventions.

PCA for Visualisation: Principal Component Analysis was employed to visually separate and validate these clusters effectively.

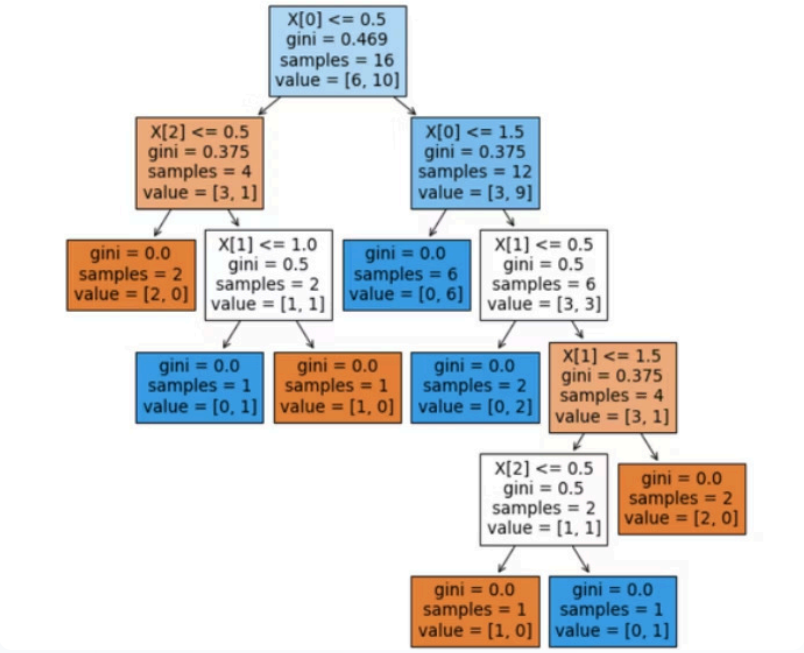
Machine Learning Model: Predicting Yield Levels

Our Decision Tree Classifier was trained to predict agricultural yield, identifying the most influential factors.

Decision Tree Classifier

A supervised learning model chosen for its interpretability and effectiveness in classification tasks.

- Predicts yield levels: Low, Medium, or High.
- Identifies key drivers of productivity.
- Provides actionable insights for policy-making.



Area

Production

Soil Type

Crop Type

These features emerged as the most significant determinants of yield variation.

Key Insights & Conclusion: Charting the Path Forward

Our analysis reveals that efficiency, not just sheer land size, drives agricultural success, opening doors for data-driven interventions.

Efficiency is Key

Productivity hinges on efficiency and optimal resource utilisation, not merely land area.

ML for Policy

Machine Learning models can inform crop planning and government agricultural policies.

Future Directions

Developing forecasting models and interactive dashboards for real-time insights.

Impact & Recommendations

Our findings provide a roadmap for enhancing agricultural productivity and fostering sustainable growth in Kerala.

Targeted Interventions

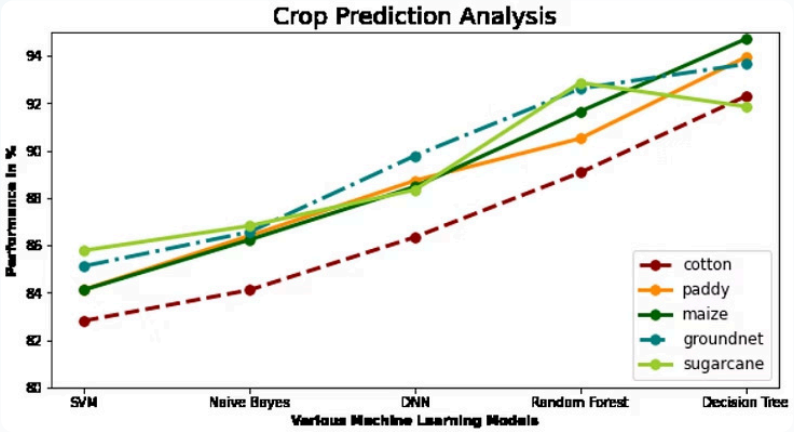
Utilise clustering results to implement district-specific support programs for low-performing areas.

- Improved seed distribution
- Modern irrigation techniques
- Soil nutrient management

Promoting Crop Diversity

Encourage farmers to adopt a variety of crops to enhance yield stability and environmental resilience.

- Government subsidies for diverse crops
- Awareness campaigns for farmers



Thank You

For further inquiries or collaboration opportunities, please feel free to connect.

Sujal S Acharya

Data Science Student

Symbiosis Institute of Technology, Nagpur

Dr. Piyush Chauhan

Guide

Symbiosis Institute of Technology, Nagpur