**Soil Analysis And Agriculture Crop Cultivation prediction**

1. Introduction:

Agriculture is a source of income for a large population in India. With the increase in the demand for production and the industrialization over the time, there has also been a huge reduction in the number of farmlands. In order to make accurate decisions based on the type of crops to be planted and to achieve a good harvest, data such as use of pesticides, fertilizers, meteorological and soil information must be made available to the farmers in an accurate and timely manner. Soil analysis is a valuable farm practice that determines the exact amount of available crop nutrients in the soil.​ It provides a visible snapshot of various chemical, physical, and biological soil properties.​ Better crop productivity can be achieved by farmers through analysis of the suitable conditions, thereby reducing the damage and loss of crops that occur due to unfavorable conditions. Decline of traditional techniques of soil management, terrestrial limitations, and population stresses have caused deterioration in the soil fertility especially in developing countries like India.

In recent days, several machine learning (ML) algorithms and models are used for effectively addressing the classification and prediction issues. The challenges faced by experts in the agricultural domain are also overcome largely by the introduction of ML techniques. A private online repository is used where real-time data is collected from the soil and crop. The values in the dataset are used for training and validating the model and then assessing and computing the results. For this purpose, the dataset is categorized into training data and testing data while building the model.

Algorithms used:

* Numpy:

NumPy is a Python library used for working with arrays.​ NumPy stands for Numerical Python​. In Python we have lists that serve the purpose of arrays, but they are slow to process.​ It aims to provide an array object that is up to 50x faster than traditional Python lists.

* Pandas:

Pandas is used to analyse data. It has functions for analysing, cleaning, exploring, and manipulating data. Pandas allows us to analyze big data and make conclusions based on statistical theories.​

* Matplotlib:

Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python.​ Its numerical mathematics extension NumPy.​ It provides an object-oriented API for embedding plots into applications using general purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK.​

* Seaborn​:

Seaborn is an amazing visualization library for statistical graphics plotting in Python.​ It provides beautiful default styles and color palettes to make statistical plots more attractive. ​It is built on the top of matplotlib library and also closely integrated to the data structures from pandas.

* Sklearn:

Scikit-learn is probably the most useful library for machine learning in Python. ​The sklearn library contains a lot of efficient tools for machine learning and statistical modelling including classification, regression, clustering and dimensionality reduction.​Simple and efficient tools for data mining and data analysis.

​

* XGBoost

​XGBoost is an open-source Python library that provides a gradient boosting framework.​It helps in producing a highly efficient, flexible, and portable model.​ When it comes to predictions, XGBoost outperforms the other algorithms or machine learning frameworks due to its accuracy and enhanced performance and ​reduces model errors.

Data Set Availability: <https://www.kaggle.com/code/theeyeschico/crop-analysis-and-prediction/data>

According to the analysis that we will perform, the available micronutrients are classified and the soil samples are categorized .

Several algorithms will be used:

* K-nearest neighbor
* Support Vector Machine Algorithm
* Random Forest Algorithm
* Decision Tree Algorithm

**References:**

* <https://www.academia.edu/46014132/Soil_Analysis_and_Crop_Fertility_Prediction_using_Machine_Learning>​
* <https://www.hindawi.com/journals/jnm/2022/5343965/>​
* Data set: <https://www.kaggle.com/code/theeyeschico/crop-analysis-and-prediction/data>​