Predicting appropriate crops based on environmental conditions using ML models.

Approach.

The dataset consists of environmental factors that the crop would grow on as well as the nutrient details of that crop. To identify the crops that would grow on certain environmental conditions the model should be trained with environmental factors rather than nutrient factors. Because if nutrients factors were also used in training the model would learn from those factors rather than learning from environmental factors. So, nutrients factors were removed from the data set.

Nutrient factors can be used to determine if considered here are in good condition. This was observed in EDA part and was able to identify that all the data points of any crop seem to have similar nutrient properties. Meaning all seem to be in good condition.

Rainfall viable was removed and log rainfall was kept in the dataset.

The best three crops wad finally predicted using model's prediction probabilities. Top three labels with highest probabilities were takes as predictions.

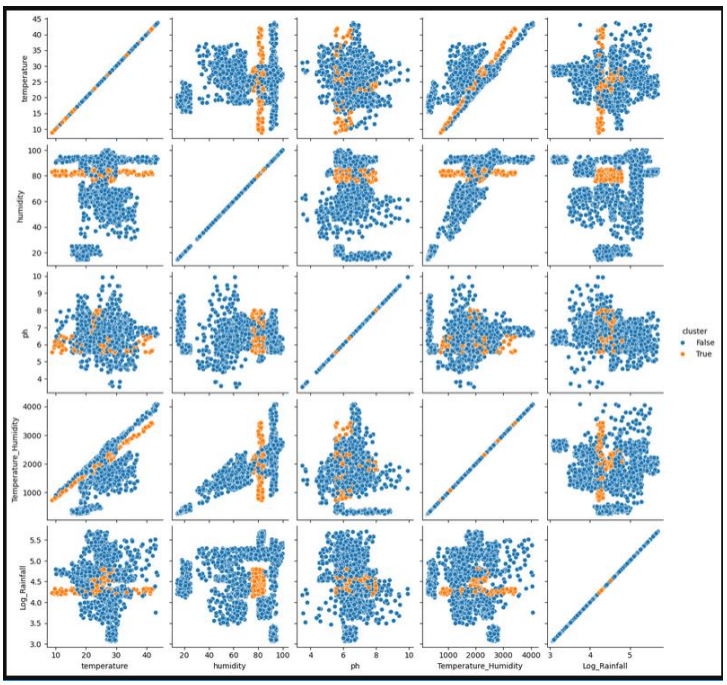
Challenges faced.

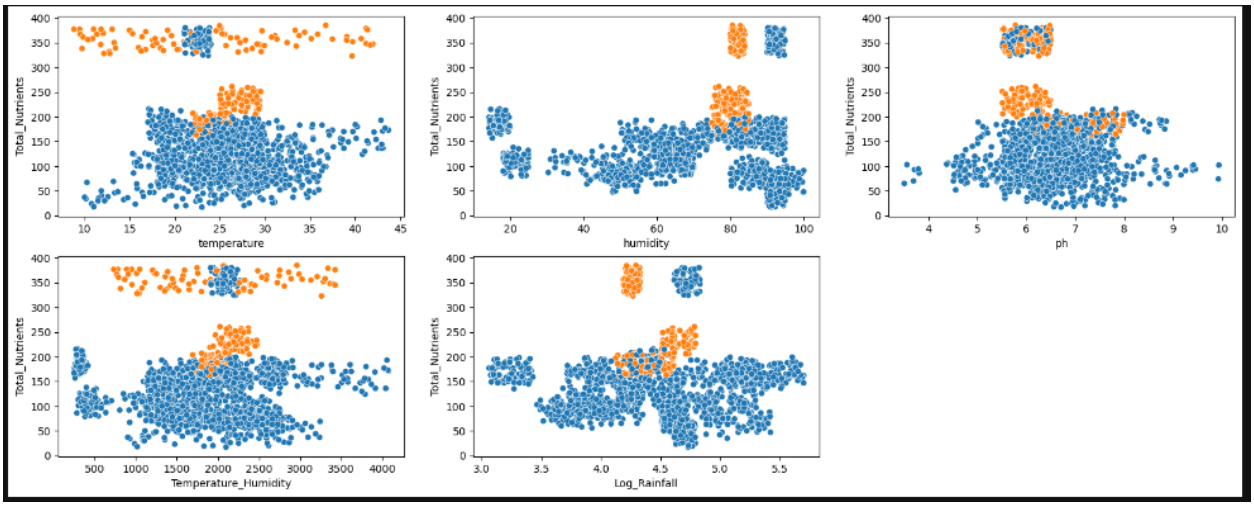
- Finding a method to identify 3 best crops from the model.
- Dealing with multiclass prediction with 22 unique labels.
- Evaluating model performance.

Insights gained from the model.

Here the predictions made through the test set was considered. Out of the 440 test predictions. 66 unique 3 crop combinations were identified. Then one of them was selected and considered as a separate cluster from the rest of the crop labels. Then

it was visualized in scatter plots between environmental variables. A clear clustering effect was identified between the selected cluster of three crops and the rest. This indicated that those crops grow on similar environmental conditions.





Suggestions for improving the model.

Use clustering algorithms to find cluster of crops that grow on similar environments. Then use cluster-based modelling for make better predictions.