

Crop prediction and Efficient fertilizer Recommendation using Machine Learning

SRI VENKATESWARA UNIVERSITY COLLEGE OF ENGINEERING

**BACHELOR OF TECHNOLOGY
IN
COMPUTER SCIENCE AND ENGINEERING**
by

Muktheswar Reddy .B (11716082)

Sai Sahith Reddy .M (11716090)

Swetha .K (11716097)

Under the guidance of

HOD, Prof. P. Venkata Subba Reddy



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING, SRI
VENKATESWARA UNIVERSITY COLLEGE OF ENGINEERING, TIRUPATI,
2020-2021.**

Crop Prediction and Efficient Fertilizer Recommendation Using Machine Learning

Abstract:

India has a large agricultural sector supporting the majority of the population for their livelihood. It employs almost half of India's working-class people. Every year huge resources in terms of land and fertilizer are used for the production of food. Most of the fertilizer being used in farms today is being wasted due to the incorrect use of a type of fertilizer. We are developing a machine learning model that analyses various soil features like Calcium, Iron concentration, pH of the soil, and various other features to recommend the type of fertilizer to use and also the type of crop needed for a particular area for maximum benefit.

Introduction:

This project aims to take in features like annual rainfall, pH of the soil, humidity, the temperature of a particular land area and train a machine learning model for accurate prediction of the type of crop to be sown to get a good yield from the crop. We also build a model to analyze soil features and recommend the type of fertilizer to be used.

Related Work:

Niketa Gandhi et al, "Rice Crop Yield Forecasting of Tropical Wet and Dry Climatic Zone of India Using Data Mining Techniques", IEEE International Conference on Advances in Computer Applications (ICACA), 2016.

Jharna Majumdar, Sneha Naraseeyappa and Shilpa Ankalaki, Analysis of agriculture data using data mining techniques: application of big data, Springer journal, 2017.

Proposed System:

DATA SET:

The dataset used for fertilizer recommendation is soil features dataset. It contains various features of soil like Calcium concentration, Magnesium concentration etc. The dataset contains a total of more than 16500 data points. It has 9 features.

Ca	Mg	K	S	N	Lime	C	P	Moisture	class
0.7	0.6	0.8	0.8	0.7	0.8	0.3	0.1	0.9	4
0.5	0.5	0.4	0.3	0.5	0.7	0.5	0.7	0.8	2
0.6	0.8	0.1	0.3	0.7	0.5	0.5	0.6	0.6	2
0.7	0.7	0.7	0.5	0.8	0.7	0.4	0.1	0.7	4
0.8	0.8	0.2	0.3	0.5	0.5	0.7	0.8	0.5	2
0.7	0.5	0.6	0.7	0.8	0.8	0.3	0.4	0.8	4

DATA SET

The dataset used for crop recommendation is the land cover dataset which includes features such as temperature, annual rainfall, pH of the soil, and humidity. It has a target variable indicating the suitable crop. The dataset has a total of more than 15000 data points.

temperatu	humidity	ph	rainfall	label
20.8797	82.0027	6.50299	202.936	rice
21.7705	80.3196	7.0381	226.656	rice
23.0045	82.3208	7.84021	263.964	rice
26.4911	80.1584	6.9804	242.864	rice
20.1302	81.6049	7.62847	262.717	rice
23.058	83.3701	7.07345	251.055	rice

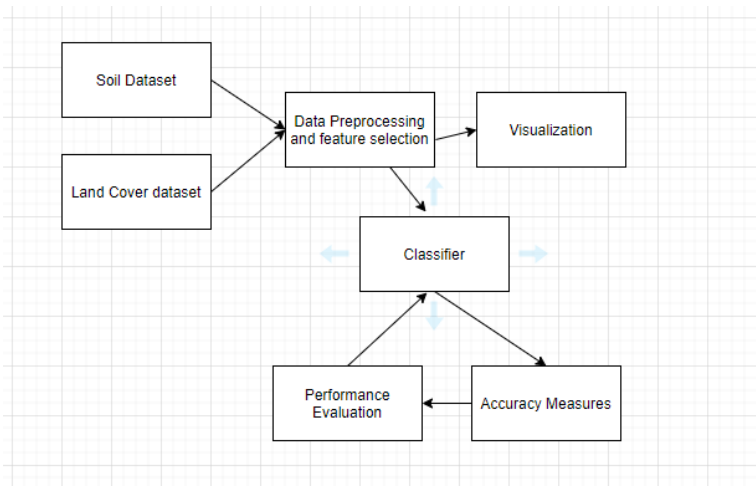
Packages required:

- ◆ Pandas
- ◆ Numpy
- ◆ Scikit-Learn
- ◆ Matplotlib

Data Preprocessing:

This process includes methods to remove any null values or infinite values which may affect the accuracy of the system. The values are also rescaled for faster training of the models. Steps in preprocessing include outlier detection, missing value treatment, rescaling. The cleaning process is used for the removal or fixing of some missing data there may be incomplete data.

Architecture Diagram:



Model Selection and Performance Evaluation:

After feature selection, the dataset is divided into pairs of xtrain ,ytrain and xtest, y test. The algorithm model is imported form sklearn. Building model is done using model. Fit (xtrain, ytrain).

Crop type and fertilizer type are our output (target). Since we are going to classify using continuous features, we are going to implement using Decision tree regression, support vector regression, and also neural network-based classifier.

After the model is built, prediction is done using the model.predict(xtest). The accuracy is calculated using accuracy_score imported from metrics - metrics.accuracy_score (ytest, predicted).

Conclusion:

We select the models with the highest possible accuracy and the accuracy is measured with various metrics like K-fold, precision, recall. We intend to develop this model so that it is beneficial for farmers in decision-making for good farm produce.