Name-Shubham Kumar Modi, 20je0933, Mine Automation and Data Analytics

Agricultural Production Optimizing Engine

Introduction

In this Model I use a dataset which is used to recommend the crop for the suitable soil. This will be very useful in crop production (Agriculture). The different parameters that I used to predict the most suitable crop to be planted here are Nitrogen(N), Phosphorous(P), Potassium(Pt), Temperature, pH and Rainfall. Like

Nitrogen is one of chemical elements that become a part of amino acids.These chemical compounds are utilized by plants to increase the production and quality of crops.

Phosphorus plays a major role in the growth of new tissue and division of cells. Plants perform complex energy transmissions, a function that requires phosphorus.

Potassium is an abundant mineral macronutrient present in both plant and animals tissues. It is necessary for the proper functionality of all living cells.

Temperature: Germination increases in higher temperatures – up to a point. Once the seeds reach optimum temperatures, which depends on the plant, germination begins to decline.

The pH range **5.5–6.5** is optimal for plant growth as the availability of nutrients is optimal.

Besides disease, **rainfall** can also determine how fast a crop will grow from seed, including when it will be ready for harvesting. A good balance of rain and proper irrigation can lead to , which can cut down on germination time and the length between seeding and harvest.

Objective of the Work

The Objective of the work is to find the suitable crop that should be planted under the given parameters i.e. nutrients present in the soil like Nitrogen ,Phosphorous ,Temperature ,pH ,rainfall at the particular area to predict the type of crop suitable for the given conditions.

Methodology:

The Machine Leaning algorithms used for the predictions are K Nearest Neighbours, Support Vector Machine, Logistic Regression and Random Forest Classifier.

Before applying and model, the data is first analysed for the presence of any non-numerical values and any outliers. Once that is verified, the instances of the models are made and applied on the 70% of the available data. The other 30 percent is used for testing.

In KNN we cluster the data points around a centre representing the base of the 21 classes. With every iteration the centre of each datapoint is re attributed to the closest centre and is given that class ID. After a few iterations when the attributes stop changing the training is stopped.

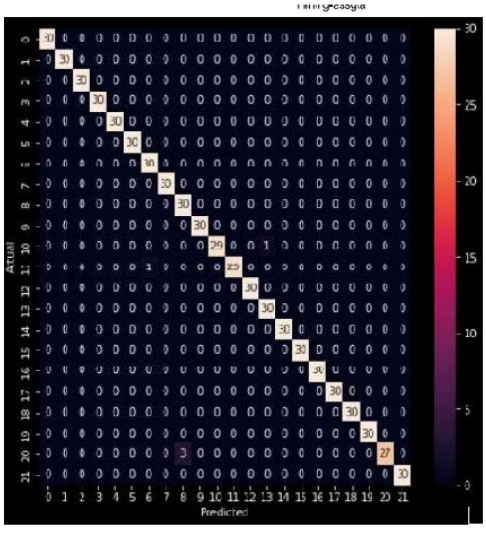
For Multiclass SVM, a range of linear and non-linear models are being applied to divide the sample space into pockets representing the classes

Logistic regression simply works based on whether the prediction output of a datapoint meets a threshold amount to be classified in a category

Random Forest is an ensemble model where a series of questions are asked based on their priority and depending on their output the range of targeted class is narrowed and finally a single class is attributed to a datapoint.

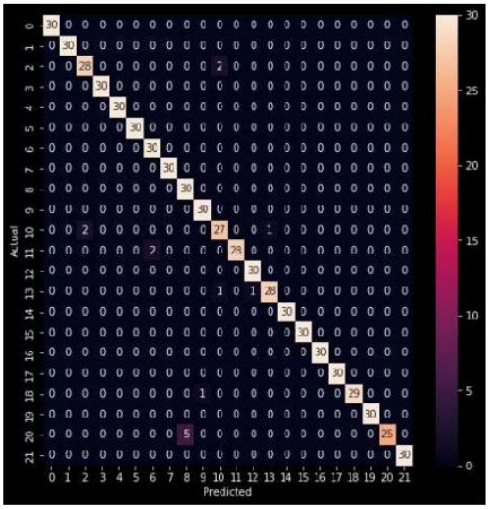
Result

accuracy			0.99	660
macro avg	0.99	0.99	0.99	660
weighted avg	0.99	0.99	0.99	660

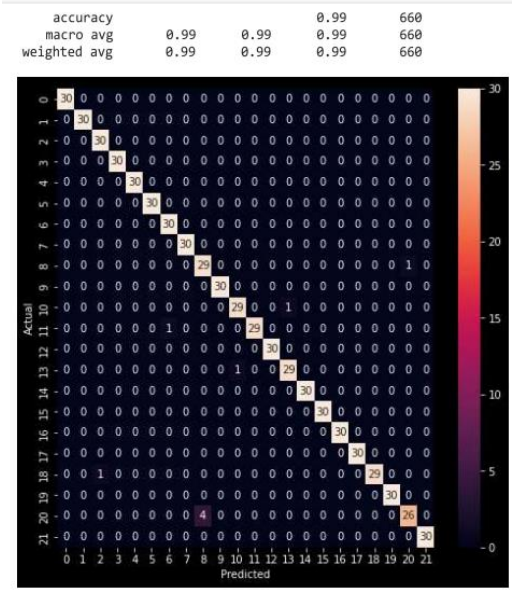


Support vector Classifier 99%

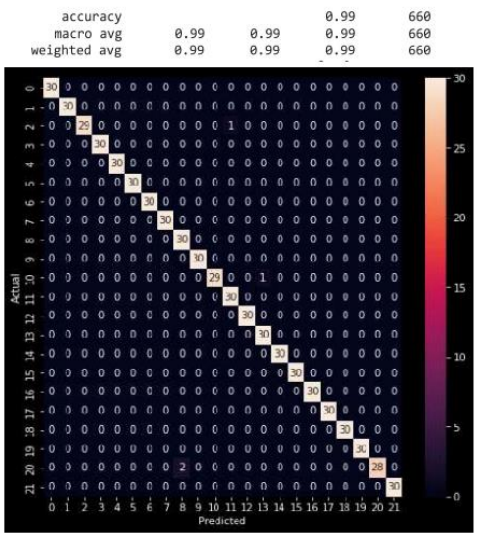
accuracy			0.98	660
macro avg	0.98	0.98	0.98	660
weighted avg	0.98	0.98	0.98	660



KNeighborsClassifier -98%



Linear Regression- 99%



Random Forest Classifier - 99%

From the observation of different data, I get the most appropriate crop that should be planted to get the maximum production by growing the crop. Like I got 99% accuracy in SVM, Linear Regression, Ramesh Forest Classifier & in KNN I got 98% accuracy.

Conclusion

From this prediction and classification I come to conclude that every crop need different amount of nutrients, water requirements and temperature to grow properly, through this model we can able to predict the crop which is most suitable in the given climatic and environmental conditions.

Reference

[GitHub - D-2204/Agricultural-Production-Optimization-Engine: This Project focuses on building a predictive model so as to suggest the most suitable crops to grow, based on the nature of soil and the given climatic conditions.]