

CHAPTER I

INTRODUCTION

Agriculture plays a very important role in national economic growth. Agriculture contributes 17-18% to India's GDP and ranks second worldwide in farm outputs. Plants need fertilizers and fertilizers replace the nutrients which crops take from the top layer of the soil. The absence of fertilizers can cause a drastic reduction in the volume of crop output. But fertilization requires precise action. Rainfall patterns and the amount of nutrients needed for a certain crop must be considered when using fertilizers. Machine learning is the current technology that can solve this problem by using available data for crop fertility and rainfall. Farmers can greatly benefit from the support of robust information about crops. The proposed model also uses a machine learning algorithm (random forest algorithm with k-fold cross-validation technique) and takes two inputs from the user that are crop and location. After applying the algorithm, the model predicts the amount of nutrients required along with the best time to use fertilizers. The website is built using Flask Python (web framework) to provide access on all platforms and can be shared among users.



Fig. 1 On-Ground Analysis

1.1 Machine Learning

The study of machine learning (ML) focuses on comprehending and developing "learning" techniques, that employ data to enhance performance on a variety of tasks. It is considered a component of artificial intelligence. Without the need for explicit programming, machine learning algorithms create a model from training data and use it to generate predictions or judgments. When it is difficult or impossible to develop conventional algorithms, many industries, including speech recognition, medicine, computer vision and email filtering use machine learning techniques. However, machine learning goes beyond simple statistical learning. Computational statistics, is a branch of machine learning that focuses on computer-based prediction. Mathematical optimization research benefits machine learning by providing tools, theory, and application fields.

What makes machine learning so crucial?

Machine learning is significant because it aids in the development of new goods and provides businesses with a picture of trends in consumer behavior and operational business patterns. Many of the largest companies operating in the world today, like Uber, Facebook, and Google, heavily rely on machine learning. Machine learning has become a major point of competitive difference for many firms.

What are the various kinds of machine learning?

How an algorithm learns to improve its prediction accuracy is a common way to classify traditional machine learning. There are four fundamental strategies: reinforcement learning, semi-supervised learning, unsupervised learning, and supervised learning. The kind of data that data scientists wish to predict determines the kind of algorithm they use. Different kinds of machine learning algorithms are:

- *Supervised learning:* With this kind of machine learning, data scientists feed algorithms labelled training data and tell them specific variables to seek for correlations between. Both the input and the output of the algorithm are given.
- *Unsupervised learning:* This kind of machine learning employs algorithms that train on unlabeled data. Data sets are searched by the algorithm for any significant relationships. Both the forecasts or suggestions generated by algorithms and the data used to train them are predefined.

- *Semi-supervised learning*: The above two forms of machine learning are combined in this method. An algorithm may be fed mostly labeled training data by data scientists, but the algorithm is allowed to analyze the data on its own and come to its own conclusions about the data set.
- *Reinforcement learning*: Data scientists typically use reinforcement learning to program a machine to complete a multi-step procedure according to pre-established norms. Data scientists design an algorithm to achieve a goal, giving it valuable feedback as it decides how to do so. The algorithm often makes the decision by itself, though.

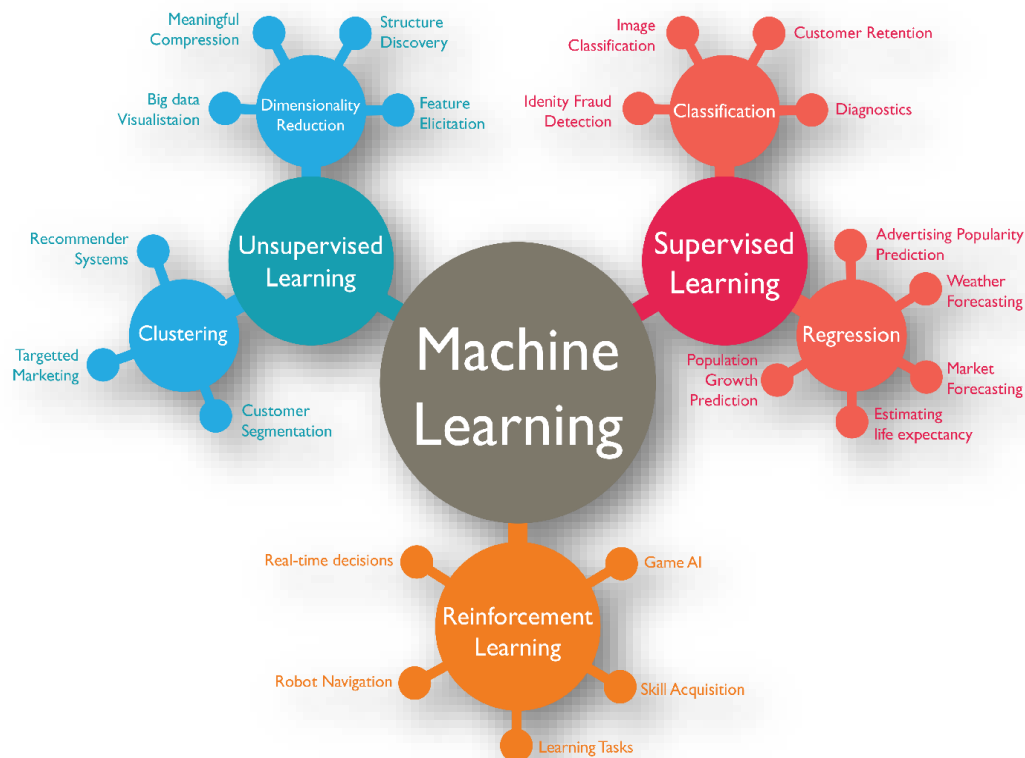


Fig 1.1 Machine Learning Classification

1.2 Plant Root Loss

Farmers face a number of challenges every year like, heavy rainfall or unpredictable weather, including high interest rates, an overreliance on traditional crops, and a lack of water. Crops may be submerged in water as a result of flooding, which could cause catastrophic losses. As soon as a plant is submerged, its foliage starts to deteriorate because its leaves are unable to exchange gases with the air above (mainly oxygen & carbon dioxide). Producers face flooding or continuously flooded soil more frequently, which hinders roots' ability to absorb nutrients. If the soil is completely soaked for an extended period of time, root loss may result. Because they can't exchange gases, root cells in waterlogged soils risk dying.

Depending on how long the soil is entirely soaked, different amounts of root loss may occur. Plant mortality and complete crop failure would ensue from total root loss. Lower plant performance and crop output would result from partial root loss. Conditions that are too moist might have other detrimental effects on crop productivity. Unusually excessive rainfall can wash away nutrients from the soil, particularly nitrogen. Granular fertilizer that has been put to the soil as nitrogen is particularly susceptible to leaching. If this happens, farmers would either have to pay more money to reapply fertilizer or see a decrease in crop yield due to nitrogen shortage.

1.3 Soil Leaching

Leaching is the downward transport of pollutants via porous soils, such as water-soluble pesticides or fertilizers. The majority of pesticides, notably clay, adhere to soil particles, become stationary, and do not drain. However, the multiple degradation mechanisms and leaching to groundwater can be seen as competitors in the fate of mobile pesticides. Groundwater does not continuously dilute the pollutants that enter it, in contrast to surface water. It could take many years to remove a contaminated plume from groundwater. Chemical deterioration is slowed by the soil's depth, the freezing temperatures, the limited microbial activity, the lack of sunlight, and the low oxygen levels. As a result, once pesticides enter an aquifer, there is little to no degradation, if any at all. This leads to water pollution.

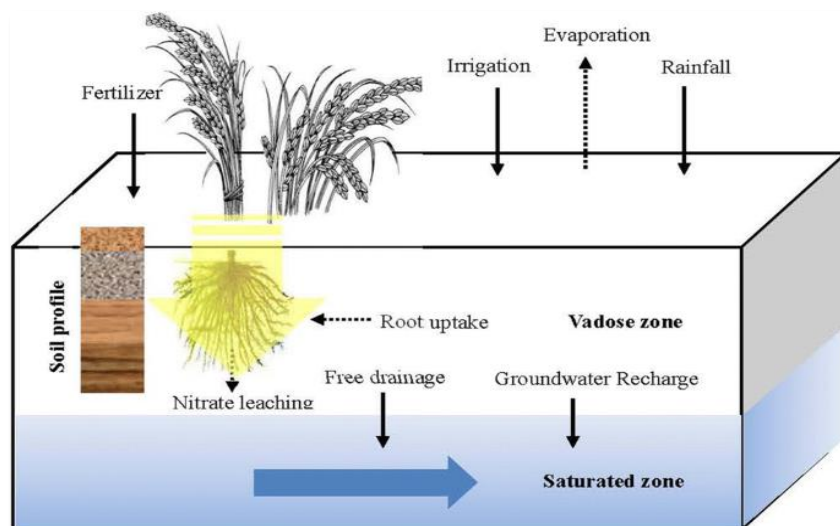


Fig 1.3 Soil Leaching

Soil Features that Influence Leaching:

- *Organic Matter:* The amount of organic matter in soil is thought to be the single factor that has the greatest impact on how microorganisms break down pesticides. Pesticides are less likely to leach into groundwater because the organic matter in the soil improves the surface area available for adsorption, boosts the soil's capacity to retain water and break down pesticides, and nourishes microorganisms. Crop leftovers can be added to the soil, manure can be added, and cover crops can be grown to increase the organic matter in the soil.
- *Soil Texture:* The amounts of sand, silt, and clay in the soil have an impact on how water moves through it. Large pores and high permeability characterize coarse-textured soils with more sand particles, which allow water to pass through quickly. Groundwater is more likely to become contaminated by pesticides delivered by water via soil with a coarse texture. Soils with a clay texture have less permeability. More water is retained and more chemicals are absorbed from the water by soil with high clay content. This lessens the likelihood of groundwater pollution, increases the likelihood of chemical breakdown and binding to soil particles, and slows the downward migration of the pollutants.
- *Soil Structure:* Water can travel through the soil quickly because of loosely packed soil particles. Tightly compacted soil acts as a dam, holding water back

and preventing its free flow. Openings and channels can be made for water flow in a number of different ways. For instance, animals and earthworms generate openings for water to flow through when they dig burrows. In soil and rock, freezing and thawing cause fissures or splits that dislodge compacted particles. When plant roots decay and die, they pierce the soil and make great water routes. Even though some clay soils, these apertures and channels might allow for a somewhat quick water flow.

- *Soil Water Content:* Rain or irrigation can recharge the groundwater and perhaps cause pesticides to seep into the aquifer, depending on how much water is already present in the soil. Once soil moisture content is getting close to or near saturation, soluble substances are much more likely to enter groundwater. When it rains and there is snowmelt in the spring, saturation is normal. Contrarily, when soils are dry, the additional water simply fills soil pores close to the soil surface, decreasing the likelihood that it will contact the groundwater supply.

1.4 Purpose

The main motive behind AgriTech is to reduce farmers losses by providing useful insights about the amount and use of fertilizers, and to reduce water pollution by slowing down the process of leaching. It serves as a link between farmers and modern technology and enables them to increase yields while using less inputs. The system is designed as a website to provide platform-independent functionality, so that the user can access it from any device. The user interface has been kept simple with more emphasis on functionality and can be used by any naive user. It takes inputs such as crop, state and city using the drop-down menus provided in the website and applies machine learning algorithms to estimate the correct amount of nitrogen, potassium and phosphorus content required. This system provides good accuracy in its decision about the nutrients required for the crop.

1.5 Objectives

Crop production is essential to the global food and biofuel economies, and ML is significantly enhancing farmers' contributions on both fronts. To enhance crop productivity

and yield, herbicides, insecticides, and fungicides must all be applied at the right time. Even if crop spraying is possible later in the season as soil moisture decreases, crop yields will almost certainly be harmed. Every year, farmers make hundreds of intricate and connected decisions that affect their risk, sustainability, and financial results.

The goal of employing machine learning in our project is to provide relevant insight for nutrient requirement for crops by taking short-term weather forecasts (specifically for seven days) into account, as well as to prevent water pollution by slowing down the leaching process.

CHAPTER II

LITERATURE SURVEY

A comprehensive study of the available literature presents a catalog of previous studies to address this issue. The authors show in [1] that predicting fertilizer usage can assist farmers to attain a proper yield with little waste by preventing toxicity and deficiency in plants to some extent. Paper [2] makes use of fuzzy logic systems that enable the reduction of fertilizer usage which results in an increase in crop productivity. Additionally, [10] shows that the enhanced efficiency of fertilizers is not sufficient for complications that can be caused by compaction. These issues can be prevented by improving the fertilizer recommendation which requires the establishment of a quantifiable relation under N and P for fertilizer usage, in terms of agricultural yield, nitrogen need, and nitrate remnant level which is shown in [11] and paper [4] seconds this by providing a comprehensive measure to estimate the weightage of nutrient requirements and also the role of the chemical properties of soil.

It is a difficult task to predict crop yield due to stochastic rainfall patterns and also temperature variation. So, we can apply different data mining techniques as propounded in [3] for crop yield prediction. Laura J.T. Hess et al. in [5] state that nitrogen leaching is prone in areas that have no-till management and this may cause crop loss. In [7] the authors suggest a novel metric for ‘soil health and quality’ including refinement of soil’s health.

The objective of the paper [8] is to examine the characteristic changes in the creation and elements of soil populaces and capabilities because of the collaboration between long haul treatment and precipitation variances, to decide if preparation history affects the water-obstruction of soil microorganisms. Also, Paper [13] predicts agricultural yield as a function of rainfall. This is accomplished by giving a general summary of how production is affected by rainfall and how much a given crop can yield given the amount of rainfall received. Because it examines all regression procedures, the suggested method of evaluation is superior to other existing methods of evaluation.

Potnuru Sai Nishant et al. in paper [6] predict the yield of practically all types of crops in India. This script makes innovative use of straightforward criteria such as state, district and area, allowing the user to forecast crop yields in any year. Paper [12] suggests

the use of Transfer Learning techniques to create a pre-trained model for detecting patterns in the dataset, which we then used to predict crop yields. In [14], supervised algorithms that boost crop yields, reduce human labor, time, and energy exerted on various agricultural tasks, and plant suggestions based on particular soil parameters are used to produce a complete way to predict crop sustainability. The study [16] demonstrated the capabilities of a machine learning model that can interpret and evaluate results, can be utilized to create the most useful information in long-term fertilizer studies, and that these methods can be employed in other long-term experiments. Paper [17] develops an interesting decision-based system on climatic, crop, and insecticide/pesticide data. This is done







Senthil Kumar Swami Durai et al. in [18] propose an integrated solution to Pre-Cultivation activities. The goal of this study is to assist a small farm in becoming more efficient and achieving a high production at a low cost. It also aids in the estimation of total growth expenses. It will assist one in planning forward. Pre-cultivation activities lead to an integrated solution in agriculture. M.S. Suchithra and Maya L. Pai propose solutions to soil nutrient classification problems utilizing the rapid learning classification technique called an Extreme Learning Machine (ELM) with various activation functions in [19].

Crop diseases are one of the primary causes that impact the overall yield. Paper [15] conducts this study using an IoT system in the Kashmir Valley, it proposes an apple disease prediction model using data analysis and machine learning. The challenges of incorporating new technology into traditional agricultural practices are discussed in this paper.





CHAPTER III

SYSTEM REQUIREMENTS SPECIFICATION

3.1 Hardware Requirements

-  Processor: Intel(R) Core(TM) i3-4005U CPU @ 1.70GHz
-  RAM: 4.00 GB
-  System type: 64-bit operating system, x64-based processor
-  Network Interface Card
-  Keyboard
-  Mouse

3.2 Software Requirements

-  Operating System (any)
-  Google Chrome (web browser)
-  Visual Studio Code
-  Jupyter Notebook

CHAPTER IV

SYSTEM DESIGN

The process of defining a system's architecture, modules, components, interfaces, and data in order to meet predetermined requirements is known as system design. We could think of systems design as the applications of systems approach to the creation of products.

4.1 Architectural Design

A conceptual model known as system architecture describes the structure and behavior of the system. It consists of the system's elements and the connections between them that explain how the whole system is implemented. Fig 4.1 below shows the system's architecture and the various components added to them.

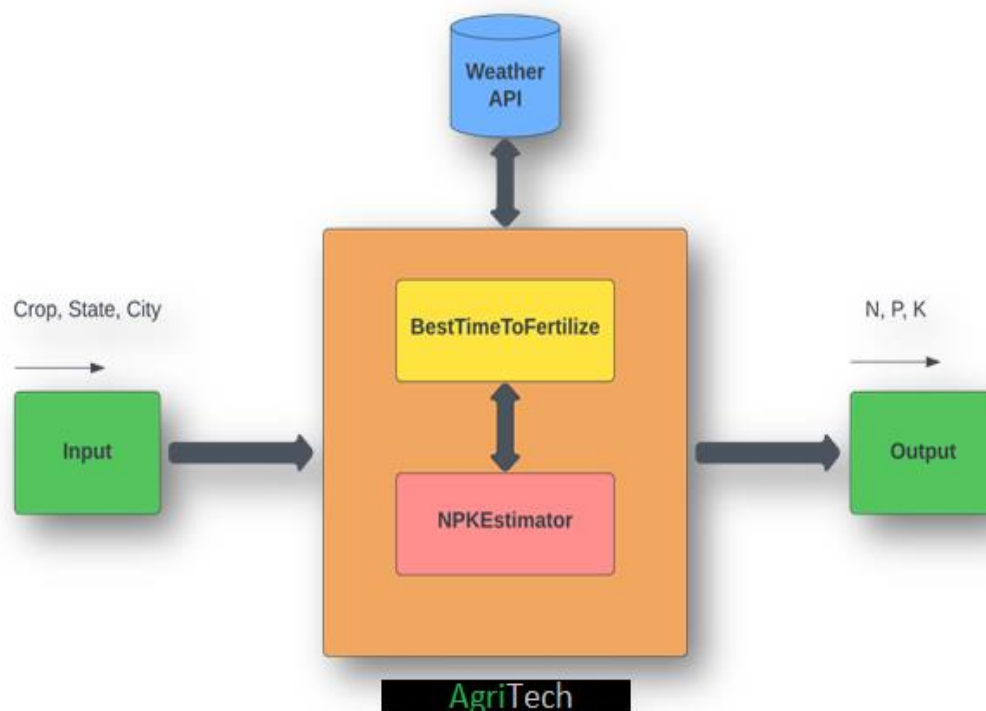


Fig 4.1 Block Diagram

The description of each component from the block diagram above and their major functionalities with respect to the AgriTech as a complete unit is described in the table below.

SI No.	Block Name	Functions
1	Input	User provides data such as crop, state and city using drop-down menu.
2	Weather API	Weather details like temerature, humidity, rainfall etc. is fetched from the weather API.
3	BestTimeToFertilize	This module provides the functionality to determine the best time to fertilize using fetched weather data and provides warning for heavy rain.
4	NPKEstimator	This module estimates the required ratio of NPK contents in the soil.
5	Output	Nitrogen, Phosphorus and Potassium content displayed on the website.

Table 4.1: Block Diagram functionalities

4.2 Data Flow Diagram

A data flow diagram is a visual representation of how data "flows" throughout a data system, simulating certain features of its operation. It is frequently used as an initial stage to develop, without going into great depth, an overview of the system that may then be expanded upon. They may also be utilized to display data processing.

The type of data that will be input into and generated by the system, how well the data will move through the system, and how the data will be kept are all displayed in a data flow diagram. Unlike a flowchart, which additionally displays information about timing and whether processes will run in succession or parallel, it does not provide this information.

As shown in Fig 4.2, the system requires input from the user (such as location and crop). The location is fed to the Weather API which will return certain characteristics (e.g. temperature, humidity, rainfall) and if there is a possibility of heavy rainfall, a precautionary message is displayed to the user, otherwise, the proposed algorithm is followed.

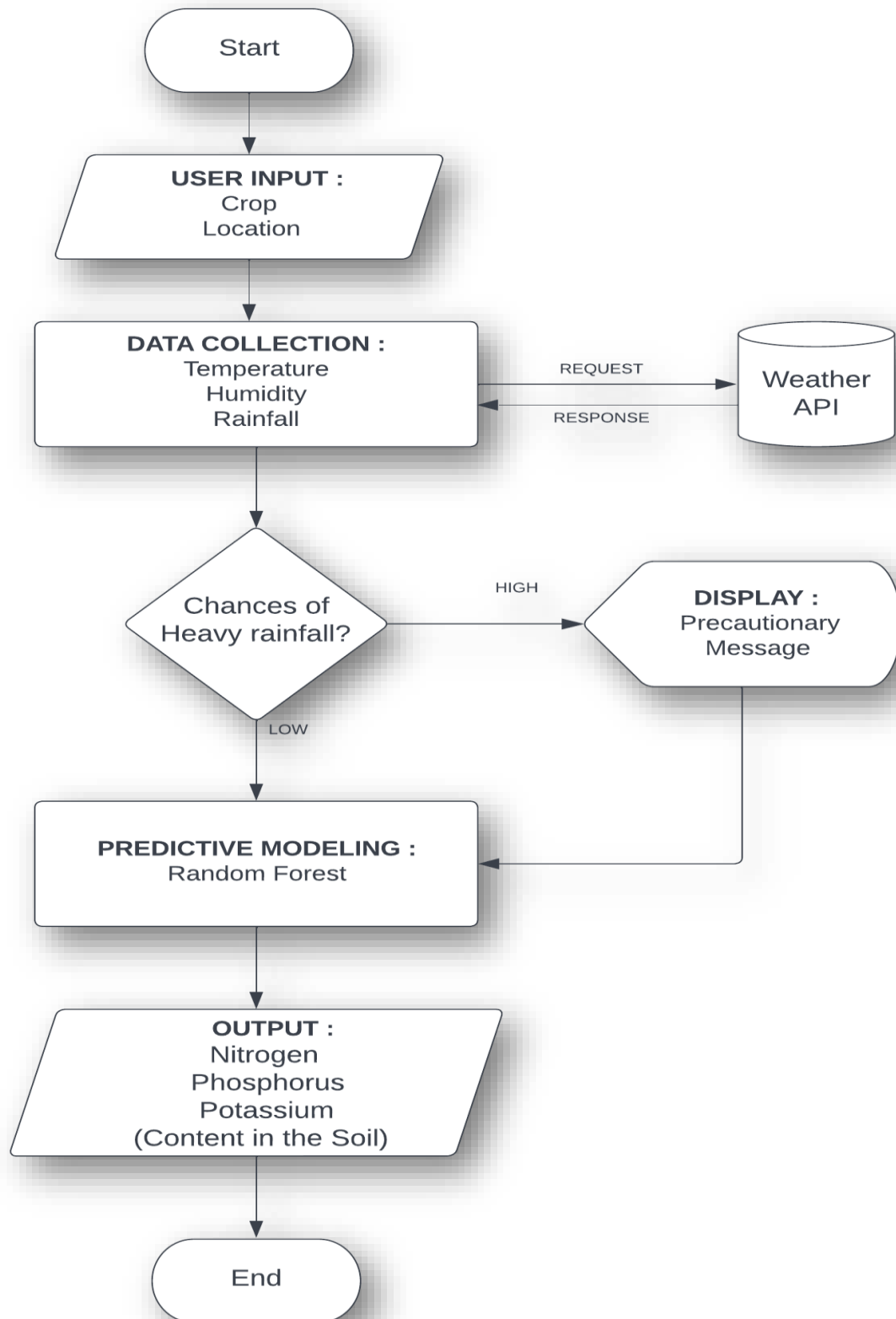


Fig 4.2 Data flow Diagram

4.3 Use Case Diagram

A **use case diagram** is a graphical depiction of a user's possible interactions with a system. A use case diagram shows various use cases and different types of users the system has and will often be accompanied by other types of diagrams as well.

The use cases are represented by either circles or ellipses. The actors are often shown as stick figures.

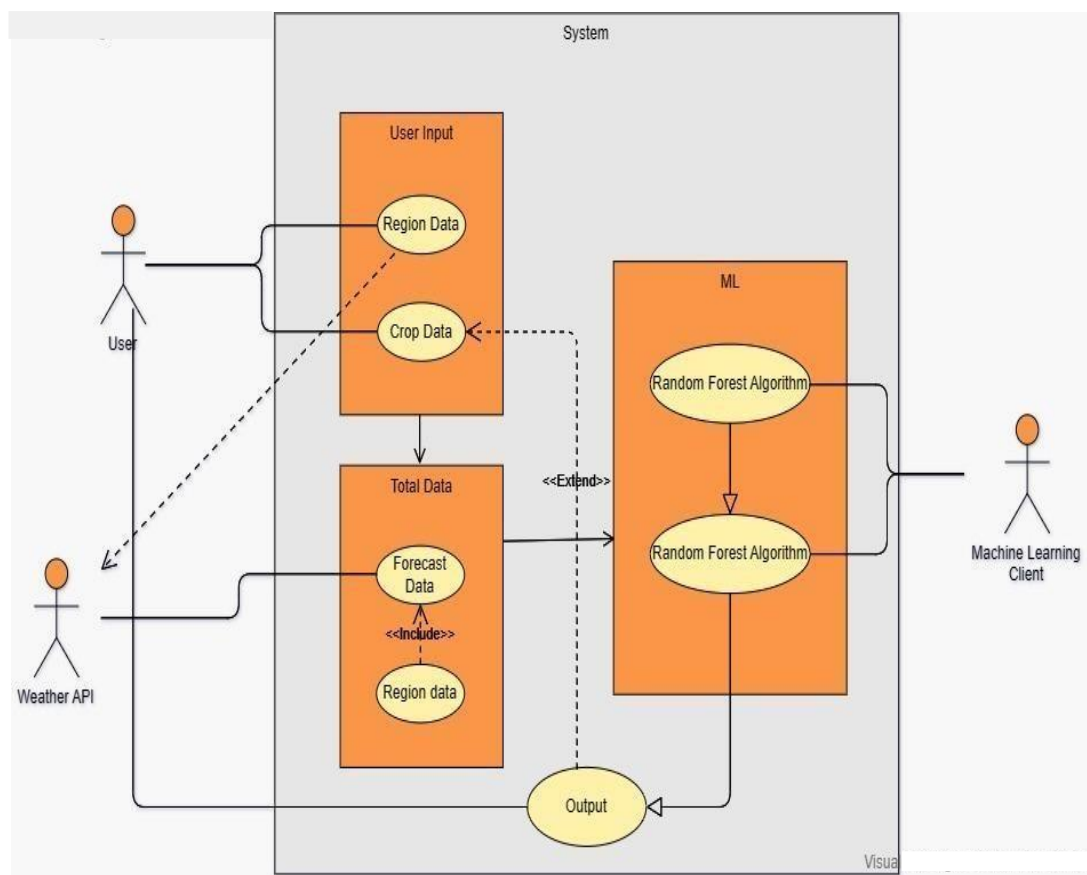


Fig 4.3 Use Case Diagram

4.3 Class Diagram

Static diagrams include class diagrams. It represents the application's static view. Class diagrams are used to create executable code for software applications as well as for visualizing, explaining, and documenting various elements of systems.

A collection of classes, interfaces, affiliations, collaborations, and constraints are displayed in a class diagram. A structural diagram is another name for it.

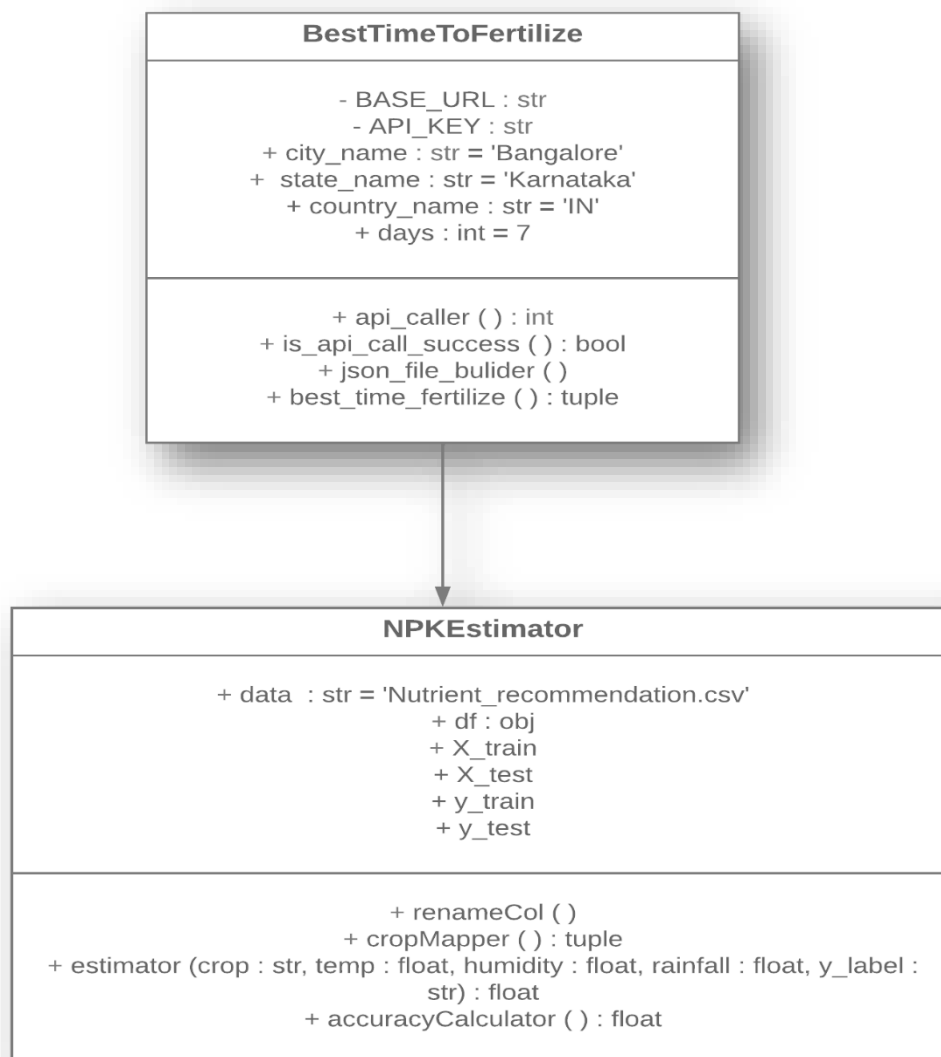


Fig 4.4 Class Diagram

4.4 Sequence Diagram

Object interactions are arranged in temporal sequence in a sequence diagram. It shows the classes and objects engaged in the scenario as well as the flow of messages that must be exchanged for the objects to work as intended. Inside the logical view of the system being developed, sequence diagrams are often connected to use case realizations. Event diagrams and event scenarios are other names for sequence diagrams.

A sequence diagram is made up of vertical parallel lines (called "lifelines") that represent several processes or things that exist at the same time and horizontal arrows that represent the messages sent between them in the chronological order in which they take place. This enables the graphical specification of straightforward runtime scenarios.

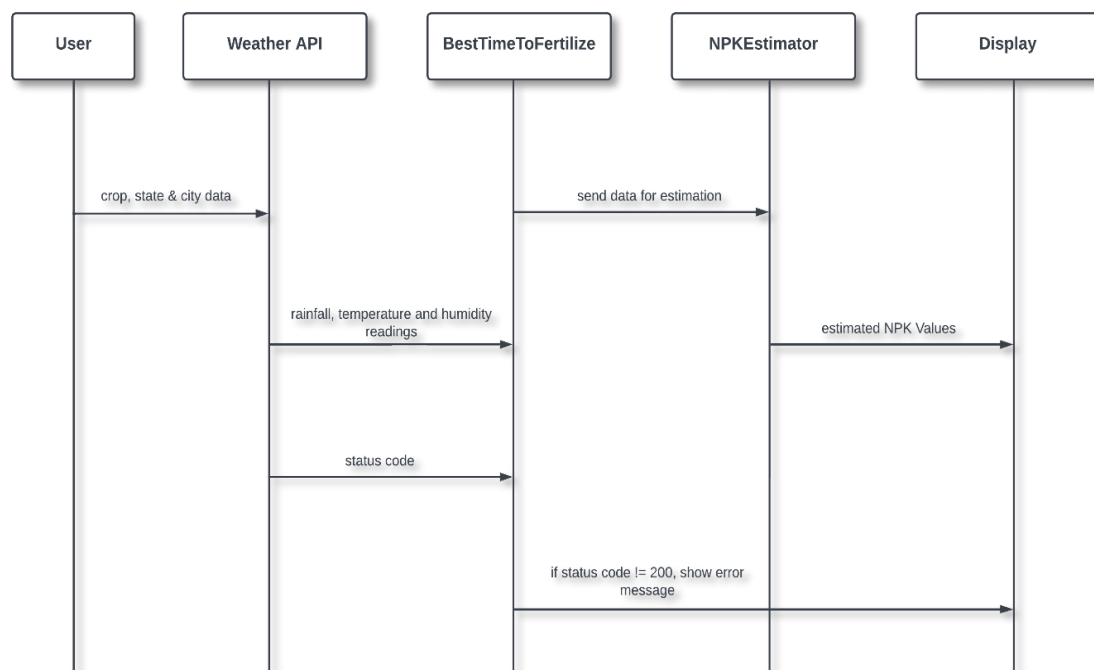


Fig 4.5 Sequence Diagram

CHAPTER V

IMPLEMENTATION

System implementation builds system pieces that adhere to user requirements of the system requirements founded in the early life cycle stages using the framework generated throughout architectural design and the outcomes of system analysis. These system components are then combined to create intermediate aggregates, which ultimately result in the entire system-of-interest (SoI). The system hierarchy's lowest-level system components are really produced by the implementation process (system breakdown structure). System components are created, purchased, or recycled. Production includes the forming, removing, connecting, and finishing processes used in hardware fabrication, the coding and testing processes used in software realization, or the processes used to build operating procedures for the duties of operators.

A design method known as "modular design," sometimes known as "modularity in design," separates a system into smaller components known as modules or skids that may be independently produced and then used in multiple systems. Functional division into distinct, scalable, reusable modules; strict usage of very well modular interfaces; and adherence to industry norms for interfaces are characteristics of a modular system.

5.1 Random Forest Regression

A group of several decision trees called a random forest (RF) are trained using different subsets of data and have changeable hyper-parameters. In our project, we are going to take crop and location as input, and based on it, we will predict the value of N, P, and K. First, we will divide our dataset into training and test datasets, where the training dataset is 80% of the original data and the rest 20% is test data. Then we will create three different random forests of size 50 (decision tree) for each N, P, and K and produces the average of the classes as the overall tree projection, shown in Table 5.1.

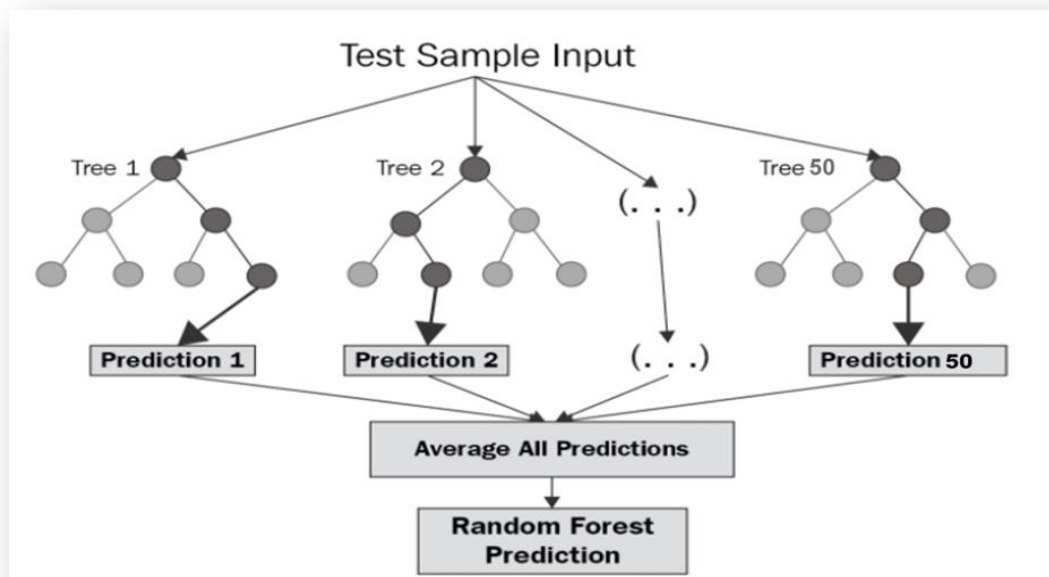


Fig.5.1.1 Random Forest Regression

BEGIN:
Step 1: The dataset of size $n = 2200$ is divided into training and test dataset (where the raining set is 80% and the test set is 20% that is training set=1,760 and the test set=240).
Step 2: Apply random forest regression to each N, P and K (Nitrogen, Phosphorus & Potassium) value with n estimators=50 (n estimators is the number of decision trees).
Step 3: Train the N Label, P Label and K Label with the training dataset and dependent variable (Where the dependent variable is N for N Label, P for P Label and K for K Label).
Step 4: Each N Label, P Label and K Label generates a 50-decision tree as an output based on the training dataset.
END

Table 5.1 Random Forest Regression Algorithm

Why we selected 50 decision trees (n_estimator = 50) for each label?

We have tested for different n_estimator values, but the upmost accuracy achieved for *N_Label* is 0.87 for two decimal digit precision. As shown in below figure Fig 5.1.1.

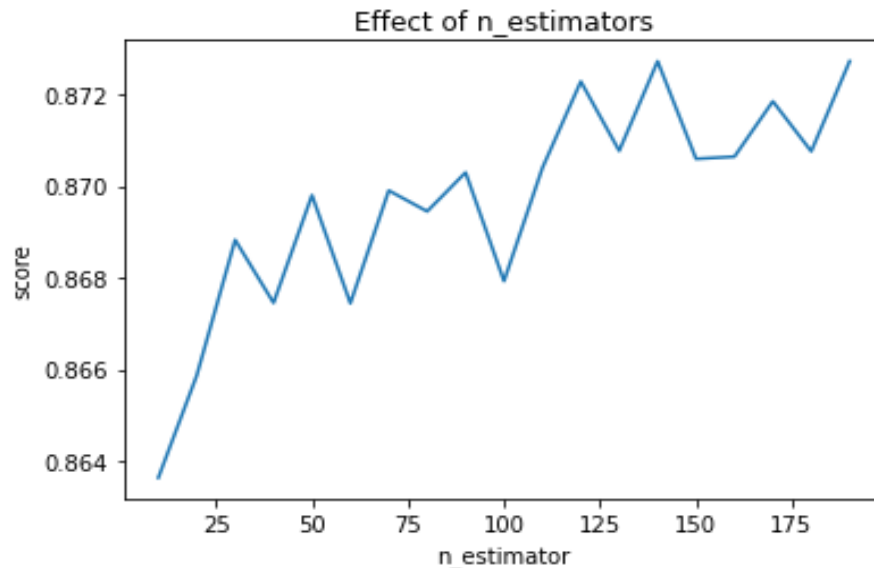


Fig 5.1.2 Effect of n_estimator

Why is a Random Forest chosen instead of a Decision Tree for this project?

Decision trees are trees that show all possible consequences of a selection using a branched technique, which is a key distinction between them and the random forest algorithm. In contrast, a set of decision trees that follow the output are produced by the random forest method.

In general, adding more trees will increase performance and predictability while decreasing calculation speed. The end solution for regression issues is the mean of all the trees. The samples in the tree target cell is the initial level of means in a random forest method regression model, followed by all trees. In contrast to linear regression, it estimates values beyond the observed range using prior observations.

The accuracy in the decision tree depends on the number of right predictions made divided by the total number of predictions, since it uses huge value attributes at each node, and it produces less accurate results when we apply an algorithm to handle the regression problem in a random forest. Decision trees are greedy and may be deterministic, meaning they produce different answers if we add or remove any additional rows. So, compared to

decision trees, random forest forecasts outcomes with higher accuracy.

The main problem with machine learning is overfitting. Overfitting may be viewed as a generic bottleneck in machine learning and occurs when we apply algorithms. When machine learning models are unable to perform well on unknown datasets, this is a sign of overfitting. This is especially true if the problem is detected mostly on testing or validation datasets and is significantly larger than the error on the training dataset. Overfitting occurs when models gain knowledge non - constant data in the training data, which has a negative effect on the performance on the new data model. Due to the employment of several decision trees in the random forest, the danger of overfitting is lower than that of the decision tree.

The accuracy increases when we employ a decision tree classifier on a data set since it contains more splits, which makes it easier to overfit the dataset and validate it.

So, that's why we decided to select a random forest as our machine learning model rather than decision tree to predict the required nutrients (Nitrogen, Phosphorus, Potassium) for the given crop. Random Forest performs well in terms of computation if we adjust the `n_estimator` value carefully. In our case we have used `n_estimator = 50` after taking readings of our model's accuracy with different `n_estimator` values, the same is shown in Fig 5.1.2.

5.2 Cross-Validation

In order to evaluate a machine learning algorithm on a small set of data, cross-validation is resampling technique. The algorithm's sole parameter, `k`, indicates how many groups should be formed from a given data sample. As a result, the technique is frequently referred to as `k`-fold cross-validation. When a precise value for `k` is given, it can be substituted for `k` in the model's regard, such as `k=4` for cross-validation which is performed four times.

In applied machine learning, cross-validation is mostly used to gauge how well a machine learning model performs on untrained data. That is, to use a small sample to assess how the model will generally perform when used to generate predictions on data that was not utilized during the model's training.

It is a well-liked technique since it is easy to comprehend and typically yields a less biased or overly optimistic assessment of the model ability than other techniques, including a straightforward train/test split.

Following is the general process:

1. Randomly shuffle the dataset.
2. Create k groups from the dataset.
3. For every distinct group:
 - The group should be used as a holdout or test data set.
 - As a training dataset contains, use the remaining groupings.
 - Adapt a model to the training set, then evaluate it against the test set.
 - Keep the evaluation result, but discard the model.
4. Using a sample of quality assessment ratings, summaries the model's skill.

It's significant that every observation in the sample data is given a unique group and remains there throughout the process. This indicates that each sample has the chance to be used k times to train the model and k times in the hold out set.

It is crucial that all data preparation done before fitting the model takes place on the loop's CV-assigned training dataset rather than the larger data set. This also holds true for any hyperparameter adjustment. Data leakage and an exaggerated assessment of the model's skill may occur from failing to carry out these procedures within the loop.

The mean of the model skill scores is frequently used to sum up the outcomes of a k -fold cross-validation run. A measure of the skill scores' volatility, like the standard error or standard deviation, should also be included.

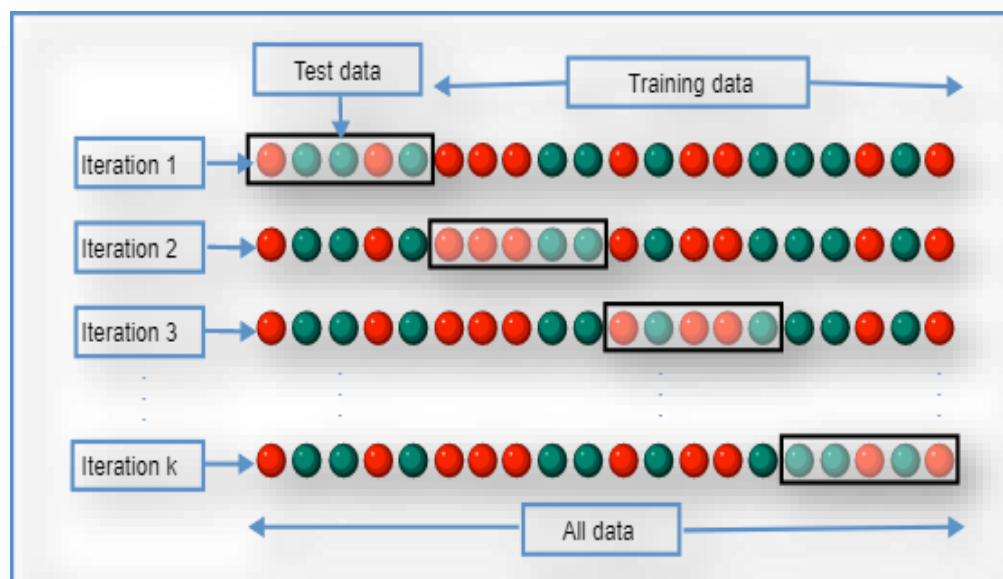


Fig.5.2.1: K-Cross Fold Validation

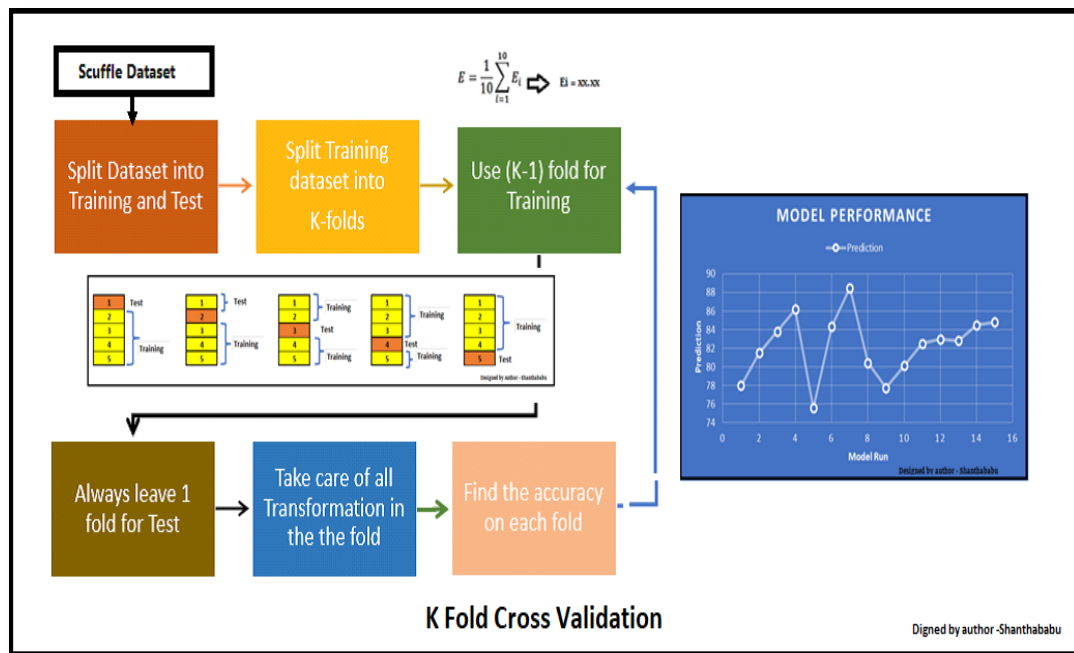


Fig 5.2.2: Cross-Validation procedure

5.3 Dataset

Dataset used in our proposed system:

	A	B	C	D	E	F	G	H
1	N	P	K	temperature	humidity	ph	rainfall	label
2	90	42	43	20.8797437	82.0027442	6.50298529	202.935536	rice
3	85	58	41	21.7704617	80.3196441	7.03809636	226.655537	rice
4	60	55	44	23.0044592	82.3207629	7.84020714	263.964248	rice
5	74	35	40	26.4910964	80.1583626	6.98040091	242.864034	rice
6	78	42	42	20.1301748	81.6048729	7.62847289	262.717341	rice
7	69	37	42	23.0580487	83.3701177	7.0734535	251.055	rice
8	69	55	38	22.708838	82.6394139	5.70080568	271.32486	rice
9	94	53	40	20.2777436	82.8940862	5.71862718	241.974195	rice
10	89	54	38	24.5158807	83.5352163	6.68534642	230.446236	rice
11	68	58	38	23.2239739	83.0332269	6.33625353	221.209196	rice
12	91	53	40	26.5272351	81.4175385	5.38616779	264.61487	rice
13	90	46	42	23.9789822	81.450616	7.50283396	250.083234	rice
14	78	58	44	26.800796	80.8868482	5.10868179	284.436457	rice
15	93	56	36	24.0149762	82.0568718	6.98435366	185.277339	rice
16	94	50	37	25.6658521	80.6638505	6.94801983	209.586971	rice
17	60	48	39	24.2820942	80.3002559	7.04229907	231.086335	rice
18	85	38	41	21.5871178	82.7883708	6.24905066	276.655246	rice
19	91	35	39	23.7939196	80.4181796	6.97085975	206.261186	rice
20	77	38	36	21.8652524	80.1923008	5.95393328	224.555017	rice
21	88	35	40	23.5794363	83.5876032	5.85393208	291.298662	rice
22	89	45	36	21.3250416	80.474764	6.44247538	185.497473	rice
23	76	40	43	25.1574553	83.1171348	5.07017567	231.384316	rice
24	67	59	41	21.9476674	80.973842	6.01263259	213.356092	rice
25	83	41	43	21.0525355	82.6783952	6.25402845	233.107582	rice
26	98	47	37	23.4838134	81.3326507	7.37548285	224.058116	rice
27	66	53	41	25.0756354	80.5238915	7.77891515	257.003887	rice
28	97	59	43	26.3592716	84.0440359	6.28650018	271.358614	rice
29	97	50	41	24.5292268	80.5449858	7.07096	260.263403	rice
30	60	49	44	20.7757615	84.497744	6.24484149	240.081065	rice

Fig.5.3 Actual Dataset

Crop Recommendation Dataset [20] last accessed on 31 Mar 2023

5.4 Data Preparation

Actual dataset contains eight features. All of the features are not useful for the proposed model. Therefore, a dimension reduction technique called feature selection is applied and seven features, then selected for evaluation.

	A	B	C	D	E	F	G
1	Crop	Temperature	Humidity	Rainfall	Label_N	Label_P	Label_K
2	rice	20.87974371	82.00274423	202.9355362	90	42	43
3	rice	21.77046169	80.31964408	226.6555374	85	58	41
4	rice	23.00445915	82.3207629	263.9642476	60	55	44
5	rice	26.49109635	80.15836264	242.8640342	74	35	40
6	rice	20.13017482	81.60487287	262.7173405	78	42	42
7	rice	23.05804872	83.37011772	251.0549998	69	37	42
8	rice	22.70883798	82.63941394	271.3248604	69	55	38
9	rice	20.27774362	82.89408619	241.9741949	94	53	40
10	rice	24.51588066	83.5352163	230.4462359	89	54	38
11	rice	23.22397386	83.03322691	221.2091958	68	58	38
12	rice	26.52723513	81.41753846	264.6148697	91	53	40
13	rice	23.97898217	81.45061596	250.0832336	90	46	42
14	rice	26.80079604	80.88684822	284.4364567	78	58	44
15	rice	24.01497622	82.05687182	185.2773389	93	56	36
16	rice	25.66585205	80.66385045	209.5869708	94	50	37
17	rice	24.28209415	80.30025587	231.0863347	60	48	39
18	rice	21.58711777	82.7883708	276.6552459	85	38	41
19	rice	23.79391957	80.41817957	206.2611855	91	35	39
20	rice	21.8652524	80.1923008	224.5550169	77	38	36
21	rice	23.57943626	83.58760316	291.2986618	88	35	40
22	rice	21.32504158	80.47476396	185.4974732	89	45	36
23	rice	25.15745531	83.11713476	231.3843163	76	40	43
24	rice	21.94766735	80.97384195	213.3560921	67	59	41
25	rice	21.0525355	82.67839517	233.1075816	83	41	43
26	rice	23.48381344	81.33265073	224.0581164	98	47	37
27	rice	25.0756354	80.52389148	257.0038865	66	53	41
28	rice	26.35927159	84.04403589	271.3586137	97	59	43
29	rice	24.52922681	80.54498576	260.2634026	97	50	41
30	rice	20.77576147	84.49774397	240.0810647	60	49	44

Fig 5.4 Customized Dataset

5.5 Input Features

Below are the input features of our system:

- **Crop:** rice, cotton, mango, orange, lentil, etc.
- **Temperature:** temperature measured in Celsius
- **Humidity:** measured relatively in percentages
- **Rainfall:** rainfall in mm

5.6 Output Features

Below are the output features of our system:

- **Label_N:** ratio of soil Nitrogen content
- **Label_P:** ratio of soil Phosphorus content
- **Label_K:** ratio of soil Potassium content

5.7 Python

Python is a high-level, interpreted, general-purpose programming language. It supports a range of programming paradigms, such as functional, object-oriented, and structured programming (particularly procedural). Python is a scripting language that may be used for web applications, including through Apache's mod wsgi. The Web Server Gateway Interface has become a standardised API for these applications. Web frameworks like Pylons, Django, web2py, Pyramid, Flask, Zope, Tornado, TurboGears, and Bottle enable programmers to create and maintain sophisticated applications. IronPython and Pyjs can be used to create the client-side of Ajax-based applications. One option for a relational database data mapper is SQLAlchemy. A framework for creating computer-to-computer communication is called Twisted.

Python is a fantastic choice for scientific computing thanks to tools like SciPy, Matplotlib, and Numpy as well as specialist libraries like Astropy and Biopython. SageMath is a notebook-based, Python-programmable computer algebra system that includes numerical mathematics, calculus, combinatorics, number theory, and algebra among many other mathematical specialties. Some of the libraries frequently used in

Python-based machine learning and AI applications are Pytorch, Scikit-learn, TensorFlow, and Keras.

Here is a collection of the main arguments in favor of Python:

- Additionally, the Python framework features packages and modules, which aid with code reuse.
- Open source software exists for Python. It is available for free download and use in applications. The source code can also be read and changed.
- Code does not need to be compiled, making the Edit-Test-Debug cycle quick.
- Supports handling of exceptions. Errors can occur in any program. Python creates exceptions which can be handled, preventing program crashes.
- Automatically managing memory. Python's memory management system uses a private heap, a data structure that acts as a queue and houses all of the language's data structures and objects.

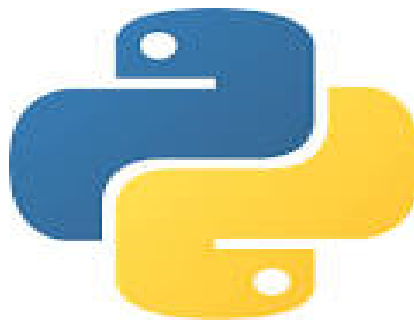


Fig 5.7: Python logo

5.8 Flask

A micro-web framework built on Python is called Flask. Because it doesn't need any particular tools or libraries, it is known as a microframework. It doesn't have a form validation layer, database abstraction layer, or any other elements that depend on already-built third-party libraries to carry out basic functions. But at the other hand, extensions can be utilized to increase application capabilities as if they had been created in Flask. The following standard framework-related tools all have extensions: upload handling, form

validation, object-relational mappers and numerous open authentication protocols.

Python is used to create the Flask web application framework. It is created by Armin Ronacher, the founder of Pocco, a global community of Python fans. The Jinja2 template engine, Werkzeug and WSGI toolkit serve as the foundation for Flask. They're both Pocco projects.

WSGI

Python web application development now adheres to the Web Server Gateway Interface (WSGI) standard. A uniform gateway between both the web applications and web server is described by the WSGI protocol.

Werkzeug

This is a WSGI toolkit that carries out utility operations like requests and response objects. As a result, a web framework may be built on top of it. Werkzeug serves as one of the foundations for the Flask framework.

Jinja2

A well-liked Python templating engine is Jinja2. To create dynamic web pages, a web templating method integrates a template with a specific data source. It's common to hear Flask referred to as a micro framework. It strives to keep an application's core straightforward but flexible. Both a form validation support and an established abstraction-layer for database processing are absent from Flask. Instead, Flask permits the use of extensions to give the application such capability.



Fig 5.8 Flask logo

5.9 HTML

- Hyper Text Markup Language is what HTML stands for.
- Website and applications are created using HTML.
- Hypertext: "Text within Text" is what hypertext is. A hypertext link can be found within a text. A link is considered a hypertext when it sends users to a new webpage when user click on it. Hypertext can be used to connect two or more web content.
- Markup language: A text document can be formatted and designed using this computer language. It can convert text into pictures, tables, and other formats.
- A web page is typically an HTML document that a web browser translates. Entering the URL will take you to the website. Both static and dynamic web pages are possible.
- Static Pages can be produced using HTML.
- HTML main tags:
 - `<!DOCTYPE>`: It describes the type of document or informs the browser of the HTML version.
 - `<html>`: This element informs the browser that the document is an HTML one.
 - `<head>`: The metadata should be contained in the first element of the html element. The body tag must open before it can be closed.
 - `<title>`: It is employed to include the HTML page's title, which is displayed just at top of the web browser.
 - `<body>`: The HTML's primary material is included in here.
 - `<h1>`: That the very first level heading of the page is described in the text inside the `<h1>` tag.
- HTML's key attributes include:
 - It is an extremely basic language. HTML is comprehensible and adaptable.
 - HTML makes it simple to create a presentation that is compelling.
 - It offers a versatile method for designing text-based web pages.
 - It gives programmers the ability to link to websites, which makes users more interested in exploring.
 - It may be seen on any platform, including Windows, Linux, and Macintosh, making it platform-independent.

- It makes it easier for programmers to enhance web sites with pictures, videos, and sound, making them more appealing and engaging.
- Both lowercase and uppercase tags are supported by HTML.



Fig 5.9 HTML logo

5.10 CSS

- Cascading style sheets, or CSS for short, is a language that governs how web pages are displayed, including their colors, fonts, and layout, making them more user-friendly.
- The primary use of CSS is to create style sheets for the internet. Even HTML is unrelated to it.
- Let's break down the CSS (acronym):
 - Falling into Style: Cascading
 - Imbuing our HTML tags: Style
 - Using style in a variety of publications' writing: Sheets
- Our HTML file can contain CSS in one of three ways:
 - Inline CSS
 - Internal CSS
 - External CSS
- Considering Priority:
 - Inline (then) Internal (then) External
- Inline CSS:
 - The only way to apply a style.
 - Independence
 - Adheres to each piece clearly.

- The concept of separating issues has been forgotten.
- Internal CSS:
 - We can use style tags to apply styles in HTML files.
 - Redundancy is eliminated.
 - But the concept of separating issues is still misunderstood.
 - Particularly used on a single document.
- External CSS:
 - We can apply styles with the aid of the link tag in the head tag.
 - References are included.
 - Documents end with a .css (extension).
 - Redundancy is eliminated.
 - The concept of distinct concerns is maintained.
 - Particularly used for each document.
- CSS characteristics:
 - A selector element and a declaration block element make up a style rule.
 - The HTML component that we want to have its style applied to is indicated by the selector.
 - One or more assertions are contained in the declaration block, separated by semicolons.
 - Every declaration that is entered has a semicolon, a value, and a name for the CSS attribute. For instance, if color is the property, then red is the value. The property is font size, and the value is 15 pixels.
 - These blocks are enclosed in curly braces, and the CSS declaration is completed with a semicolon.



Fig 5.10 CSS logo

5.11 JavaScript

- Website programming languages include JavaScript.
- Recognized for creating websites that are also often used outside of browsers.
- JavaScript is a declarative as well as an imperative language. There is a standard library of objects in JavaScript, including arrays, dates, math, and other types.
- Client-side: Provides objects to manipulate the browser's Document Object Model (DOM).
- For instance, client-side extensions let your program include HTML components.
- React to user events including form submission, mouse clicks, page navigation, etc. as well as forms and user events.
- Convenient client-side libraries include AngularJS, ReactJS, and VueJS.
- Server-side: Provides objects related to the execution of JavaScript on the server.
- Imperative Languages-In this type of language, we are primarily interested in how to do it.
- Controls only the calculation flow. The object-oriented approach, which is a procedural programming approach, is equivalent to asynchronous wait when considering what to do next after an asynchronous call.
- Declarative programming: I'm worried about how to do that in this kind of language. Basically, it requires logical operations. The main goal here is to explain the desired result without directly defining the acquisition method like the arrow function.
- There are two ways to include JavaScript in HTML:
 - Internal JS: As necessary, tags can be inserted inside of other tags..
 - External JS: The JavaScript code can be written in a separate file with a.js extension and linked within the tag of the HTML file we wish to add it to.

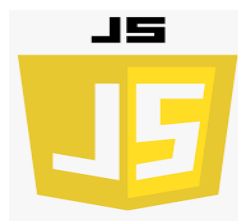


Fig 5.12 JavaScript logo

5.12 Backend Code

app.py

```
from flask import Flask, render_template, request, url_for
from BestTimeToFertilizeModule import BestTimeToFertilize
from NPKEstimatorModule import NPKEstimator

app = Flask(__name__)

@app.route('/processing/', methods=['GET', 'POST'])
def processing():
    # print('Processing.....')
    if request.method == "GET":
        print("The URL /processing is accessed directly.")
        return url_for('index.html')

    if request.method == "POST":
        form_data = request.form
        call_success = []
        npk_list_dict = []
        popup_data = []
        seven_days = []

        crop = form_data['crop']
        state = form_data['state']
        city = form_data['city']

        with open("InputData.csv", "w") as fh:
            input_data = "%s,%s,%s" % (crop.strip(), state.strip(), city.strip())
            fh.write(input_data)

        btff = BestTimeToFertilize(city_name = city, state_name = state)
        btff.api_caller()

        if btff.is_api_call_success():
```



```
category, heading, desc = btbf.best_time_fertilize()

call_success.append(1)
popup_data.append([category, heading, desc])
seven_days = btbf.weather_data[:]
# print(seven_days)

# today's weather data
di = btbf.weather_data[0]
temp = di['Temperature']
humidity = di['Relative Humidity']
rainfall = di['Rainfall']

est = NPKEstimator()
est.renameCol()

npk = {'Label_N':0, 'Label_P':0, 'Label_K':0}
for y_label in ['Label_N', 'Label_P', 'Label_K']:
    npk[y_label] = est.estimate(crop, temp, humidity, rainfall, y_label)
# print(npk)

npk_list_dict.append(npk)

output_data = category + "\n" + heading + "\n" + desc + "\n" + str(npk['Label_N']) + "\n" +
str(npk['Label_P']) + "\n" + str(npk['Label_K'])
with open("output.txt", "w") as fh:
    fh.write(output_data)
else:
    print("Error Occured")
#print(call_success, npk_list_dict, form_data, popup_data)
return render_template('update.html', CALL_SUCCESS = call_success, NPK = npk_list_dict, FORM_DATA
= form_data, POPUP_DATA = popup_data, SEVEN_DAYS = seven_days)

@app.route('/', methods=['POST', 'GET'])
def index():
```

```
return render_template('index.html')

if __name__ == "__main__":
    app.run(debug=True)
```

BestTimeToFertilizeModule.py

```
import requests as rq
import json as js
from time import sleep

class BestTimeToFertilize:
    __BASE_URL = "https://api.weatherbit.io/v2.0/forecast/daily?"
    __API_KEY = "480589e42e7c4352abe4fe25bd398ab0"

    def __init__(self, city_name = 'Bangalore', state_name = 'Karnataka', days = 7):
        self.city_name = '+'.join(city_name.lower().strip().split())
        self.state_name = '+'.join(state_name.lower().strip().split())
        self.country_name = 'IN'
        self.days = days
        self.response = None
        self.response_code = None
        self.weather_data = list()

    def api_caller(self):
        try:
            complete_url = "{0}city={1}&state={2}&country={3}&key={4}&days={5}".format(self.__BASE_URL,
self.city_name, self.state_name, self.country_name, self.__API_KEY, self.days)
            # print(complete_url)
            # while self.response == None:
            self.response = rq.get(complete_url)
            sleep(5)
            self.response_code = self.response.status_code
            return self.response_code
        except Exception as msg:
            print("api_caller():", msg)
            return -1
```

```
def is_api_call_success(self):
    if self.response_code == 200:
        return True
    elif self.response_code == 204:
        print('Content Not available, error code: 204')
    return False

def json_file_builder(self):
    try:
        json_obj = self.response.json()
        with open('weather_data.json', 'w') as file:
            js.dump(json_obj, file, indent = 1, sort_keys = True)
        print("weather_data.json file build successfully")
    except Exception as msg:
        print("json_builder():", msg)

def best_time_fertilize(self):
    json_obj = self.response.json()

    # print("City:", json_obj['city_name'], "\n")

    prolonged_precip = 0
    prolonged_prob = 0

    heavy_rain_2d = False
    heavy_rain_chance_2d = 0
    precip_2d = 0
    precip_chance_2d = 0

    for i in range(self.days):
        date = json_obj['data'][i]['datetime']
        temp = json_obj['data'][i]['temp']
        rh = json_obj['data'][i]['rh']
        precip = json_obj['data'][i]['precip']
```

```

prob = json_obj['data'][i]['pop']
w_code = json_obj['data'][i]['weather']['code']
w_desc = json_obj['data'][i]['weather']['description']
i_code = json_obj['data'][i]['weather']['icon']
prolonged_precip += precip
prolonged_prob += prob

count_2d = 0
if i < 2:
    precip_2d += precip
    precip_chance_2d += prob
    if w_code in [202, 233, 502, 521, 522]:
        heavy_rain_2d = True
        heavy_rain_chance_2d += prob
        count_2d += 1
        heavy_rain_chance_2d //= count_2d

di = {
    "Date":date,
    "Temperature":temp,
    "Relative Humidity":rh,
    "Rainfall":precip,
    "Probability of Precipitation":prob,
    "Weather Description": w_desc
}

self.weather_data.append(di)

prolonged_prob //= self.days
precip_chance_2d //= 2

if heavy_rain_2d:
    print("***21, "Warning !!!", "***21)
    print("Heavy Rain Chances within 2 days:", heavy_rain_chance_2d)
    print("Heavy Rainfall puts your fertilizer at risk.")
    print("***21, "Warning !!!", "***21)

```

```
        return ('Warning', 'Heavy Rain Alert', 'Heavy Rain Chances within two days from now is %d%%' %
                (heavy_rain_chance_2d))

    elif prolonged_precip > 12.7 and prolonged_prob >= 50:
        print("***21, \"Warning !!!\", \"***21)
        print(\"Prolonged Rainfall of greater than 12.7 mm puts your fertilizer at risk.\")
        print("***21, \"Warning !!!\", \"***21)
        return ('Warning', 'Prolonged Rainfall Alert', 'Prolonged Rainfall of greater than 12.7 mm puts your
fertilizer at risk. From now %.2f mm rainfall will receive for upcoming seven days, chances %d%%' %
                (prolonged_precip, prolonged_prob))
    else:
        print("-"*80)
        print("The amount of rain for 2 days, counting today:", precip_2d)
        print("Chances of rain for 2 days, counting today:", precip_chance_2d)
        print()
        return ('Message', 'Precipitation Amount', 'The amount of rain for 2 days, counting today is %.2f mm
and chances is %d%%' % (precip_2d, precip_chance_2d))
```

NPKEstimatorModule.py

```
import warnings
import numpy as np
import pandas as pd
from sklearn import metrics
import category_encoders as ce
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor

warnings.filterwarnings('ignore')

class NPKEstimator:
    def __init__(self, data = 'Nutrient_recommendation.csv', ):
        self.df = pd.read_csv(data, header=None)
        self.X_train = None
        self.X_test = None
        self.y_train = None
        self.y_test = None
```

```
def renameCol(self):
    self.df.columns = ['Crop', 'Temperature', 'Humidity', 'Rainfall', 'Label_N', 'Label_P', 'Label_K']
    self.df.drop(self.df.index[:1], inplace=True)

def cropMapper(self):
    # create mapping of crop(string) to int type

    mapping = dict()

    with open("mapped_crops.csv", "w") as fh:
        fh.write("Crops,Key\n")
        for i, crop in enumerate(np.unique(self.df[['Crop']]), 1):
            mapping[crop] = i
            fh.write("%s,%d\n" % (crop, i))
        mapping['NA'] = np.nan
        fh.write("NA,nan")
    # print(mapping)

    ordinal_cols_mapping = [{"col": "Crop", "mapping": mapping}, ]
    encoder = ce.OrdinalEncoder(cols = 'Crop', mapping = ordinal_cols_mapping, return_df = True)
    return mapping, encoder

def estimator(self, crop, temp, humidity, rainfall, y_label):
    X = self.df.drop(['Label_N', 'Label_P', 'Label_K'], axis=1)
    y = self.df[y_label]

    self.X_train, self.X_test, self.y_train, self.y_test = train_test_split(X, y, test_size = 0.20, random_state =
42)

    mapping, encoder = self.cropMapper()
    self.X_train = encoder.fit_transform(self.X_train)
    self.X_test = encoder.transform(self.X_test)
```

```
regressor = RandomForestRegressor(n_estimators = 50, random_state = 0)
regressor.fit(self.X_train, self.y_train)

# y_pred = regressor.predict(self.X_test)
query = [mapping[crop.strip().lower()], temp, humidity, rainfall]
y_pred = regressor.predict([query])
return y_pred[0]
```

```
def accuracyCalculator(self):
    model = RandomForestRegressor(n_jobs=-1)
    estimators = np.arange(10, 200, 10)
    scores = []
    for n in estimators:
        model.set_params(n_estimators=n)
        model.fit(self.X_train, self.y_train)
        scores.append(model.score(self.X_test, self.y_test))

    scores_arr = [round(sc, 3) for sc in scores]
    unique, counts = np.unique(scores_arr, return_counts = True)

    max_count = max(counts)

    accuracy = -1
    for uni, count in zip(unique, counts):
        # print(uni, count)
        if count == max_count:
            accuracy = uni

    # print("Model accuracy: %.2f" % (accuracy))
    return accuracy
```

CHAPTER VI

SOFTWARE DESCRIPTION

The software utilized in this development to satisfy our demands is described in detail in this chapter. The following is a description of the software:

6.1 VS Code

The Code - OSS repository is made available under a standard Microsoft product license, but with the Microsoft-specific modification of Visual Studio Code. The functionality that developers need to finish their edit-build-debug cycle are combined with a code editor's simplicity of use in Visual Studio Code. It offers comprehensive code editing, navigation, and comprehension support along with a deep extension architecture, lightweight debugging, and a lightweight interface with existing tools. Every month, Visual Studio Code receives updates with bug fixes and new features. On the Visual Studio Code website, we may download Visual Studio Code for Linux, Windows and macOS.



Fig.6.1 VS Code logo

6.2 Git

Git is an open-source, free version control system that can manage tasks of any scale, from modest to massive. Git is easy to use, leaves a small digital footprint, and works effectively.

It beats SCM solutions like CVS, Subversion, ClearCase and Perforce thanks to features like affordable local branching, practical staging zones, and many processes. Git encourages and supports the development of numerous independent local branches. Sophisticated production lines can be created, merged, and removed in a matter of a few seconds.



Fig.6.2 Git logo

6.3 Jupyter Notebook

A free online application called Jupyter Notebook allows users to create and share documents that include text, live code, equations, and visualizations. Jupyter Notebook is maintained by Project Jupyter. The IPython project, which formerly had its own IPython Notebook project, gave birth to Jupyter Notebooks. Julia, Python, and R are the three primary programming languages that Jupyter supports. Although Jupyter includes the IPython kernel, which enables you to create Python program, there are already over 100 other kernels available.



Fig.6.3 Jupyter Notebook logo

CHAPTER VII

RESULTS

AgriTech, a user-friendly system, has been implemented in the form of a website to provide cross-platform functionality and suggest appropriate timings and amount of nutrients required for an inputted crop with alert system for heavy rainfall (as shown in Fig 7.1-7.6).

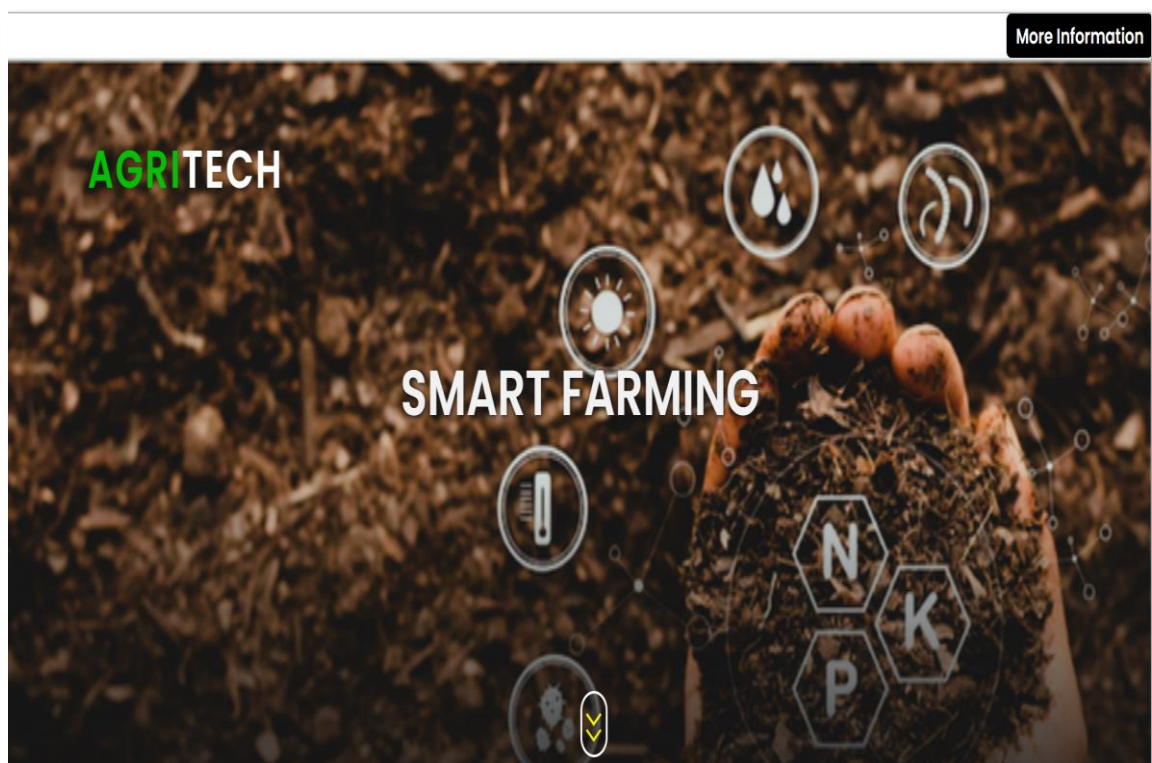


Fig 7.1 Homepage

What are NPK Values?

What do the numbers on fertilizer mean?

N	P	K
NITROGEN	PHOSPHORUS	POTASSIUM
greens up plants	reaches down to the roots and helps produce blooms	promotes all around wellbeing
JUST THINK: ↑ UP ↑	↓ DOWN ↓	↻ ALL AROUND ↻
NITROGEN	PHOSPHORUS	POTASSIUM

NPK values are a way of describing the nutrient content of fertilizers that are used to help plants grow. The three letters N, P, and K stand for nitrogen, phosphorus, and potassium, which are the three main nutrients that plants need to thrive.

The NPK values on fertilizer bags are typically shown as a set of three numbers separated by hyphens. For example, a fertilizer with an NPK value of 10-5-5 means that it contains 10% nitrogen, 5% phosphorus, and 5% potassium.

Fig 7.2 More information page

Fill out the Details

Select Crop

Select State

SUBMIT



Fig 7.3 Input Form

Fill out the Details

rice

Karnataka

Bidar

SUBMIT



Fig 7.4 Details filled using the drop-down menu

Fill out the Details

rice

Karnataka

Bidar

LOADING..



Fig 7.5 Applying Algorithm to inputted details

Entered Details

rice

Karnataka

Bidar

Required Nutrient Ratio

N : 103.76

P : 25.26

K : 19.74

Message

Precipitation Amount

The amount of rain for 2 days, counting today is 0.00 mm and chances is 0%

< BACK

7 DAYS REPORT

Date : 2023-02-09

Temperature : 25.2

Relative Humidity : 25

Rainfall : 0

Probability of Precipitation : 0

Weather Description : Few clouds

Date : 2023-02-10

Temperature : 26.1

Relative Humidity : 16

Rainfall : 0

Probability of Precipitation : 0

Weather Description : Few clouds

Date : 2023-02-11

Temperature : 27.1

Relative Humidity : 15

Rainfall : 0

Probability of Precipitation : 0

Weather Description : Few clouds

Date : 2023-02-12

Temperature : 27

Relative Humidity : 16

Rainfall : 0

Probability of Precipitation : 0

Weather Description : Few clouds

Date : 2023-02-13

Temperature : 25.9

Relative Humidity : 19

Rainfall : 0

Probability of Precipitation : 0

Weather Description : Few clouds

Date : 2023-02-14

Temperature : 25.8

Relative Humidity : 13

Rainfall : 0

Probability of Precipitation : 0

Weather Description : Clear Sky

Date : 2023-02-15

Temperature : 26.8

Relative Humidity : 12

Rainfall : 0

Probability of Precipitation : 0

Weather Description : Clear Sky

Fig 7.6 Output with seven days of weather forecasts & alerts/messages

CHAPTER VIII

CONCLUSION AND FUTURE SCOPE

8.1 Conclusion

The proposed system is able to provide information about the amount of nutrients required by the crops for satisfactory crop growth and production with respect to weather conditions. It provides weather alerts and messages. Alerts are displayed in the output of this application in case of bad weather conditions. The accuracy can be improved further with the development of technologies.

8.2 Future Scope

The proposed system provides a helping hand to our farmers. It gives information about the use and quantity of nutrients required by the crops. There is scope for improvement in the system by providing user interface in the native language, so that the user can operate the system easily if he or she is unfamiliar with the English language. In addition, speech recognition systems can be added to handle illiterate users.

REFERENCES

- [1] Krutika Hampannavar, Vijay Bhajantri, Shashikumar G. Totad “Prediction of Crop Fertilizer Consumption,” Fourth International Conference on Computing Communication Control and Automation (ICCUBE),2018, PP.1-5
- [2] G. Prabakaran, D. Vaithiyanathan, Madhavi Ganesa, “Fuzzy decision support system for improving the crop productivity and efficient use of fertilizers,” Computers and Electronics in Agriculture, vol-150, 2018, PP. 88-97
- [3] Shital Bhojani, Nirav Bhatt, “Data Mining Techniques for Crop Yield Prediction,” Computers and Electronics in Agriculture, vol-6, 2018, PP. 357-358
- [4] Yulong Yin, Hao Ying, Huifang Zhen, Qingsong Zhang, Yanfang Xue, Zhenling I, “Estimation of NPK requirements for rice production in diverse Chinese environments under optimal fertilization rate,” Agricultural and Forest Meteorology, vol-279, 2019, PP. 1-6
- [5] Laura J.T. Hess, Eve-Lyn S. Hinckley, G. Philip Robertson, Pamela A. Matson, “Rainfall intensification increases nitrate leaching from tilled but not no-till cropping systems in the U.S. Midwest,” Agriculture, Ecosystems & Environment, vol-290, 2020, PP. 1-10
- [6] Potnuru Sai Nishant, Pinapa Sai Venkat, Bollu Lakshmi Avinash, B. Jabber, “Crop Yield Prediction Based on Indian Agriculture using Machine Learning,” 2020 International Conference for Emerging Technology (INCET), 2020, PP. 1-4
- [7] Tony Yang, Kadambot H.M., Siddique, Kui Liu, “Cropping systems in agriculture and their impact on soil health,” Global Ecology and Conservation, vol-23, year, PP. 1-13
- [8] János Káta, Ágnes Oláh Zsuposné, Magdolna Tállai, Tarek Alshaal, “Would

fertilization history render the soil microbial communities and their activities more resistant to rainfall fluctuations? ,” *Ecotoxicology and Environmental Safety*, vol-201, 2020, PP. 1-11

[9] Usman Ahmed, Jerry Chun-Wei Lin, Gautam Srivastava, Youcef Djenouri, “A nutrient recommendation system for soil fertilization based on Evolutionary Computation,” *Computers and Electronics in Agriculture*, vol-189, 2021, PP. 1-7

[10] A.Hussein, Diogenes L. Antille, Shreevatsa Kodur, GuangnanChen, Jeff N.Tullberg, “Controlled traffic farming effects on productivity of grain sorghum, rainfall and fertilizer nitrogen use efficiency,” *Journal of Agriculture and Food Research*, vol-3, 2021, PP. 1-17

[11] Zujiao Shi, Donghua Liu, Miao Liu, Muhammad Bilal Hafeez, Pengfei Wen, Xiaoli Wang, Rui Wang, Xudong Zhang, Jun Li, “Optimized fertilizer recommendation method for nitrate residue control in a wheat–maize double cropping system in dryland farming,” *Field Crops Research*, vol-271, 2021, PP. 1-10

[12] Janmejy Pant, R.P. Pant, Manoj Kumar Singh, Devesh Pratap Singh, Himanshu Pant, “Analysis of agricultural crop yield prediction using statistical techniques of machine learning,” *Materials Today: Proceedings*, vol-46, 2021, PP. 1-10

[13] Benny Antony, “Prediction of the production of crops with respect to rainfall,” *Environmental Research*, vol-202, 2021, PP. 1-5

[14] Akash Manish Lad, K. Mani Bharathi, B. Akash Saravanan, R. Karthik, “Factors affecting agriculture and estimation of crop yield using supervised learning algorithms,” *Materials Today: Proceedings*, 2022, PP. 1-10

[15] Raves Akhtar, Shabbir Ahmad Sofi, “Precision agriculture using IoT data analytics and machine learning,” *Journal of King Saud University - Computer and Information Sciences*, 2021, PP. 1-17

[16] Saheed Garnaik, Prasanna Kumar Samant, Mitali Mandal, Tushar Ranjan Mohanty,

Sanat Kumar Dwibedi, Ranjan Kumar Patra, Kiran Kumar Mohapatra, R.H. Wanjari, Debadatta Sethi, Dipaka Ranjan Sena, Tek Bahadur Sapkota, Jagmohan Nayak, Sridhar Patra, Chiter Mal Parihar, Hari Sankar Nayak, “Untangling the effect of soil quality on rice productivity under a 16-years long-term fertilizer experiment using conditional random forest,” *Computers and Electronics in Agriculture*, vol-197,2022, PP. 1-10

[17] Rubby Aworka, Lontsi Saadio Cedric, Wilfried Yves Hamilton Adoni, Jérémie Thouakessseh Zoueu, Franck Kalala Mutombo, Charles Lebon Mberi Kimpolo, Tarik Nahhal, Moez Krichen, “Agricultural decision system based on advanced machine learning models for yield prediction: Case of East African countries,” *Smart Agricultural Technology*, vol-3, 2022, PP. 1-9

[18] Senthil Kumar Swami Durai, Mary Divya Shamili, “Smart farming using Machine Learning and Deep Learning techniques,” *Decision Analytics Journal*, vol-2, 2022, PP. 1-30

[19] M.S. Suchithra, Maya L. Pai, “Improving the prediction accuracy of soil nutrient classification by optimizing extreme learning machine parameters,” *Information Processing in Agriculture*, vol-7, 2022, PP. 1-11

[20] Kaggle, <https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset> last accessed on 31 Mar. 2023.

PAPER PUBLICATION DETAILS

Paper published in IJRMETS (International Journal) :



e-ISSN: 2582-5208

International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)

Volume:05/Issue:04/April-2023

Impact Factor- 7.868

www.irjmets.com

SMART USE OF FERTILIZER IN AGRICULTURE USING MACHINE LEARNING

**Suresh Chimkode^{*1}, Ankush Kanjekar^{*2}, Arun Patil^{*3}, Kalyan Kumar^{*4},
Amish Bemelkedar^{*5}**

^{*1}Professor, Department of Computer Science and Engineering, Guru Nanak Dev Engineering College, Bidar,
Karnataka, India

^{*2,3,4,5}Student, Department of Computer Science and Engineering, Guru Nanak Dev Engineering College, Bidar,
Karnataka, India

DOI : <https://www.doi.org/10.56726/IJRMETS37071>

ABSTRACT

Farmers typically have limited control over fertilizer use, and competent guidance is needed for them to achieve higher yields while reducing fertilizer loss. The amount of rainfall after each fertilizer application is linked to nutrient loss, with moderate rainfall at the right time being beneficial for nutrient penetration into the soil's rooting zone and dissolution of dry fertilizer. However, excessive rainfall can increase the risk of nutrient runoff and leaching, affecting essential nutrients such as nitrogen (N), phosphorus (P), potassium (K), manganese (Mn), and boron (B) present in the soil. To address this, this study proposes an updated version of the random forest algorithm based on time-series data, to predict the required amount of nutrients for various crops by analyzing rainfall patterns and crop fertility. The method suggested in this study is useful for improving soil fertility by providing recommendations for optimal nutrient conditions for crop growth and reducing the potential for leaching and runoff.

Keywords: Fertilizer use, rainfall patterns, nutrient loss, random forest algorithm, time-series data, soil fertility.
