A PROJECT REPORT

on

CROP AND FERTILIZER PREDICTION

Submitted to

KIIT Deemed to be University

In Partial Fulfilment of the Requirement for the Award of

BACHELOR'S DEGREE IN
Computer Science and Engineering

BY

SAKSHAM RAJ	2005473
NIVEDITA SUTRADHAR	20051550
SUMITA CHAKRAVARTY	2005621
KV PRASHANT	2005102

UNDER THE GUIDANCE OF
HRUDAYA KUMAR TRIPATHY

SCHOOL OF COMPUTER ENGINEERING
KALINGA INSTITUTE OF INDUSTRIAL TECHNOLOGY
BHUBANESWAR, ODISHA - 751024

May 2023

KIIT Deemed to be University

School of Computer Engineering
Bhubaneswar, ODISHA 751024

CERTIFICATE

This is certify that the project entitled

"CROP AND FERTILIZER PREDICTION"

submitted by

SAKSHAM RAJ	2005473
NIVEDITA SUTRADHAR	20051550
SUMITA CHAKRAVARTY	2005621
KV PRASHANT	2005102

is a record of bonafide work carried out by them, in the partial fulfilment of the requirement for the award of Degree of Bachelor of Engineering (Computer Science and Engineering) at KIIT Deemed to be university, Bhubaneswar. This work is done during year 2022-2023, under our guidance.

Date: 03/ 05/ 2023

Hrudaya Kumar Tripathy
Project Guide

Acknowledgement

We would like to express our sincere gratitude to School of Computer Science and Engineering KIIT University for giving us a chance and all necessary support to do this project.

First and foremost, We are deeply grateful to our project guide Dr. Hrudaya kumar Tripathy for providing us with the necessary guidance and support throughout the project. Their expert knowledge, valuable feedback, and suggestions helped me to complete this project successfully.

We would also like to thank the faculty and staff of the Department of Computer Science for their constant support and encouragement towards our project.

We extend our heartfelt appreciation to our family and friends for their unwavering support, motivation, and understanding throughout the project.

We would also like to thank our predecessors who have done such thorough research in predicting the attrition of crop prediction and fertilizer prediction using machine learning which helped us to formulate a path on how we were supposed to pursue this project.

Lastly, we would like to acknowledge the vital contributions made by the various online resources, research papers, and literature that played an essential role in developing and enhancing the project.

Thank you all for making this project a success!

SAKSHAM RAJ

NIVEDITA SUTRADHAR

KV PRASHANT

SUMITA CHAKRAVARTY

ABSTRACT

Since ancient times, agriculture has been India's primary industry. India has a long history of being an agricultural country. In order to secure the food security of all Indians as well as their ability to work and support themselves, agriculture contributes a significant amount of the domestic product to the Indian economy. Since the turn of the century, farmers, particularly Indian farmers, have been making a good living. But because of the debt and loss they are experiencing, many farmers are now taking their own lives. Now because the weather and soil condition is always changing, food production and forecasting made by farmers in the past based on experience are becoming less accurate. The results of this study will aid beginning farmers by providing them with instructions on how to plant appropriate crops and how much fertilizer to use by applying.

Crop Prediction systems have gained significant attention in recent years, with the increasing need to optimize agricultural production while reducing costs and maximizing yield. These systems use different machine learning models to predict the best crop for a particular location based on soil fertility, location, cost of yield, and other factors.

The primary goal of this project is to help farmers select the best crop that they can grow in their specific location and soil conditions, leading to a significant improvement in the quality and quantity of their harvest. The machine learning models are trained on various data sets, including soil characteristics, locations and other factors affecting crop growth.

The system generates crop Predictions based on the input provided by farmers, including location, soil type, and other variables. The models analyze the data and provide a list of crops suitable for the specific location, taking into account the various factors that affect crop growth. This helps farmers make informed decisions on what crops to plant, reducing the risk of crop failure and increasing their overall yield.

Overall, Crop Prediction systems have the potential to revolutionize the agricultural sector by providing farmers with the necessary information to optimize their crop production and increase their income. By leveraging the power of machine learning, these systems can help farmers select the best crop for their land, improving their livelihoods while contributing to sustainable agriculture.

Keywords : Machine Learning, Crop Prediction, Fertilizer Prediction, Data analysis, Data Preprocessing, Supervised Learning

Contents


Fig.1 - Agronomist discussing with farmers in the outdoor field regarding farming

Fig. 2- Implementation of Artificial Intelligence in Farming in coming years

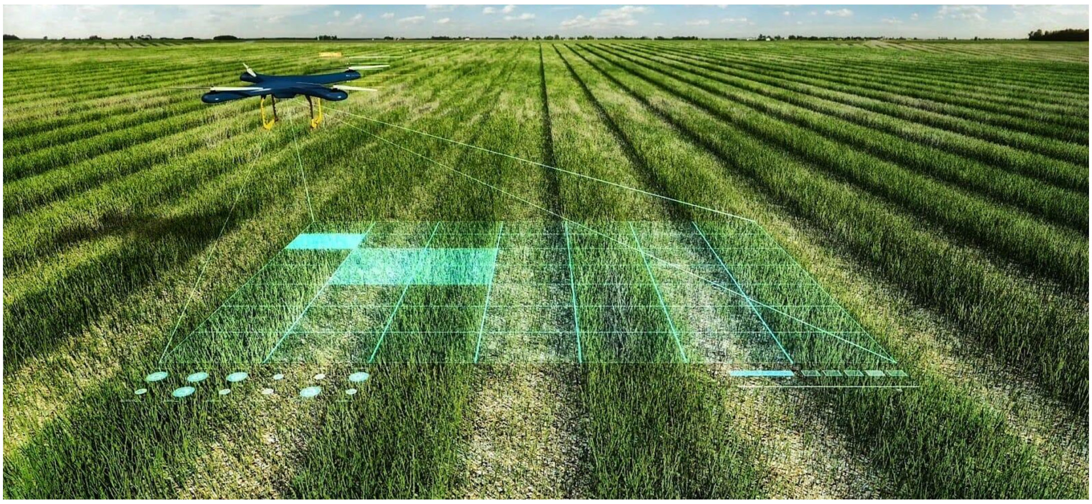


Fig. 3-Drones on open plains scanning the area for the development of cultivation using Image Annotation in ML

Chapter 1

Introduction

Agriculture has long been regarded as the primary source of supplies for meeting people's basic requirements. It is regarded as a basic employment as well as one of India's important industrial sectors. In order to maintain a healthy diversity, farmers should practise traditional naked eye observation and produce healthy crops without using pesticides on their cultivation field or on the animals that eat those products. But in today's world, weather patterns are shifting quickly in opposition to the availability of natural resources, which decreases food availability and increases security. The agricultural sector's GDP is also steadily falling. About 17.2% of the population experienced it in 2005, 11.1% in 2012, 5% in 2018, and only 2% in the first quarter of 2020. Approximately 80% of farmers come from rural areas, and farms as an industry will have an impact on their way of life if crop production revenues decline.

It makes sense for Indian farmers to exhibit a specific interest in precise and effective farming. There are several ways to boost crop output and profit while simultaneously boosting crop quality in order to preserve agricultural economic growth in India. Therefore, one of the solutions for crop prediction in relation to atmospheric & soil parameters of agricultural land is to implement a recent technological advance like machine learning. because climate variability has increased during the past few decades. It is changing as a result of globalisation every day.

The proposed approach explores the use of supervised machine learning and takes into account the most likely class as a prospective class. In this instance, the category consists just of the crop that was forecast using the input data. Farmers will find it simpler to choose an economical crop for each parcel of land they own if the crop is predicted. Then, through a mobile application, the farmers are given instructions to assist them in understanding what kinds of seeds should be planted in order to increase yields. In order to forecast crops using previous data, it was necessary to look at farmers' prior knowledge of the weather. As a result, choosing the right crop requires comprehensive knowledge of historical climatic circumstances. This study recommends a technique for using machine learning to forecast a crop that is cheap for the farmers that are in need of assistance for the given input parameter. Consequently, the proposed work will provide farmers with useful Predictions for more lucrative farming.

Chapter 2

Literature Review

Expert agronomists have traditionally predicted crops based on their expertise and understanding of a variety of variables, including weather, soil, and crop characteristics. However, this approach required a lot of time, money, and judgement. Due to its capacity to handle vast volumes of data and spot patterns that may not be obvious to human specialists, machine learning techniques have recently attracted a lot of attention for crop prediction.

K-Nearest Neighbors (KNN), a machine learning technique that uses distance-based learning, is utilized in crop Prediction systems. The foundation of KNN is the idea that predictions should be based mostly on the labels of the K data points that are closest to the input data point.

Another method utilized in crop prediction systems is Support Vector Machine (SVM). With the objective of maximizing the margin between the hyperplane and the data points, SVM is a probabilistic-based learning algorithm that seeks to identify a hyperplane that divides the data into several classes.

Making judgments based on a set of rules is done using a tree-based model called a decision tree. Decision trees are employed in crop Prediction systems to pinpoint the critical elements that affect crop yield and to suggest the best crop for a given area.

A bagging-based ensemble technique called Random Forest makes predictions by using several different decision trees. To get beyond the difficulties of analyzing big datasets, crop Prediction systems employ it because it can handle complex datasets.To increase accuracy and manage big datasets, crop Prediction systems use the ensemble algorithm XgBoost. In XG Boost, many decision trees are combined to generate forecasts, with each tree learning from the mistakes of the one before it.

Crop Prediction systems use the probabilistic method Naive Bayes to produce predictions about the likelihood that an event will occur. The likelihood of each feature happening independently is calculated by Naive Bayes under the assumption that the features are independent.

In conclusion, crop Prediction systems are a promising technology that aids farmers in maximising crop yield by using a variety of machine learning algorithms. Some of the most popular algorithms in crop Prediction systems include KNN, SVM, decision trees, random forests, XGBoost, and naive Bayes. Each approach has advantages and disadvantages.

CHAPTER 3

Problem Statement

This project's primary goal is to predict crops, which can be very helpful to farmers as they make plans for grain harvest and sale. Implement a machine learning algorithm that provides a more accurate forecast of the best crop for the appropriate region and growing season in our nation. Based on location, this study seeks to forecast crop. This study's objective is to examine the prediction of crops that will provide high yield in the specified site taking into account the location and soil attributes. It is a vital duty for guaranteeing sustainable agriculture and food security. The time-consuming, expensive, and subjective traditional techniques of crop and fertilizer forecast rely on specialist agronomists. In order to forecast crop production and fertilizer needs, accurate and effective machine learning models are required.

3.1 Project Planning

When preparing to implement the project development of a machine learning-based crop prediction system, there are a number of steps that need to be taken. The following are these steps:

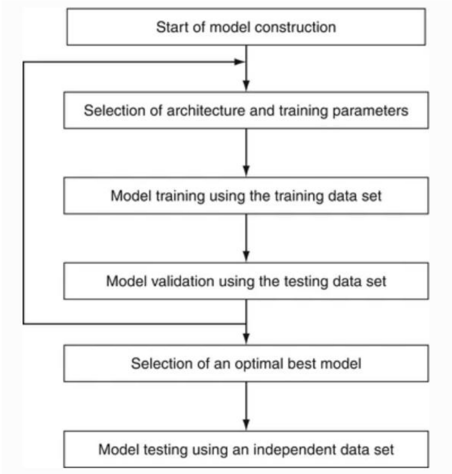
Data gathering and preprocessing: The fundamental phase in the machine learning pipeline is data gathering for training the ML model. The accuracy of the predictions provided by ML systems is only as good as the training data. The following are a few issues that can occur during data collection: This entails preparing the data for machine learning by cleaning, processing, and structuring it.Any machine learning algorithm's accuracy is influenced by the training dataset's correctness and parameter count. The datasets were chosen based on their soil characteristics, state-level production with cost of cultivation and cost of production/yield of crops in under five years or less, and data of fertilizers by using data of temperature, humidity, moisture, soil type, crop type, nitrogen, phosphorus, and potassium levels in soil and recommending fertilizer based on the values.
Data per-processing changes the data into a format that can be processed in data mining, machine learning, and other data science tasks more quickly and efficiently. The methods are typically applied as soon as possible.

The information is transformed into a distribution with a mean of 0 and a typical deviation of 1 during normalization using the z-score method, also known as standardization. The mean of the related feature is subtracted from each standardized value before being divided by the quality deviation.

Model selection and training: The model chosen will depend on the qualities of the data as well as the particular specifications of the current issue. The supervised learning algorithms group includes the decision-tree algorithm. It is a tree-structured classifier, where internal nodes stand in for a dataset's features, branches for the decision-making process, and each leaf node for the classification result. Decision trees can be used with ensemble techniques like gradient boosting and random forests to enhance performance and lessen over fitting. Decision trees are a suitable option for big data applications since they can handle enormous datasets with many attributes.When the relationship between the input features and the target variable is complex or challenging to model using a linear function, this is especially helpful. To comprehend how the learned classifier behaves when applied to test or validation datasets, confusion matrix is used.They are taught with ML algorithms after analyzing 4 datasets. You can use the same algorithms for training the k datasets with or without changing the hyper-parameters, or you can use other techniques.With a tweak in hyper-parameters like "depth," Decision Tree methods can be used as basis models. SVM, Naive Bayes, and Logistic Regression are just a few examples of the various algorithms that can be combined.

Model validation is the process of assessing a machine learning model's performance using a different collection of data, called the validation dataset.

The model's capacity to generalize to new data is assessed using the validation datasets, which differs from the training datasets. The methods listed below have been applied to model validation:



Flow chart

Model Optimization: Model optimization is the process of fine-tuning a machine learning model's hyper-parameters to enhance performance. Hyper-parameters are parameters that must be established prior to training since they cannot be learned during training. For this model's optimisation, Grid Search and Gradient Descent have been used:

In this method, a grid of hyper-parameters is established, and the model is trained and assessed for each set of hyper-parameters in the grid. Based on how well the model performs, the ideal set of hyper-parameters is chosen.
By iteratively modifying a model's parameters in the opposite direction as the gradient, the optimization process known as gradient descent is used in machine learning to minimize the cost function of a model. The gradient is the slope of the cost function, and the algorithm can converge to the best set of parameters by going in the direction of a negative gradient. Numerous machine learning algorithms, such as neural networks, logistic regression,support vector machine and linear regression, can be used with gradient descent.

The performance was tested using fresh and untested data.

Heat map, Confusion Matrix, and Correlation Matrix are used by Seaborn to analyse the model's results and visualize them.

By following these procedures, the project development of a crop prediction system utilizing machine learning can be carried out successfully, producing accurate predictions and enabling farmers and policymakers to make well-informed decisions.

3.2 Project Analysis

An AI-based system called the Crop Prediction System makes crop Predictions to farmers based on a variety of environmental criteria, including soil type, weather, and other climatic circumstances. A machine learning technique called ensemble learning combines a number of different separate models to provide more accurate prediction results. The goal of this research is to create an ensemble learning-based crop Prediction system that will aid farmers in maximizing crop productivity and minimising losses brought on by improper crop selection.

Project Objectives The project's objectives are as follows:

- Build a solid model that can accurately estimate crop sustainability in a given state under specific climatic and soil conditions.
- Give advice on the area's best crops to ensure that the farmer doesn't suffer any losses.
- Give a profit analysis of different crops based on data from prior years.

Create an algorithm for crop Predictions using ensemble learning. Gather and prepare information on the different types of soil, the weather, and other environmental elements.For a probability-based strategy, train multiple different models using KNN, Decision Tree, Random Forest, XGBoost, and Naive Bayes.Utilising techniques for ensemble learning, combine the individual models. Analyse the ensemble model's performance and contrast it with that of individual models.

Project Scope: The following is included in the project's scope:

gathering information on the different types of soil, the weather, and other environmental aspects preparing and cleaning the data to get rid of any discrepancies or mistakes choosing and utilizing various machine learning techniques to train distinct models Combining the individual models using ensemble learning approaches assessing the system's performance using data from the actual world

Project Execution technique: The following technique will be used to complete the project:

Data on soil types, weather patterns, and other environmental aspects will be gathered from reputable sources, including government organizations, research institutions, and NGOs, as the first stage. The information will be kept in a database in a structured fashion.

To eliminate any anomalies or discrepancies, the collected data will be preprocessed and cleansed. Data minimization, data transformation, and data cleansing will all be required.

Different machine learning algorithms, including Decision Trees, Random Forests, SVM, KNN, and Naive Bayes, will be used to train a number of different models. A portion of the preprocessed data will be used for the training of each model.

Development of an Ensemble Model: Bagging, Boosting, and Stacking are a few ensemble learning techniques that will be used to merge the individual models. The ensemble model's effectiveness will be assessed using cross validation

Conclusion: The Crop Prediction System using Ensemble Learning is a project that aims to help farmers maximize their crop yield by recommending the most suitable crops based on various environmental factors. The project involves collecting and preprocessing data, training individual models using different machine learning algorithms, combining the individual models using ensemble learning techniques and comparing the performance through accuracy of all the models.

3.2 Project Analysis (SRS)

1. Initialization

Based on a farmer's location, soil type, and weather conditions, this software will suggest the best crop for them. Additionally, the software will advise farmers on how to increase crop yield and suggest crop rotations.

Functional prerequisites:

2.1. Predictions for Crops Based on Location

Farmers will receive crop predictions from the software based on their region, including temperature, humidity, and rainfall.

2.2. Predictions for Soil-based Crops

Based on the type of soil and the amount of nutrients present in the soil, the software will suggest crops to farmers.

2.3 Weather-based Crop prediction

Based on information from the weather forecast, the software will offer crop predictions.

2.4 Fertilizer Prediction

Based on the crop type, N, P, and K values, pH, temperature, moisture, and humidity, the project will recommend fertiliser.

3. Requirements that aren't functional

3.1. Safety

3.2 Performance .

The project will be quick and effective, able to manage many of users, and able to process data instantly.

3.3 Dependability

Real-time data updates, reliable error handling, and recovery methods will make the project dependable.

4. Finality

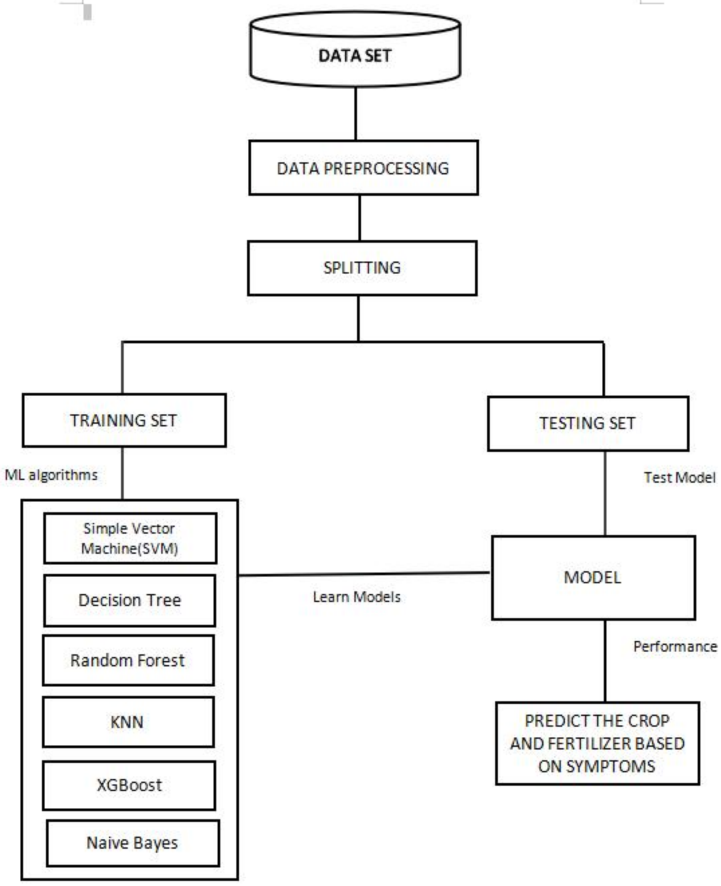
Finally, this research will advise farmers on the appropriate fertiliser to use and assist them in making informed decisions regarding their crops. Farmers will benefit from time and money savings, and overall agricultural output will increase.

3.3 System Design

3.3.1 Design Constraints

- Operating systems: Mac, Linux, Windows XP/7/8/8.1/10, and 11.
- Python is the coding language
- Pandas is the first tool.
  2. Numpy
  3. Tensorflow
  4. Scikit Learn
  5. Seaborn
- IDE : Jupyter

3.3.2 Block Diagram



Block diagram

Chapter 4

Implementation

4.1 Methodology OR Proposal

The following steps can be used to summaries the technique of a crop prediction system that employs machine learning models like logistic regression, SVM, and decision trees:

Data gathering:.
We have features in Fertilizer_prediction.csv that are dependent on the variables, including temperature, humidity, moisture, soil type, crop type, nitrogen, phosphorus, and fertilizer name.
For crop prediction systems, we take into account NPK (nitrogen, phosphorus, and potassium) values, pH (soil acidity/alkalinity), temperature, moisture, humidity, crop type, and various fertilizer values because each of these variables can have a big impact on the growth and yield of crops.

We have taken yield, area, and cost of production into consideration in state_wise_crop _production.csv. We chose these attributes because cost of cultivation and cost of production: These variables give a ballpark figure for production costs, including the price of inputs including seeds, fertiliser, pesticides, labour, and equipment. Making informed decisions regarding resource allocation and figuring out how profitable crop production is requires accurate cost estimation. We have features like State_Name, District_Name, CROP_Year, Season, Crop Area, and Production in crop_production.csv. In order to give extra context and information about the crops being farmed, crop prediction systems frequently include the attributes State_Name, District_Name, CROP_Year, Season, Crop Area, and Production.

In Crop_Prediction.csv, we have features like N, P, K, Temperature, Humidity, PH, rainfall, and Label. Temperature, humidity, pH, and rainfall are important variables that affect crop development and yield, hence crop prediction algorithms typically incorporate them.

Machine learning data preprocessing involves splitting up our datasets into a training set and a test set. One of the most important data pretreatment stages, since it allows us to improve the functionality of our machine learning model.The dataset is divided into training and testing data in an 80:20 ratio.The StandardScaler() function in the Python Sklearn module is used to standardize the data values into a standard format before encoding the categorical data into numerical data. Only data values that adhere to the Normal Distribution can be standardized.

A table of correlation coefficients between variables is called a correlation matrix. The correlation between any two variables is shown in each cell of the table. The value ranges from -1 to 1. Using the pandas library, data is summarized using a correlation matrix, which is also used as a diagnostic tool and as an input for more sophisticated analyses.We use the corr() method on data-frames in order to generate a correlation matrix for a given datasets.

We have used the Pandas library's iloc[] method to extract an independent variable. It is used to remove the necessary columns and rows from the datasets. The handling of missing data in the datasets is the next stage of data preprocessing. Our machine learning model may run into serious issues if our datasets has some missing data.To guarantee that the model generates precise Predictions, these data are dependable and thorough.You can always patch up the missing information using Python's adaptable tools rather than deleting the records or columns.The simplest way for filling in NaNs with precise values is pandas.Data-frames.fillna().

Model selection: KNN, SVM, Decision Tree, Random Forest, and Naive Bayes are just a few of the machine learning models that can be used to predict crops. The particulars of the problem and the qualities of the data will determine which model is chosen.

Model Training: The non-parametric supervised machine learning technique k-nearest neighbors classifier (kNN) is used. It is distance-based and categorizes things according to the classes of their close neighbors. Although it can be used for regression issues as well, kNN is most frequently used for classification tasks.The K in this classifier's name stands for the k nearest neighbors, where k is an integer value that the user specifies. Data determines what value of k should be used. It is calculated via a simple majority vote among each point's closest neighbors.It is a form of non-generalizing learning because it merely maintains instances of the training data.

SVMs are frequently employed in classification issues and are extremely effective in large dimensional spaces. Because they only use a portion of the training points in the decision function, SVMs are well-liked and memory-efficient algorithms.To find a maximum marginal hyperplane (MMH), SVMs partition datasets into a number of classes. This can be accomplished in the following two steps: Support Vector Machines will initially iteratively create hyper-planes that best divide the classes. The hyperplane that properly separates the classes will then be selected.

The Naive Bayes classifier has proven to be effective. It is the simplest and fastest classification approach currently known, and it is ideally suited for dealing with vast amounts of information. The Bayes theory of probability is used to generate predictions about unknowable classes.

Model Evaluation: Using several error metrics, we may assess the performance of the model using this technique. There are many different error measures that are used to evaluate performance in classification and regression tasks, including accuracy, confusion matrix, mean squared error, and mean absolute error. We created our model with classification in mind, therefore we would compute metrics for measuring the effectiveness of the classification model.

Model Optimization: The model can be improved by modifying the hyper-parameters based on the findings of the evaluation. Here, the model's performance has been optimized using GridSearchCV and Gradient Descent.

GridSearchCV is a method for adjusting hyper-parameters to find the best values for a particular model.The value of a model's hyper-parameters has a substantial impact on its performance. Noting that there is no way to determine the best values for hyper-parameters in advance, it is ideal to explore every conceivable value before deciding what the best ones are. GridSearchCV was used to automate the tuning of hyper-parameters because doing so manually may take a lot of time and resources.The model_selection package of Scikit-learn (aka SK-learn) contains a function called GridSearchCV.Therefore, it's crucial to note that the Scikit-learn library must be installed on the computer. This function aids in fitting your estimator (model) to your training set by looping through predefined hyper-parameters. So, from the list of hyper-parameters, we can choose the best parameters in the end.

By iteratively modifying a model's parameters in the opposite direction as the gradient, the optimization process known as gradient descent is used in machine learning to minimize the cost function of a model. The gradient is the slope of the cost function, and the algorithm can converge to the best set of parameters by going in the direction of a negative gradient.

Visualization: A graphical representation of data is used to examine outliers, trends, the distribution of the data, and other factors. Python has a variety of libraries for data visualization, such as matplotlib, seaborn, etc. To visualize the pair-plots, we'll utilize the Seaborn library. The Seaborn library will be imported first, and the pair-plot will then be printed. The values of the matrix are depicted as colors in a heat-map, which is a graphic representation of the matrix. A heat-map can be used in a machine learning setting to show the relationship between several datasets elements.

A table called a confusion matrix is used to assess how well a categorization model is working. It contrasts the test set's actual labels with the model's anticipated labels.

A statistic called the F1 score combines recall and precision into one number. It is determined as the harmonic mean of recall and precision:
F1 Score is equal to 2 * (Precision * Recall) / (Precision + Recall).
The F1 score runs from 0 to 1, with 1 denoting flawless recall and precision. A higher F1 score suggests better performance in categorization tasks.

The Jaccard score, commonly referred to as the Jaccard index, is a different metric used to assess how well categorization models work. It is determined by dividing the area of the intersection by the combined area of the expected and actual labels.

Jaccard Score is the intersection of the predicted and the actual labels, or the union of the two.

The performance of a binary classifier as the discrimination threshold is changed is depicted on a figure called the ROC curve, which stands for receiver operating characteristic curve. By comparing the true positive rate (TPR) and false positive rate (FPR) at various threshold settings, it is produced.

A typical statistic for assessing the effectiveness of binary classifiers is the area under the ROC curve (AUC-ROC). It has a scale from 0 to 1, with 1 denoting ideal performance.
Recall Rating
The recall score, sometimes referred to as sensitivity, gauges how many actual positives the model accurately detected. The ratio of true positives to the total of true positives and false negatives is used to compute it.
Recall is defined as True Positives / (False Negatives + True Positives).
The recall score ranges from 0 to 1, with 1 representing flawless performance in correctly detecting every genuine positive.

Precision Score:
The proportion of anticipated positives that the model properly identifies is represented by the accuracy score. The ratio of true positives to the total of both true and false positives is used to compute it.
Precision is equal to the product of true positives and false positives.
A precision score of 1 indicates perfect accuracy in properly detecting all anticipated positives. The precision score has a range of 0 to 1.
In general, data gathering, preprocessing, model selection, model training, model evaluation, model optimization, and deployment are all steps in the approach of a crop prediction system employing machine learning models. The system's success will be based on the reliability of the data, the choice of the best model, and the precision of the forecasts.

4.2 Testing OR Verification Plan

We have divided our datasets in 80:20 ratio in data preprocessing in train-test split. We used our test data to measure the accuracy of different models and we got a good accuracy on each datasets on each model.

Decision Tree-
Decision Tree model accuracy of Dataset 1 on training set and test set are :
1.0
7621212121212121
Decision Tree model accuracy of Dataset 2 on training set and test set are :
1.0
0.9333333333333333
Decision Tree model accuracy of Dataset 3 on training set and test set are :

0.1895586512172894
0.028857837181044958

Decision Tree model accuracy of Dataset 4 on training set and test set are :
1.00.6

Random Forest-
Random Forest model accuracy of Dataset 1 on training set and test set are :
1.0
0.803030303030303
Random Forest model accuracy of Dataset 2 on training set and test set are :
1.0
1.0
Random Forest model accuracy of Dataset 3 on training set and test set are :
0.1895586512172894
0.028857837181044958
Random Forest model accuracy of Data set 4 on training set and test set are :
1.0
0.4666666666666667

XgBoost-
XgBoost model accuracy of Dataset 1 on training set and test set are :
1.0
0.7803030303030303
XgBoost model accuracy of Dataset 2 on training set and test set are :
1.0
0.9666666666666667
XgBoost model accuracy of Dataset 3 on training set and test set are :
0.1883869287853144
0.02976913730255164
XgBoost model accuracy of Dataset 4 on training set and test set are
1.0
0.5333333333333333

SVM-
SVM model accuracy of Dataset 1 training set and test set are :
0.7383116883116884
0.7075757575757575SVM model accuracy of Dataset 2 training set and test set are :
1.0
1.0

SVM model accuracy of Dataset 3 training set and test set are :
0.08501497200885301

0.07229647630619684
SVM model accuracy of Dataset 4 training set and test set are :
0.6470588235294118

0.3333333333333333

KNN-
KNN model accuracy of Dataset 1 training set and test set are :
0.8175324675324676
0.7636363636363637

KNN model accuracy of Dataset 2 training set and test set are
0.9130434782608695
0.6333333333333333

KNN model accuracy of Dataset 3 training set and test set are
0.12172894154406978
0.015188335358444714

KNN model accuracy of Dataset 4 training set and test set are
0.6470588235294118
0.6

Naive Bayes-
Naive Bayes model accuracy of Dataset 1 on training set and test set are
0.7967532467532468
0.7772727272727272

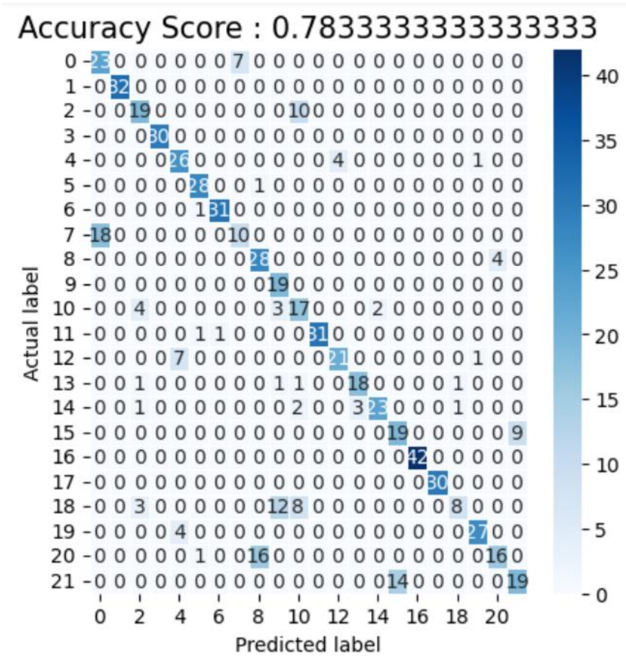
Naive Bayes model accuracy of Dataset 2 training set and test set are :
1.0
1.0

Naive Bayes model accuracy of Dataset 3 training set and test set are :
0.09425856008332248
0.08262454434993925

Naive Bayes model accuracy of Dataset 4 training set and test set are :
0.8529411764705882
0.6666666666666666

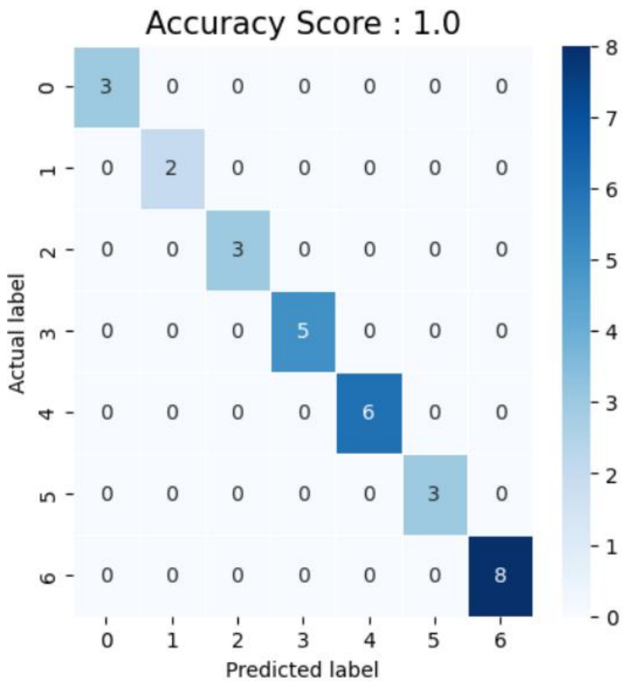
4.3 Result Analysis OR Screenshots

Accuracy of the combined model for dataset 1: 0.7833333333333
Score for accuracy: 78.33% F-1 Score: 0.7833333333333333
Score for precision: 0.7833333333333333
Score for Recall: 0.7833333333333
Score on the Jaccard: 0.6438356164383562

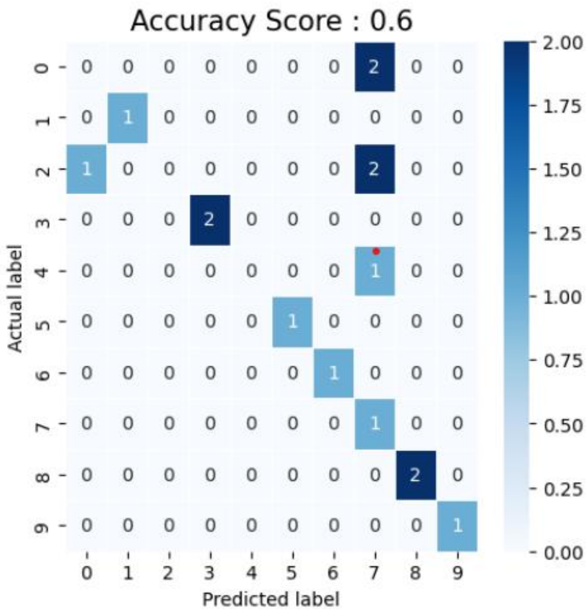


Accuracy of the combined model for dataset 2: 1.0
Score for Accuracy : 100% F-1 Rating: 1.0
Score for precision: 1.0
Score for Recall: 1

Score for Jaccard: 1



Accuracy of the combined model for dataset 3: 60%
Score for Accuracy : .6 % F-1 Rating: 0.6
Score for precision: 0.6
Score for Recall: 0.6
Score on the Jaccard: 0.42857142857142855



4.4 Quality Assurance

Members of our team come from a variety of backgrounds, and we explore potential problems with a focus on potential harm to vulnerable groups (such as local farmers).

A literature search on industry standards and Predictions for data science and agriculture.

A plan for mitigation was made.

A report were made on forecasts that were accurate and effective.

Chapter 5

Conclusion and Future Scope of the project

5.1 Conclusion

A ML model have been made that has achieved a good accuracy on test data that is accuracy of 1 on fertilizer datasets accuracy of .8 on crop datasets and accuracy of .77 on location datasets.we have achieved this accuracy by using different ml models in ensemble learning. This project has a huge potential it can serve millions of farmers in future and solve their most of the problems.

5.2 Future Scope

The future scope of this project is that we can also use weather, climate , and others varying features also we can deploy it on web or app that can be used by farmers and they can utilize it to maximize their profit. We can also collect more and accurate data from farmers by survey or google form.

References

[1]R. M. Osgood, Jr., Ed., S. M. Metev and V. P. Veiko, Laser Assisted Microtechnology, 2nd ed. Berlin, Germany: Springer-Verlag, 1998.

Lecture Notes in Statistics, ed. Breckling, The Analysis of Directional Time Series: Applications to Wind Speed and Direction. 1989, vol. 61, Berlin, Germany: Springer.

"A novel ultrathin elevated channel low-temperature poly-Si TFT," IEEE Electron Device Lett., vol. 20, no. 11, Nov. 1999, pp. 569–571. S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok.

"High resolution fibre distributed measurements with coherent OFDR," Proc. ECOC'00, 2000, paper 11.3.4, p. 109, by M. Wegmuller, J. P. von der Weid, P. Oberson, and N. Gisin.

"High-speed digital-to-RF converter," U.S. Patent 5 668 842, September 16, 1997, R. E. Sorace, V. S. Reinhardt, and S. A. Vaughn.

The IEEE website, 2002. [Online]. Accessible at: www.ieee.org

[Online] [7] M. Shell. IEEEtran homepage on CTAN. Easily accessible at: www.ctan.org/tex-archive/macros/latex/contrib/supported/IEEEtran

Crop and Fertilizer Prediction

Saksham Raj
2005473

Abstract: Crop and fertilizer Recommmendation System will help farmers to optimize their crop yield by providing personalized Predictions for the type and amount of fertilizers to use. The project is built using different machine learning algorithms trained on soil data, location, crop-specific information and cost of cultivation and production of crops. The project focuses on individual contribution, which means that each farmer will have their own personalized Prediction based on their specific crop and soil conditions.The system will also provide real-time feedback to farmers, allowing them to adjust their fertilizer usage as necessary.Overall, the goal of the project is to improve crop yields, decrease environmental impact, and increase profitability for small-scale farmers.

Individual contribution and findings: Once the data was cleaned and processed, We moved on to model training, where I have used various supervised machine learning algorithms(since out dataset have output labels) such as Random Forest, Decision tree, Gaussian Naive Bayes, Knn, XGBoost, and SVM to build predictive model. i experimented with different model architectures and tuned hyperparameters using grid search and gradient descent to find the best-performing model.

In this project we want to use one model in outcome of a new data point is predicted based on its proximity(distance metric) to other data points in the training dataset hence we used k Nearest Neighbour model.

The likelihood that a new data point will belong to each class in the training dataset is used to classify new data points in a probabilistic machine learning method. The predicted class label for the test data point is then given to the class with the highest probability. Therefore, to calculate the conditional probabilities, i used the Naive Bayes model, in which each feature is considered to be independent of and unrelated to the other features.Large datasets can benefit from naive Bayes because it trains rapidly and uses less computer power. I used Gaussian Naive Bayes because the features follow a normal distribution.

I have also employed the Support Vector Machine (SVM) model, which divides the data into various classes by determining the best hyperplane, with each data point being represented as a vector in an n-dimensional space (number of features). The hyperplane that maximizes the distance between the closest points of the two classes is found by the algorithm.Since SVM can handle both linearly and non-linearly separable datasets, we used it.

I have also used several tree-based models such as Decision Trees, Random Forests and Extreme Gradient Boosting (XGBoost) because tree based models can handle high-dimensional data and nonlinear relationships between features

Decision Trees: Decision Trees are a simple tree-based model that creates a single tree with nodes representing decisions based on features. Decision Trees produce white-box models meaning the decision process is easy to understand and visualize.

Random Forests: Since decision trees are prone to overfitting, i employed an ensemble method, which mixes a number of decision trees that have been trained on various samples of the training data. Additionally, each tree in the forest has been trained using a random subset of features. By averaging the predictions of several trees, this technique improves accuracy and decreases variance, which reduces overfitting.

Extreme Gradient Boosting (XGBoost): XGBoost is an optimized variant of GBMs that uses a more robust regularized loss function and faster computational techniques to build better models faster. XGBoost also allows for parallelization of the training process, making it salable to larger datasets.

Finally, I used Ensemble learning, a machine learning technique that enhances the accuracy and resilience of predictive models by integrating several models, to aggregate the results from all four types of machine learning models. Ensemble models typically perform better in terms of accuracy and generalization. Additionally, they can withstand noisy data and outliers better. On the training set, ensemble models are less likely to overfit.

Model Optimization:

After developing initial models, I focused on optimizing them further to achieve higher accuracy and reduce overfitting.

I employed regularization techniques to prevent overfitting and improve generalization performance.

I have used grid search to find optimal value of k, locations of initial centroids.

For models like SVM, I have used gradient descent as an optimisation algorithm to identify the best set of parameters. The purpose of gradient descent is to reduce the model's cost function or objective function.

Conclusion: Our team used a variety of machine learning algorithms to create a reliable crop and fertilizer Prediction system. Model training and optimization were the main focuses of my contribution to the project. We were able to create a trustworthy predictive model that could give farmers precise Predictions by merging different machine learning approaches and optimization methodologies. The system has the potential to transform agricultural methods, boosting crop yields and enhancing the sector's bottom line.

Individual contribution to project report preparation: In project implementation I tried to choose best possible ways to do all the things to achieve the best possible result. In Implementation I have specified all the tools/models that we are using in our project

Methodology: The purpose of this project was to create a machine learning-based Prediction system for crops and fertilizers. The suggested methodology included data collection, feature engineering and preprocessing, model construction, tuning of the hyper-parameters, and testing.

We collected data from various sources on soil types, climate patterns, crop varieties, and fertilizer compositions. We cleaned and processed the data, handling missing values, and ensuring it was in a format suitable for machine learning algorithms.

Next, we worked on feature engineering, where we selected relevant features from the dataset to improve the model's accuracy. We identified factors such as location, moisture, soil pH level, n, p, k value, cost of cultivation and production and crop type to understand how they impact crop growth and yield. We also used domain knowledge to engineer new features related to plant physiology and soil chemistry.

Using various machine learning methods, including Random Forest, Decision Trees, Naive Bayes, Support Vector Machines, Knn, and Extreme Gradient Boosting (XGBoost), we created a number of models. To choose the model with the highest performance, we experimented with several model topologies and parameter choices. We tested the model to ensure its capacity for generalisation and to guard against over fitting.

We utilized grid search, gridant descent to tune hyper-parameters and optimize the model performance. Once the optimal set of hyper-parameters were identified, we trained the model on the entire dataset.

Testing / Verification Plan:To test the model's performance, we conducted model accuracy testing on test data. We evaluated our model's generalization ability by comparing the training and testing performance metrics. We also performed error analysis to identify specific areas for improvement.

Result Analysis / Screenshots: Our model achieved an accuracy of 100% on dataset 2 and 4 and accuracy of 80% on dataset1.The optimized models outperformed baseline models and provided accurate crop and fertilizer Predictions.We implemented the Prediction system to suggest farmers personalized Predictions for crops and fertilizers. Screenshots of the application were taken.

Quality Assurance: To ensure the quality of the project, we followed best practices in software development and machine learning. We wrote clean and well-documented code, conducted thorough testing, and implemented version control using Git. We also ensured that the data collection and preprocessing were accurate and followed ethical standards. Additionally, we documented our methodology and results to ensure reproducibility of the project

Individual contribution for project presentation and demonstration: I have prepared the model training and model optimization slides in the ppt and demonstrated the same as well.

Full Signature of Supervisor:

……………………………

Full signature of the student:

Saksham raj

Crop and Fertilizer Prediction System

NIVEDITA SUTRADHAR
20051550

Abstract: Crop and fertilizer Prediction Project will help farmers to optimize their crop yield by providing personalized Predictions for the type and amount of fertilizers to use. The project is built using different machine learning algorithms trained on soil data, location, crop-specific information and cost of cultivation and production of crops. The project focuses on individual contribution, which means that each farmer will have their own personalized Prediction based on their specific crop and soil conditions.The system will also provide real-time feedback to farmers, allowing them to adjust their fertilizer usage as necessary.Overall, the goal of the project is to improve crop yields, decrease environmental impact, and increase profitability for small-scale farmers.

Individual contribution and findings: As a data preprocesser for the crop and fertilizer Predictions system, my primary responsibility was to prepare the dataset for subsequent analysis. To achieve this, I performed the following tasks:

Loading Data: I loaded the data into notebook using pandas that is collected by my teammate from different sources.I have looked at description of each datasets such as mean, standard deviation, range of values

Data cleaning: I cleaned the raw data by removing irrelevant information, correcting errors, and filling in missing values. This involved using statistical techniques such as mean imputation and outlier detection. Additionally, I also created a unique identifier for each record to facilitate future analysis. For missing values if feature contain numerical values I assigned median of that feature and if feature contain categorical value I assigned mode of that feature. For outlier detection I have used iqr method.

Data transformation: I transformed the cleaned data into a format that could be easily analyzed. This involved converting categorical variables into numerical values, scaling the data, and normalizing the distribution of the data.for data transformation I have used LabelEncoder(). for normalizing the distribution in same range I have used StandardScaler().

Feature engineering: I engineered new features from the existing ones to enhance the predictive power of the model. I removed the features that are not relevant for the project.

Input- output split: for each dataset I have divided it into input features that is expected to value received from user/farmers and output features/class that is the class of crop predicted or fertilizer suggested.

Test and train split: I have divided all dataset in 70, 30 ratio for training and testing purpose.

Data visualization: I have drawn pairwise plot to see relationship between each two features to analyse their relationship. I have also drawn correlation matrix to see how all features are related if two features are highly correlated and its value in heatmap of correlation matrix is close to 1 then I have removed one o9f the features

Conclusion:
In conclusion, my contribution to the crop and fertilizer Predictions system was crucial in ensuring that the data was accurate, reliable, and well-prepared for analysis. By performing data collection, cleaning, transformation, feature engineering,input-output split, test-train split, and visualization. I helped to create a robust dataset that was used to build a highly effective Prediction system.

Individual contribution to project report preparation: Literature review is an essential section of any research project report that involves analyzing published works relevant to the topic of study. My aim during literature review is to gain an overview of existing literature related to the research question, identify gaps in knowledge, and see areas for further study.

While Preparing a literature review I have gone thorough various research papers and analyzed them. It involves searching scholarly articles, conference papers, and other relevant sources of information through academic libraries, online databases, and search engines.

Crop and fertilizer Prediction systems are becoming increasingly popular as farmers, agronomists, and researchers seek more efficient ways to optimize crop yields while reducing the environmental impact(pollution) of agriculture. In this literature review, i explored the current state-of-the-art in crop and fertilizer Prediction systems, including their benefits, limitations, and potential for future development.

One of the key benefits of crop and fertilizer Prediction systems is the ability to tailor Predictions to specific soil and crop conditions. This allows farmers to optimize crop yields while minimizing nutrient runoff and other environmental impacts. Additionally, these systems can help reduce costs associated with over- or under-fertilization, as well as provide a basis for precision agriculture techniques such as variable rate application.

However, there are also challenges associated with developing accurate and effective crop and fertilizer Prediction systems. For example, factors such as weather, pests, and disease can significantly affect crop performance, making it difficult to predict optimal fertilization rates. Additionally, different crops may have different nutrient requirements, which can further complicate the development of universal Prediction systems.

Recent research has explored a variety of approaches to address these challenges, including machine learning algorithms, sensor-based technologies, and improved remote sensing techniques. These approaches have shown promising results in increasing the accuracy and efficiency of crop and fertilizer Prediction systems, but further research is needed to fully realize their potential.

Overall, crop and fertilizer Prediction systems have the potential to revolutionize the way we approach agriculture by improving crop yields and reducing environmental impacts. Continued research and development in this area will be crucial for realizing these benefits and addressing the challenges associated with developing effective Prediction systems.

Individual contribution for project presentation and demonstration:
I prepared the data preprocessing slides in the presentation and demonstrated the same as well.

Full Signature of Supervisor: ………………………………

Full signature of the student:

Nivedita Sutradhar

Crop and Fertilizer prediction System

KV Prashant
2005102

Abstract: Crop and fertilizer prediction System will help farmers to optimize their crop yield by providing personalized predictions for the type and amount of fertilizers to use. The project is built using different machine learning algorithms trained on soil data, location, crop-specific information and cost of cultivation and production of crops. The project focuses on individual contribution, which means that each farmer will have their own personalized prediction based on their specific crop and soil conditions.The system will also provide real-time feedback to farmers, allowing them to adjust their fertilizer usage as necessary.Overall, the goal of the project is to improve crop yields, decrease environmental impact, and increase profitability for small-scale farmers

Individual contribution and findings:
We have collected four recent datasets from Kaggle for the last 5 years for exploring the crop production in India through researching the state wise crop_production.csv, Fertilizer_prediction.csv, crop_prediction.csv, state_wise_crop _production.csv.

I have worked on two of the datasets : (1)Fertilizer_prediction.csv
(2) state_wise_crop _production.csv

We have also worked on the result analysis of the different machine leaarning algorithms.

In Fertilizer_prediction.csv we have features like Temperature,Humidity,Moisture,Soil type, Crop Type, Nitrogen, Phosphorus, Potassium and Fertilizer name depending on the variables. We take NPK (Nitrogen, Phosphorus, and Potassium) values, pH (soil acidity/alkalinity), temperature, moisture, humidity, crop type, and different fertilizer values into consideration for crop prediction systems because each of these factors can have a significant impact on the growth and yield of crops.

NPK values: Nitrogen, Phosphorus, and Potassium are essential nutrients required by plants in varying amounts. Nitrogen is critical for plant growth and is a component of chlorophyll, which is necessary for photosynthesis.

pH: . Plants require specific pH ranges to absorb nutrients effectively, and soil pH that is too high or too low can make essential nutrients unavailable to plants. The pH values will help determine the acidity and alkalinity which can affect the availability of nutrients to plants.

Different crops have varying nutrient and fertilizer requirements. Fertilizers can supply essential nutrients to crops, but their application needs to be appropriate for each crop and soil type. Using the wrong fertilizer or applying it incorrectly

can cause nutrient imbalances or even damage plants.
In summary, taking these factors into account can help to predict the growth and yield of crops accurately, and recommend appropriate fertilizers to maximize crop yield while minimizing environmental impact.

In state_wise_crop _production.csv we have features like cost of production, cost of cultivation, yield , area into consideration, we have selected such features beacause:

Cost of cultivation and cost of production: These factors provide an estimate of the expenses incurred in the production process, including the cost of inputs such as seeds, fertilizers, pesticides, labor, and machinery. Accurately estimating these costs is important in determining the profitability of crop production and in making informed decisions about resource allocation.

The area of the land under cultivation is a critical factor in crop prediction systems, as it determines the total production volume and the potential revenue. The area also affects the amount of inputs required, such as fertilizers and pesticides, and the labor needed for planting, harvesting, and other activities.

By taking these factors into account, crop prediction systems can provide farmers with accurate information on the potential profitability of different crops and help them to make informed decisions about resource allocation, input management, and marketing.

Result analysis:
We have split the datasets into Testing data and Training data in the ratio of 80:20.

Different models like KNN for distance based approach, SVM, Decision Tree, Random Forest, Naive Bayes for probability based approach have been applied and the accuracy is been analyzed.

Individual contribution to project report preparation:
I have written the future scope of our project and the problems it can resolve for the betterment of the society.Here are some potential future developments in this field:
As machine learning algorithms become more sophisticated and advanced, they will be able to predict crop yields with greater accuracy. This will enable farmers to plan their planting and harvesting schedules more effectively and optimize resource utilization.

Crop prediction systems can be enhanced by integrating data from other sources such as weather forecasts, soil moisture sensors, and satellite imagery. This will enable the systems to account for a wider range of factors that influence crop yields, and provide more accurate predictions.

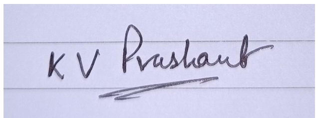
Crop prediction systems using machine learning can be scaled up to handle larger volumes of data and provide predictions for larger areas With the help of machine learning, crop prediction systems can provide personalized
predictions to individual farmers based on their specific soil types, weather patterns, and other factors. This will enable farmers to optimize their crop yields and maximize their profits.

Crop prediction systems can be integrated with farm equipment and automated
farming systems, enabling farmers to take actions based on the predictions generated by the system.In conclusion we can say that the future of crop prediction systems using machine learning is promising, with potential for increased accuracy, scalability, and personalized predictions, as well as integration with other data sources and automation.

Individual contribution for project presentation and demonstration: I have done the result analysis of the ML models and have added the conclusion of the project in the presentation and will demonstrate the same part as well.

Full Signature of Supervisor: ……………………………

Full signature of the student:



Crop and Fertilizer Prediction System

SUMITA CHAKRAVARTY
2005621

Abstract: Crop and fertilizer prediction System will help farmers to optimize their crop yield by providing personalized predictions for the type and amount of fertilizers to use. The project is built using different machine learning algorithms trained on soil data, location, crop-specific information and cost of cultivation and production of crops. The project focuses on individual contribution, which means that each farmer will have their own personalized prediction based on their specific crop and soil conditions.The system will also provide real-time feedback to farmers, allowing them to adjust their fertilizer usage as necessary. Overall, the goal of the project is to improve crop yields, decrease environmental impact, and increase profitability for small-scale farmers

Individual contribution and findings:

My role in the project was to collect the dataset for crop and fertilizer prediction system and to analyse the result of different machine learning algorithms of K-Nearest Neighbours (KNN), Support Vector Machine(SVM), Naive Bayes. We have collected four recent datasets from Kaggle for the last 5 years for exploring the crop production in India

I have worked on the the following two data sets
1) Crop_prediction.csv
2) crop_production.csv

In Crop_prediction.csv we have features like N, P, K, Temperature, Humidity, PH, rainfall, Label

The attributes N, P, K, temperature, humidity, pH, and rainfall are commonly used in crop prediction systems because they are important factors that influence crop growth and yield.

N, P, and K are the three primary macronutrients required by plants for healthy growth. The presence or absence of these nutrients in the soil can have a significant impact on crop yield and quality.

Temperature and humidity are important environmental factors that affect plant growth and development. Different crops have different temperature and humidity requirements, and these factors can vary significantly depending on the season and location.

pH is a measure of the acidity or alkalinity of the soil, and it can have a significant impact on plant growth and nutrient availability. Different crops have different pH requirements, and adjusting the pH of the soil can be an important factor in maximizing crop yield and quality.

The label attribute is used to identify the crop being grown and is the output of the prediction system. By analyzing the input attributes (N, P, K, temperature, humidity, pH, and rainfall) and comparing them to known patterns of crop growth and yield, a prediction system can identify the most likely crop and provide predictions for optimal growing conditions.

Overall, these attributes are chosen because they are critical factors in determining crop growth and yield and are relatively easy to measure or estimate. By analyzing these attributes, a crop prediction system can help farmers make more informed decisions about how to optimize their crop yields and improve their overall agricultural productivity.

In crop_production.csv we have features like State_Name, District _Name, CROP_Year, Season, Crop Area, Production

The attributes State_Name, District_Name, CROP_Year, Season, Crop Area, and Production are commonly used in crop prediction systems to provide additional context and information about the crops being grown.

State_Name and District_Name help to identify the specific location where the crops are being grown. Different regions can have different climate conditions, soil types, and other environmental factors that can affect crop growth and yield. Knowing the location can help to provide more accurate predictions and predictions.

CROP_Year provides information about the year in which the crops were grown. This can be important because different years can have different weather patterns, disease and pest outbreaks, and other factors that can affect crop growth and yield.

Season helps to identify the specific time of year when the crops are being grown. Different crops have different growing seasons, and the timing of planting and harvesting can be critical factors in determining crop yield and quality.

Crop Area and Production provide information about the size of the crop and the amount of yield produced. This information can be used to analyze trends over time and identify patterns in crop growth and yield.

By analyzing these additional attributes, a crop prediction system can provide more accurate and personalized predictions to farmers. For example, the system may recommend different planting times or fertilizer application rates based on the specific location, season, and previous crop yields. This can help farmers to optimize their crop yields and improve their overall agricultural productivity.

Result Analysis

We have split the datasets into Testing data and Training data in the ratio of 80:20.

Different models like KNN for distance based approach, SVM, Decision Tree,

Random Forest, Naive Bayes for probability based approach have been applied and
the accuracy is been analyzed.

I have tested on the following algorithms

Individual contribution to project report preparation:

My individual contribution to project report preparation was the introduction part, Problem Statement / Requirement Specifications, Project Planning, Project Analysis and Block Diagram

Problem Statement

In the problem statement we have written the need to improve crop yields by providing accurate and personalized predictions for farmers. Farmers face the challenge of selecting the right crops to plant and the appropriate fertilizers to use to maximize yield, while minimizing costs and environmental impact.

Project Planning

In project planning I have written about the are various steps that should be followed while planning to execute the project development of a crop prediction system using machine learning

Data collection and pre processing
Model selection and training:
Model validation

Model Optimization Here Grid Search and Gradient Descent has been used for model optimization:

Testing: Test its performance on new, unseen data.

The interpretation of the results of the model and visualization them is done through heatmap,confusion matrix and correlation matrix using seaborn

By following these steps, the project development of a crop prediction system using machine learning can be executed efficiently and effectively, leading to accurate predictions and informed decision-making for farmers and policymakers.

Software Requirements Specification (SRS)
Software Requirements Specification (SRS) is a critical component of software development. It is a document that outlines the functional and non-functional requirements of a software system. In the SRS section I have briefly explained about the functional and non functional requirements.

Block Diagram

A block diagram is a diagram of a system in which the principal parts or functions are represented by blocks connected by lines that show the relationships of the blocks. They are heavily used in engineering in hardware design, electronic design, software design, and process flow diagrams.

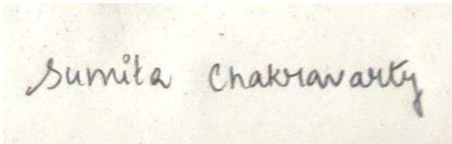
We have made block diagram through Ms Word using different shapes and blocks connected through lines.

Individual contribution for project presentation and project demonstration

My role in preparing presentation and demonstration was preparing the first 5 slides and demonstrating the introduction and data collection part.

Full Signature of Supervisor:

Full signature of the student:



TURNITIN PLAGIARISM REPORT

CROP AND FERTILIZERPREDICTION

ORIGINALITY REPORT

18%	9%	9%	11%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	J. Liu, L. Zhou. "QSAR modeling for thiolactomycin analogues using genetic algorithm optimized artificial neural networks", Molecular Simulation, 2007 Publication	2%
2	Submitted to University of North Texas Student Paper	1%
3	www.worldleadershipacademy.live Internet Source	1%
4	Submitted to Asia Pacific University College of Technology and Innovation (UCTI) Student Paper	1%
5	sron.nl Internet Source	1%
6	www.coursehero.com Internet Source	1%
7	Submitted to Cerritos College Student Paper	1%

8	Submitted to University of Wales Institute, Cardiff Student Paper	1%
9	Submitted to University of Leicester Student Paper	1%
10	Submitted to North South University Student Paper	<1%
11	Silakari, P.. "QSAR analysis of 1,3-diaryl-4,5,6,7-tetrahydro-2H-isoindole derivatives as selective COX-2 inhibitors", European Journal of Medicinal Chemistry, 200807 Publication	<1%
12	Submitted to Rivier University Student Paper	<1%
13	asianssr.org Internet Source	<1%
14	Submitted to Nottingham Trent University Student Paper	<1%
15	Submitted to University of Lincoln Student Paper	<1%
16	www.science.gov Internet Source	<1%
17	Submitted to Berlin School of Business and Innovation Student Paper	<1%

18	Premasudha B G, Thara D K, Tara K N. "ML based methods XGBoost and Random Forest for Crop and Fertilizer Prediction", 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN), 2022 Publication	<1 %
19	Submitted to The Robert Gordon University Student Paper	<1 %
20	Submitted to Gitam University Student Paper	<1 %
21	Submitted to Queen Mary and Westfield College Student Paper	<1 %
22	academic-accelerator.com Internet Source	<1 %
23	www.ijraset.com Internet Source	<1 %
24	Submitted to Indian Institute of Technology, Bombay Student Paper	<1 %
25	Submitted to Liverpool John Moores University Student Paper	<1 %
26	Submitted to University of Salford Student Paper	<1 %