Robust-Yield-Prediction-on-various-Farm-Processing-Units

Robust Yield Prediction Development Plan.

Gadhasu Akhil – Emp.ID: 4880

Introduction:

Agriculture is the practice of cultivating crops for food and other essential resources. While agriculture has been a key driver of human civilization for thousands of years, it is also facing significant challenges in the modern era.

Biggest Problem* - Global population continues to grow & demand for food rises.

Thus heavily relies on data and technology to improve productivity, efficiency, and sustainability.

Data science, which involves the use of statistical methods, ML algorithms, and other analytical tools to extract insights from large datasets, can play a vital role in advancing agriculture.

Data analysis can then be used to optimize planting strategies, fertilizer application, and irrigation, resulting in higher yields and reduced input costs. By analyzing data on soil health, water use, and other factors, farmers can adopt practices that reduce environmental impact while maintaining profitability can improve productivity, sustainability,

By harnessing the power of data, the agriculture industry can meet the challenges of the 21st century and ensure a secure and stable global food supply.

Problem Statement:

A new fast-food chain is seeing rapid expansion over the past couple of years. They are now trying to optimize their supply chain to ensure that there are no shortages of ingredients.

Goal: Build a model that could predict the output of each food processing farm over the next few years. These predictions could further increase the efficiency of their current supply chain management systems.

Objectives:

a) Explore the data and engineer new features.

b) Predict the yield for each farm during the given timestamps.

c) Given the forecasted demand for the next few months for a particular ingredient, device a strategy to source it, creating a strategy to source ‘ing_w’ ingredient type.

Approach/Flow Chart:

- The First stage is understanding the data. There are 5 curated datasets. The datasets consist of Crop Information mainly around different types of ingredients, Farm Data and Field Management Information and also consists information of the weather data provided by timestamp for each location where the farms are present.
- Data Cleaning and new features were introduced in the data for better model
- Exploratory Data Analysis was performed, and eventually all the Insights and Observations are validated to see if the approach is feasible. In addition to this, the initial search also included grouping the variables, plotting relations in the weather data, measures of various weather fields were also defined.
- The Second stage was the Data Preparation. When preparing the data, the data was initially cleaned and merged which means that their information regarding farming companies, type of ingredients, timestamp, weather data and more information regarding the problem were stored. After all the necessary data was extracted correctly, the data was synthesized in order to provide it to an ML model.
- In the Final stage, Regression model was built error metric (RMSE) is evaluated checked for any performance tuning and with best hyperparameters model is build and used to make the yield predictions for next year.

Tools and Libraries used:

- Jupyter notebooks
- Python
- Libraries used : Pandas
  Seaborn
  Matplotlib.pyplot
  Sklearn
  Klib

Results and Discussions

- Using Random Forest Regressor, the training Score and test Score were 98.56% and 89.50% respectively.

Evaluation metric used is the Root Mean Squared Error (RMSE): 312.035.

- Using Gradient Boosting the scores are 85.07% and 82.18% respectively.

Test RMSE Score: 406.551917096756.

- Using RF Regressor with Hyperparameters the scores are 90.99% and 84.82% respectively.

Test RMSE_Score: 375.25472220732246

Conclusion:

- The project has successfully demonstrated the use of Random Forest Regressor in the development of Robust Yield Prediction model on a dataset consisting of various parameters related to the obtaining of expected yield. The results were very realistic in nature.
- Artificial intelligence (AI) has become increasingly important in the agriculture industry as a tool to improve efficiency, sustainability, and productivity.

  There are several ways AI is being used in agriculture - Precision agriculture, Crop monitoring, Yield prediction, Supply chain management, Autonomous vehicles.
- Conclusion: AI is transforming the agriculture industry, enabling farmers to make more informed decisions and improve productivity, sustainability, and profitability. As the technology continues to develop, we can expect even more innovative applications in the future.