Precision Agriculture: Tailored crop solutions through Soil and Climate Assessment

Harsh Raj(21Bec1699)
Sakshi(21BEC1568)
Kumar Harsh(21BEC1697)
Guide: Dr. Upender Patri

Outline of the project

- Introduction
- Problem Statement
- Objective
- Literature Survey
- Proposed System(Introduction)
- Proposed System(Diagram)
- List of Modules
- What is to be done next?
- References

Introduction

- Precision agriculture and its relevance to soil mapping, and prediction of crop based soil properties are highlighted in the introduction .
- The study focuses on precision agriculture and soil fertility management, specifically looking at Nitrogen,Phosphorus ,Pottasium and soil pH, texture for precise crop prediction .
- The expertise of the authors in areas such as infrared spectroscopy and geostatistics is mentioned .
- Funding for the study is acknowledged, and the importance of digital soil mapping in agriculture is emphasized .
- The introduction sets the stage for the study's focus on sensor-based soil mapping, sampling design optimization, and prediction model comparison for precision agriculture applications

Problem Statement

- The goal here is to create a system that recommends suitable crops to farmers based on various parameters. These parameters include soil quality, previously cultivated crops, rainfall, humidity and ph value of soil.

Objective

- **Optimizing Crop Selection**:
  - A crop recommendation system aims to **maximize yield and profit** by suggesting crops that are well-suited to the specific soil, climate, and other environmental conditions of a particular farm.
  - It considers factors such as soil type, pH levels, nutrient content, temperature, and rainfall patterns.
- **Reducing Risk and Uncertainty**:
  - By providing accurate recommendations, the system helps farmers reduce the risk of crop failure or poor yields.
  - It minimizes the uncertainty associated with choosing crops based solely on intuition or tradition.
- **Sustainable Agriculture**:
  - Crop recommendations take into account sustainable practices, such as crop rotation and soil health management.
  - The system encourages diversification and balanced land use, promoting long-term sustainability.
- **Efficient Resource Utilization**:
  - Proper crop selection ensures efficient utilization of resources like water, fertilizers, and pesticides.
  - Farmers can avoid overuse or underuse of inputs, leading to cost savings and environmental benefits.

Literature Survey

1.Improving soil fertility with lime and phosphogypsum enhances soybean yield and physiological characteristics

- The paper focused on the effects of soil amendments, specifically lime and phosphogypsum, on soybean crops in Brazil. Various parameters were studied, including crop sowing, meteorological data, soil analysis, root dry matter, nutrient concentrations in leaves and grains, gas exchange parameters, and physiological analysis of plant leaves. The results consistently showed significant improvements in nutrient availability, soil quality, root growth, and plant physiological parameters with the application of lime and phosphogypsum.
- The combined application of lime and phosphogypsum led to enhanced soil pH, nutrient availability, root growth, and plant nutrition compared to individual amendments. This improvement in soil conditions resulted in increased soybean shoot dry matter production, grain yield, and crude protein content in grains. Additionally, the amendments reduced oxidative stress and increased antioxidant enzyme activities in soybean plants, indicating improved stress tolerance.
- Overall, the study highlighted the importance of soil management practices, particularly the long-term application of lime and phosphogypsum, in enhancing crop productivity, nutrient uptake, and overall soybean production in tropical regions with unpredictable weather conditions. The findings suggest that these soil amendments can be a promising strategy for sustainable soybean production in tropical no-till systems, especially in areas prone to dry spells.

2.Assessing the performance of machine learning algorithms for soil salinity mapping in Google Earth Engine platform using Sentinel-2A and Landsat-8 OLI data

The paper focused on utilizing machine learning algorithms, specifically Random Forest (RF), Classification and Regression Trees (CART), and Support Vector Regression (SVR), to predict soil salinity in the Urmia Lake region. The RF model outperformed the other algorithms, accurately categorizing different salinity classes. The study employed a k-fold cross-validation method with k=5 to ensure optimal results. Comparisons with other studies using RF, CART, and SVR algorithms for soil salinity prediction in various regions showed varying levels of accuracy, with the RF algorithm consistently providing the most accurate estimation of soil salinity in the Urmia Lake region.

- The research emphasized the importance of integrating field measurements with remote sensing (RS) data for improved mapping of salt-affected lands. It highlighted the Random Forest method as superior for predicting soil salinity classes compared to traditional methods. The study also utilized cloud computing platforms like Google Earth Engine to facilitate the analysis of environmental issues. By assessing the performance of machine learning algorithms (CART, RF, SVR) in the Urmia Playa Lake Basin using Sentinel-2A and Landsat-8 OLI data, the study found that RF yielded the most reliable results. The study area's socio-economic and ecological significance led to the collection and laboratory analysis of soil samples for electrical conductivity (EC), aiming to provide accurate and rapid soil salinity mapping for improved monitoring and management of salt-affected lands.
- Moreover, the paper utilized soil samples analyzed in previous research to generate soil salinity maps using Landsat-8 OLI and Sentinel-2A satellite images. Environmental covariates were leveraged to predict soil salinity, and the study conducted variable reduction to select the most relevant predictors. Overall, the RF, CART, and SVR algorithms provided accurate predictions for soil salinity mapping. The references in the paper collectively highlighted the significance of soil salinity mapping using remote sensing and machine learning techniques, emphasizing comparisons of different algorithms, spatial prediction of soil salinity, and the utilization of satellite data like Landsat and Sentinel-2 for addressing soil salinity issues in agriculture and the environment.

3.Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel-2 spectral imaging

The paper focused on predicting soil properties, specifically soil organic carbon (SOC), clay, silt, and sand content, using different remote sensing platforms such as laboratory spectroscopy, hyperspectral airborne sensors, and superspectral spaceborne sensors in various locations including New Hampshire, USA, and the Czech Republic.

1. **Prediction Performance**:

   - Laboratory data provided good results for all properties except silt in one location.

   - Hyperspectral airborne data showed good prediction accuracy for SOC and clay in most locations.

   - Superspectral satellite data had good results for SOC but lower accuracy for other properties.

   - Sand and silt prediction was particularly challenging with satellite data.

2. **Comparison of Platforms**:

- In the Czech Republic, the study found that the superspectral Sentinel-2 satellite showed potential for estimating and mapping SOC and clay over large areas, but not silt and sand content.

- Sentinel-2 imagery could accurately predict SOC and clay content, but not silt and sand content, showing slightly less precision than lab spectroscopy and airborne sensors.
- The spatial distribution maps of SOC and clay derived from different platforms showed consistent trends, with Sentinel-2 satellite data performing well in differentiating various classes of SOC and clay.

3. **Importance of Platform Selection**:

- The results highlighted the importance of platform selection for accurate soil attribute prediction, with clay content being more accurately predicted due to its spectral response.
- While satellite data like Sentinel-2 may have limitations in predicting certain soil properties like silt and sand, its extensive coverage and frequent revisit characteristics make it a valuable tool for monitoring soil attributes.
- Overall, the study demonstrated the potential of remote sensing platforms in predicting soil properties, emphasized the importance of accurate soil property mapping for various applications in agriculture and environmental monitoring, and called for further research to address limitations in using spaceborne data for soil assessments.

4.Use of quick and hydrated lime in stabilization of lateritic soil: comparative analysis of laboratory data.

The paper compared the performance of soil stabilized with quicklime and hydrated lime, specifically focusing on lateritic soil. Quicklime consistently showed superior performance in various aspects compared to hydrated lime. Quicklime resulted in lower plasticity, higher strength, lower swell values, better durability, and slightly lower hydraulic conductivity. In terms of engineering properties, quicklime exhibited lower plasticity, higher dry unit weight, and higher compressive strength compared to hydrated lime. Both types of lime were effective in reducing hydraulic conductivity and showed similar results in swelling potential. Overall, the study recommended quicklime as a more effective stabilization alternative for lateritic soil based on its superior performance in most characterizations. The experiments were conducted by AAA who developed the concept and prepared the manuscript, while AO performed the experiments. No competing interests were declared, and the study referenced various research on lime stabilization of soils.

5.Impact of silicate and lime application on soil fertility and temporal changes in soil properties and carbon stocks in a temperate ecosystem.

- The paper analyzed the impact of different soil treatments, including silicate, calcium, and a combination of both, on soil fertility indicators over a 67-year period. The results indicated that soil pH increased in all treated soils, with the combined treatment showing the highest increase. Soil organic carbon content varied among treatments, with calcium resulting in the highest increase. Available phosphorus decreased in all treated soils, with the combined treatment showing the largest decrease. Exchangeable potassium, magnesium, and calcium increased in all treated soils, with the combined treatment showing the highest increase in calcium. Available silicate concentration increased in all treated soils, with the silicate treatment showing the highest increase.
- Furthermore, the study found that the combined application of silicate and lime led to increased soil pH and fertility but decreased carbon stocks. The combined treatment resulted in the highest positive soil degradation index, indicating improved soil quality. The relative change ratios varied among treatments, with silicate treatments showing lower improvement in soil organic carbon but higher improvement in potassium content.
- Additionally, the study highlighted that the combined application of silicate and lime had the most significant effects on soil properties and fertility improvement, although it also reduced carbon stocks. The combined treatment showed the highest impact on soil organic carbon stocks and led to the highest improvement in soil properties such as calcium, SiO_2, and Mg. The annual growth rate of the soil fertility index was highest in the combined silicate and lime treated soils.
- Overall, the paper suggests that the combined application of silicate and lime is the most effective in improving soil fertility and organic carbon content over the long term. However, it also points out the need for further research to understand the long-term effects on microbial communities and carbon lability, as well as the importance of co-applying silicate and lime with soil amendments that can stabilize and sequester soil carbon.

Granular and powdered lime improves soil properties and maize (Zea mays l.) performance in humic Nitisols of central highlands in Kenya

The paper investigated the impact of lime treatments on soil properties and maize productivity in two study sites with clay soils in Kenya. The application of lime resulted in an increase in soil pH and a decrease in exchangeable acidity, which are beneficial for maize production. Powdered lime was

found to be more effective than granular lime in increasing soil pH and decreasing exchangeable acidity. Additionally, the application of powdered $CaCO_3$ lime significantly increased soil available phosphorus, exchangeable calcium, and magnesium, although there was no significant change in total organic carbon and potassium levels. Treatments with fertilizer, either alone or in combination with lime, led to the highest maize yields in both study sites.

- The study suggests further research on the long-term effects of different lime types and emphasizes the importance of informed soil management practices for sustainable agricultural development in the region. Overall, the results highlight the significance of lime application in improving soil quality and crop productivity in acidic Nitisols in Kenya.
- The experiment was conducted in Kirege and Kangutu in Meru South Sub-County of Tharaka Nithi County, Kenya, which are characterized by bimodal rainfall patterns. The study sites received different cumulative rainfall, with Kirege being extremely acidic and experiencing higher rainfall compared to the moderately acidic Kangutu. Drought spells occurred during the cropping season in both sites, affecting the water availability for the maize crop.
- The experimental design involved a randomized block design with eight treatments replicated four times per site, including control, fertilizer only, sole lime, and fertilizer plus lime treatments. The liming materials used in the experiment included powdered lime, which showed significant improvements in soil properties and maize productivity.
- The authors of the paper contributed to designing and conducting the experiments, analyzing and interpreting the data, and writing the paper. Data associated with the study has been deposited at the Kenyatta University Repository.
- In conclusion, the study demonstrates the positive effects of lime treatments on soil properties and maize productivity in acidic clay soils in Kenya, highlighting the potential for improved agricultural practices in the region.

Parameter	Quicklime (CaO)	Hydrated Lime (Ca(OH)2)	Study Findings
Dosages (%)	0, 2.5, 5, 7.5, 10	0, 2.5, 5, 7.5, 10	Dosages were kept the same for a one-to-one comparison [12].
Plasticity	Lower plasticity (1.4 times lower at 10% treatment)	Higher plasticity	Quicklime reduces soil plasticity more effectively [1].
Strength (UCS)	Higher strengths (approx. 50% higher at 10% treatment)	Lower strengths	Quicklime yields higher unconfined compressive strength, especially at higher dosages [12].
Swell Values	Lower swell values (2.2 times lower at 10% treatment)	Higher swell values	Quicklime results in lower swell values [1].
Durability	Marginal superiority (108% durability for mixtures)	Slightly lower (96% durability for mixtures)	Quicklime shows slightly better durability [1].
Dry Unit Weight	Not specified	Higher dry unit weight	Hydrated lime yields a higher dry unit weight [2].
Hydraulic Conductivity	Not specified	Not specified	Both quicklime and hydrated lime were used in tests for hydraulic conductivity [2].

Curing Time	Cured for 28 days	Cured for 28 days	Both types of lime were cured for the same duration 2.
pH Effect	Increases pH, aiding in forming cementitious compounds	Increases pH, aiding in forming cementitious compounds	Both types increase pH, which is crucial for the stabilization reaction 12.
Environmental Sustainability	Not specified	May be combined with curing agents for cost-effective and environmentally sustainable stabilization 3.	Hydrated lime can be used with a curing agent to reduce lime usage and enhance sustainability 3.

Proposed System

The proposed system in the "Crop Reccomendation System" can be defined in chronological order as follows:

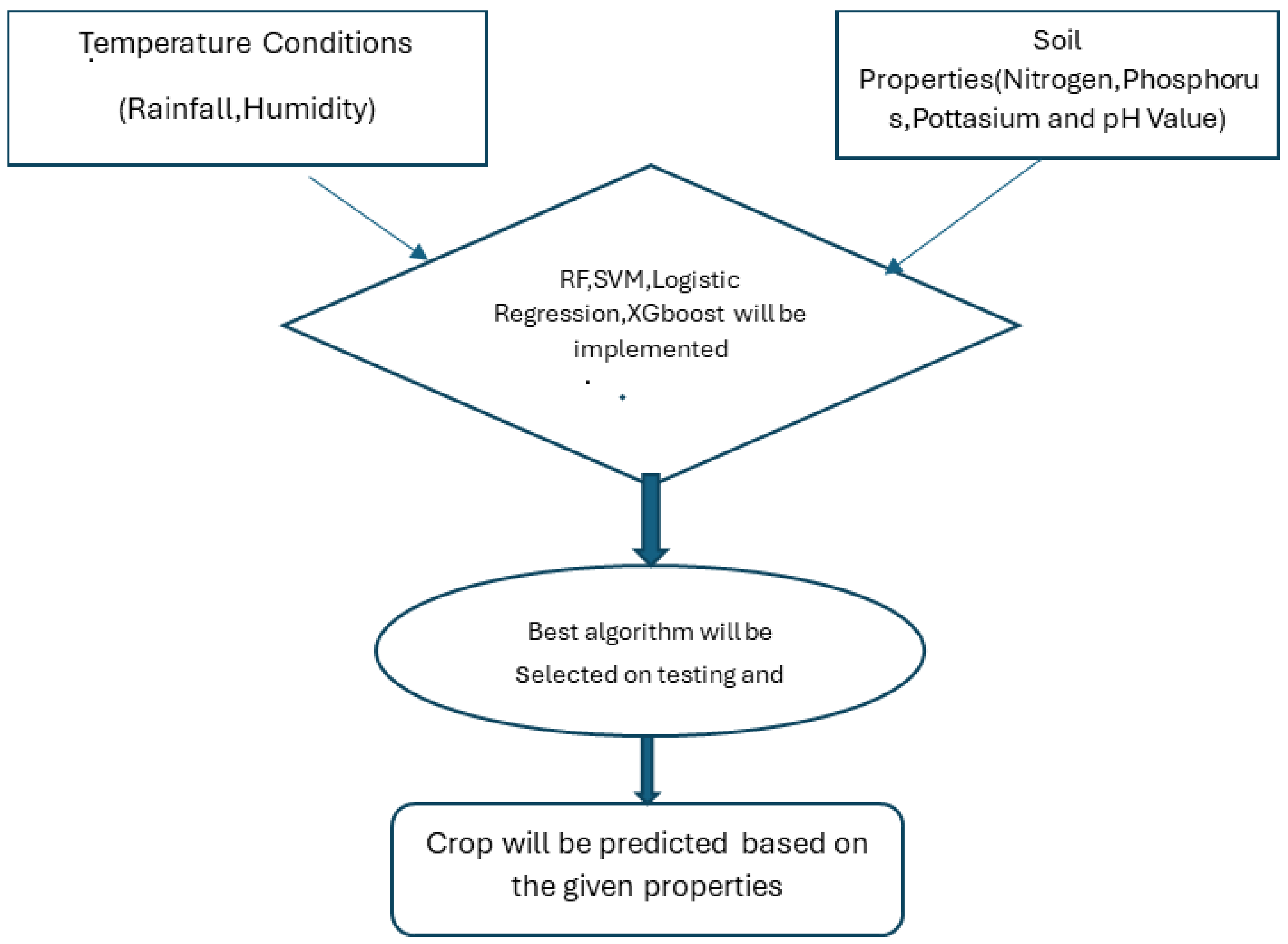
Dataset of different crops based on soil properties, rainfall, humidity and pH value of soil will be collected.

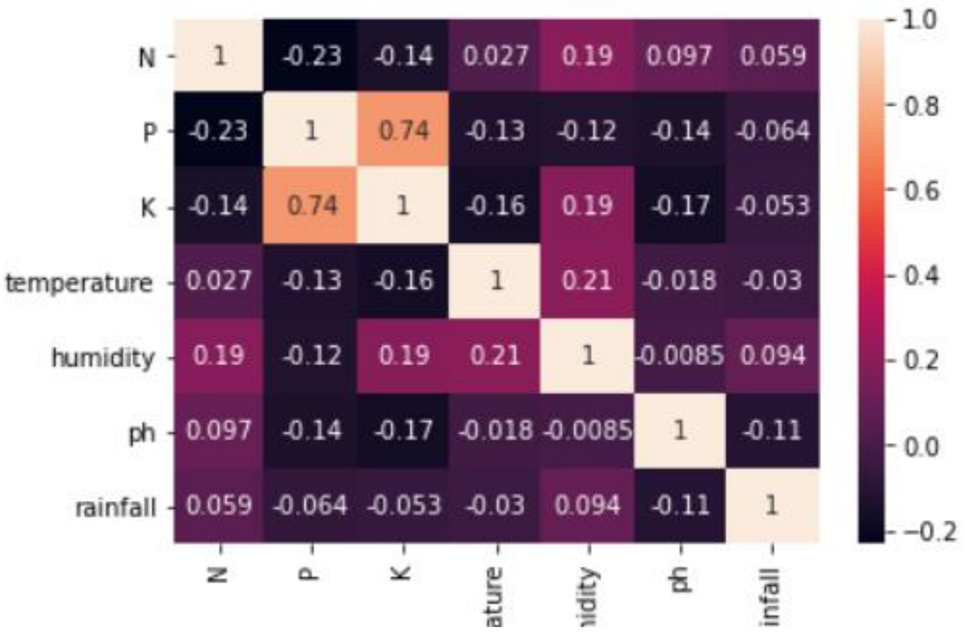
Based on the dataset ,We will perform different Machine learning and Deep Learning algorithms .

On Comparison of different models, We will determine the Best suit ed algorithm for our dataset.

The algorithm that we will be selecting will be predicting the Crop suited for the given conditions.

Proposed System Diagram





The correlation Graph shows that phosphorus and pottasium is 74% related.

```
from __future__ import print_function
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report
from sklearn import metrics
from sklearn import tree
import warnings
import plotly.graph_objects as go
import plotly.express as px
from plotly.subplots import make_subplots
warnings.filterwarnings('ignore')
df=pd.read_csv('/content/Crop_recommendation.csv')
df.head()
print('A sample set of rows for dataframe is:\n')
display(df.sample(6))
```

- **Importing libraries and dataset:**

Data Preprocessing :

```
df.isnull().sum()
```

```
N              0
P              0
K              0
temperature    0
humidity       0
ph             0
rainfall       0
label          0
dtype: int64
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2200 entries, 0 to 2199
Data columns (total 8 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   N            2200 non-null   int64
 1   P            2200 non-null   int64
 2   K            2200 non-null   int64
 3   temperature  2200 non-null   float64
 4   humidity     2200 non-null   float64
 5   ph           2200 non-null   float64
 6   rainfall     2200 non-null   float64
 7   label        2200 non-null   object
dtypes: float64(4), int64(3), object(1)
memory usage: 137.6+ KB
```

```
df.describe()
```

	N	P	K	temperature	humidity	ph	rainfall
count	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000
mean	50.551818	53.362727	48.149091	25.616244	71.481779	6.469480	103.463655
std	36.917334	32.985883	50.647931	5.063749	22.263812	0.773938	54.958389
min	0.000000	5.000000	5.000000	8.825675	14.258040	3.504752	20.211267
25%	21.000000	28.000000	20.000000	22.769375	60.261953	5.971693	64.551686
50%	37.000000	51.000000	32.000000	25.598693	80.473146	6.425045	94.867624
75%	84.250000	68.000000	49.000000	28.561654	89.948771	6.923643	124.267508
max	140.000000	145.000000	205.000000	43.675493	99.981876	9.935091	298.560117

```
df.size
```

```
17600
```

```
df.shape
```

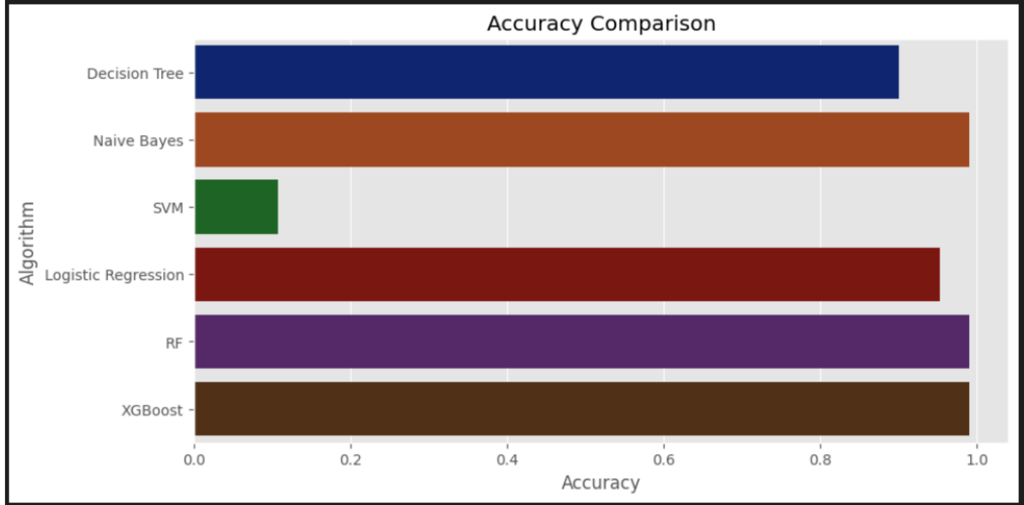
```
(2200, 8)
```

```
df.columns
```

```
Index(['N', 'P', 'K', 'temperature', 'humidity', 'ph', 'rainfall', 'label'], dtype='object')
```

Comparison of different models:

```
<Axes: title={'center': 'Accuracy Comparison'}, xlabel='Accuracy', ylabel='Algorithm'>
```




```
Decision Tree --> 0.9
Naive Bayes --> 0.990909090909091
SVM --> 0.10681818181818181
Logistic Regression --> 0.9522727272727273
RF --> 0.990909090909091
XGBoost --> 0.990909090909091
```

Output/Results:

```
(variable) prediction: Any
, 23.603016, 60.3, 6.7, 140.91]])
prediction = RF.predict(data)
print(prediction)
```

```
['coffee']
```

```
data = np.array([[83, 45, 60, 28, 70.3, 7.0, 150.9]])
prediction = RF.predict(data)
print(prediction)
```

```
['jute']
```

List of Modules

The modules utilized in the proposed system are as follows:

- **Random Forest:**During the training phase, Random Forest creates a **number of Decision Trees**.Each tree is constructed using a **random subset of the dataset** to measure a random subset of features in each partition..This randomness introduces variability among individual trees, reducing the risk of **overfitting** and improving overall prediction performance.In prediction, the algorithm **aggregates the results of all trees**: For **classification tasks**, it uses **voting** (combining the predictions of individual trees).

  For **regression tasks**, it uses **averaging** (combining the predicted values of individual trees).

- **Logistic Regression**: **Logistic regression** estimates the **probability of an event occurring** based on a given dataset of independent variables. It's often used for **classification** and **predictive analytics**. The model analyzes the relationship between two data factors and predicts the output of a **categorical dependent variable**.Instead of giving exact values (like 0 or 1), logistic regression provides **probabilistic values** that lie between 0 and 1

- **XGBoost :** XGBoost is an optimized distributed gradient boosting library designed to be highly efficient, flexible, and portable .It falls under the family of **boosting algorithms**, which are ensemble learning techniques that combine the predictions of multiple weak learners (usually decision trees) to create a strong predictive model .XGBoost is widely used for **supervised learning tasks**, such as regression and classification.
- **SVM :Support Vector Machine (SVM)** is a powerful **supervised machine learning algorithm** used for both **classification** and **regression** tasks. Its primary goal is to find the **optimal hyperplane** in an N-dimensional space that can **separate data points into different classes**.SVMs are versatile and efficient, making them suitable for various applications, including text classification, image classification, spam detection, and more.
- **Decision tree:** A **decision tree** is a powerful tool used in **supervised learning** for both **classification** and **regression** tasks. It constructs a flowchart-like tree structure where each internal node represents a test on an attribute, each branch corresponds to an outcome of the test, and each leaf node (terminal node) holds a class label or a numerical value
- **Naïve Bayes:** Naive Bayes classifiers are a collection of algorithms based on Bayes' Theorem. Although not a single algorithm, they all share a common principle: every pair of features being classified is independent of each other. These classifiers are particularly useful for text classification, where data often has high dimensionality (each word represents a feature). Naive Bayes is commonly used in tasks like spam filtering, sentiment detection, and rating classification due to its speed and ease of prediction.

What is to be done next?

Some future improvements in the field of sensor-based precision agriculture and soil mapping could include

We can develop a web app for our proposed system to help the farmers in a more efficient way.

Exploring the effectiveness of different prediction models, such as machine learning techniques like Random Forest, to enhance prediction accuracy for mapping soil properties in agricultural fields.

Determining the total cost for implementation of web app and sensors needed for collecting the dataset.

Developing clear guidelines on the ideal number of training samples needed for sensor-based precision agriculture applications to ensure well-fitted prediction models without unnecessary costs.

These future improvements can contribute to advancing the field of sensor-based precision agriculture and soil mapping, leading to more accurate and efficient soil property

REFRENCES:

- Adamchuk, V. I., Morgan, M. T., & Lowenberg-Deboer, J. M. (2004). A model for agro-economic analysis of soil pH mapping. Precision Agriculture, 5, 111–129. https://doi.org/10.1023/B:PRAG.0000022357. 28154.eb
- Adamchuk, V. I., Viscarra Rossel, R. A., Marx, D. B., & Samal, A. K. (2011). Using targeted sampling to process multivariate soil sensing data. Geoderma, 163, 63–73. https://doi.org/10.1016/j.geoderma. 2011.04.004
- Bertsimas, D., & Tsitsiklis, J. (1993). Simulated annealing. Statistical Science, 8(4), 10–15. https://doi.org/ 10.1214/ss/1177011077
- Biswas, A., & Zhang, Y. (2018). Sampling designs for validating digital soil maps: A review. Pedosphere, 28, 1–15. https://doi.org/10.1016/S1002-0160(18)60001-3
- Bönecke, E., Meyer, S., Vogel, S., Schröter, I., Gebbers, R., Kling, C., Kramer, E., Lück, K., Nagel, A., Philipp, G., Gerlach, F., Palme, S., Scheibe, D., Zieger, K., & Rühlmann, J. (2021). Guidelines for precise lime management based on high-resolution soil pH, texture and SOM maps generated from proximal soil sensing data. Precision Agriculture, 22, 493–523. https://doi.org/10.1007/ s11119-020-09766-8

REFRENCES:

- Breiman, L. (2001). Random forests. Machine Learning, 45, 5–32. https://doi.org/10.1023/A:10109 33404324
- Brus, D. J. (2019). Sampling for digital soil mapping: A tutorial supported by R scripts. Geoderma, 338, 464–480. https://doi.org/10.1016/j.geoderma.2018.07.036
- Brus, D. J., de Gruijter, J. J., & van Groenigen, J. W. (2006). Designing spatial coverage samples using the k-means clustering algorithm. Developments in Soil Science, 31, 183–192. https://doi.org/10. 1016/S0166-2481(06)31014-8
- Castro-Franco, M., Costa, J. L., Peralta, N., & Aparicio, V. (2015). Prediction of soil properties at farm scale using a model-based soil sampling scheme and random forest. Soil Science, 180, 74–85. https://doi.org/10.1097/SS.0000000000000115
- Chen, S., Arrouays, D., Mulder, V. L., Poggio, L., Minasny, B., Roudier, P., Libohova, Z., Lagacherie, P., Shi, Z., Hannam, J., Meersmans, J., Richer-de-Forges, C. A., & Walter, C. (2022). Digital mapping of GlobalSoilMap soil properties at a broad scale: A review. Geoderma, 409, 115567. https://doi.org/10. 1016/j.geoderma.2021.115567

THANK YOU!