**Mini Project Report on**



**Crop Recommendation System**



**Submitted in partial fulfillment of the requirement for the award of the degree of**

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE & ENGINEERING**

**Submitted by:**

**Student Name: Rameshwari Kapoor University Roll.No.:** 2119003

***Under the Mentorship of***

**Ms. Preeti Chaudhary**

**Assistant Professor**



**Department of Computer Science and Engineering**

**Graphic Era Hill University**

**Dehradun, Uttarakhand**



**CANDIDATE’S DECLARATION**

I hereby certify that the work which is being presented in the project report entitled **“Crop Recommendation System”** in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineeringof the Graphic Era Hill University, Dehradun shall be carried out by myself under the mentorship of **Ms. Preeti Chaudhary, Assistant Professor**, Department of Computer Science and Engineering, Graphic Era Hill University, Dehradun.

**Name: Rameshwari Kapoor** **University Roll no.:**2119003

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **Chapter No.** | **Description** | **Page No.** |
| Chapter 1 | Introduction about Machine Learning | 4 |
| Chapter 2 | Crop Recommendation System Briefing | 6 |
| Chapter 3  Chapter 4 | Methodology  Code | 9  11 |
| Chapter 5 | Need of Crop Recommendation System | 19 |
| Chapter 6 | Application | 20 |
| Chapter 7 | Conclusion | 21 |

**Chapter 1**

**Introduction**

**Machine Learning**

The term ‘Machine Learning’ was coined by Arthur Samuel in 1959. Machine learning is a rapidly expanding field of artificial intelligence that allows computers to automatically infer information from historical data. Machine learning employs a variety of methods to create mathematical models and make predictions using knowledge or historical data. Currently, it is being used for various tasks such as house price prediction, crop recommendation system, image/speech recognition, email filtering, Facebook auto-tagging, recommender system, Natural Language Processing and many more.

* Classification

There are several types of machine learning, including supervised learning, unsupervised learning, and reinforcement learning.

* Supervised learning

Supervised learning, also known as supervised machine learning, is a subcategory of  machine learning and artificial intelligence. It uses labelled datasets to train algorithms to classify data or predict outcomes accurately.

* Unsupervised learning

Unsupervised Machine Learning is a type of machine learning that utilizes unsupervised algorithms to classify and group un-labelled datasets. These algorithms can identify hidden patterns or clusters of data without human intervention.

* Reinforcement learning

RL is a form of machine learning in which an agent learns through trial and error based on feedback from its actions and experiences in an interactive setting.

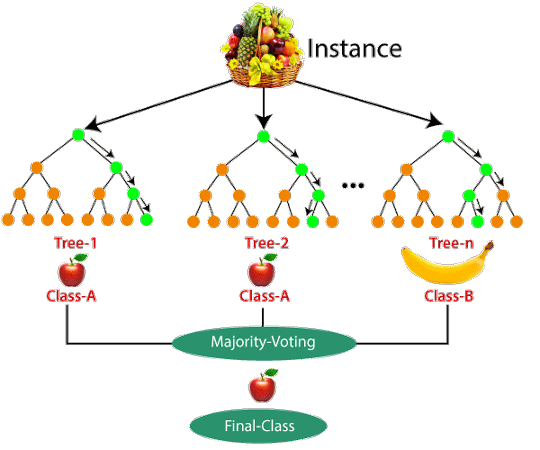
**Chapter 2**

**Crop Recommendation System**

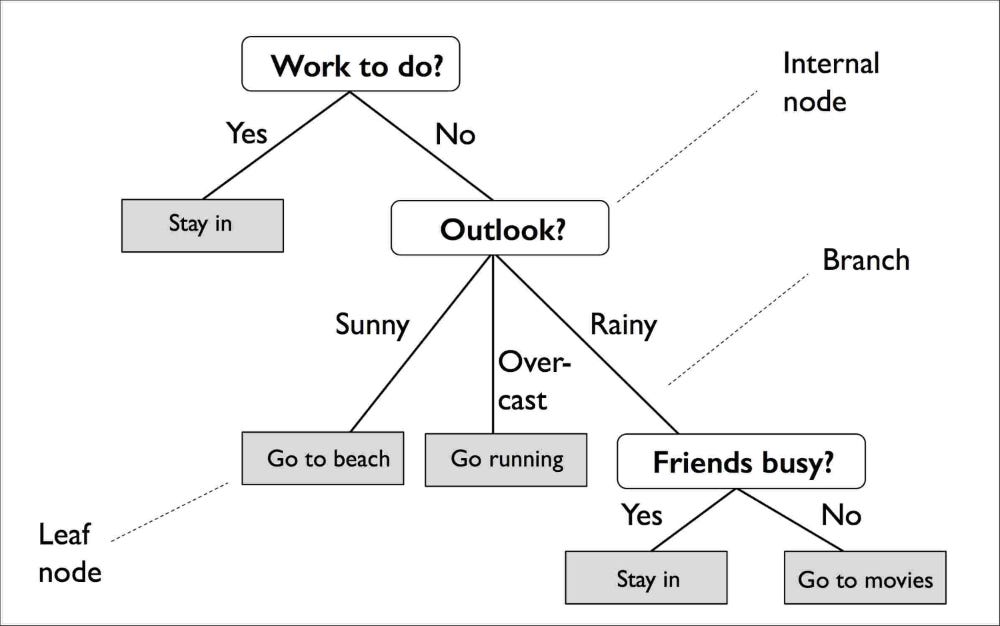
A Crop Recommendation System uses machine learning algorithms to provide recommendations to the farmers on the best crops to plant based on specific conditions. It considers various factors like amount of NPK (Nitrogen, Phosphorous, Potassium) present in the soil, temperature, and humidity in that area, ph of soil and the amount of rainfall received during that time. The system provides a connectivity to farmers via a mobile application. Farmers can use this system to improve their yield and thus improving their agricultural productivity and profitability.

In this project I have merged two crop recommendation datasets to increase the data and handle the problem of overfitting. The main objective/aim of this project is to use different ML Algorithms and check which algorithm(s) gives the best result. Some of the algorithms that I have used in this project for making model includes:

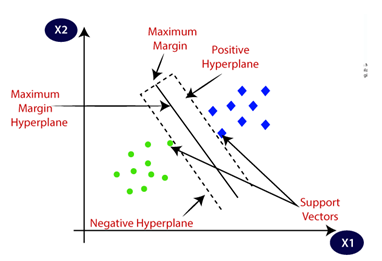
* **Random Forest: -**  It is a popular supervised ML algorithm which can be used for both regression and classification. It produces predictions based on the majority votes from several decision trees rather than relying just on one.



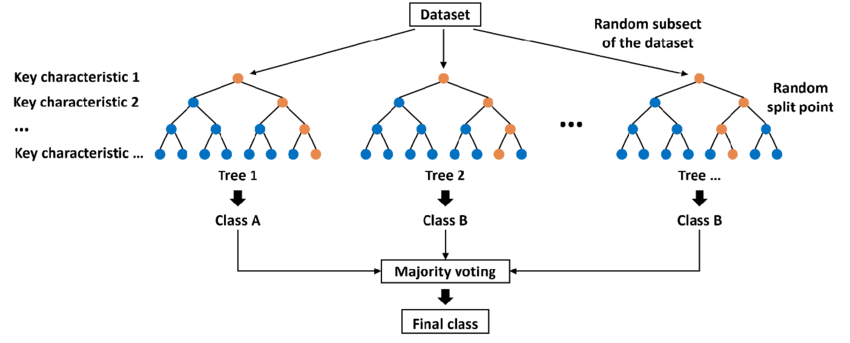
* **Decision Tree:-**  It is also a supervised ML algorithm that can be used for both regression and classification. In a decision tree the internal nodes represent the features, branches represent the decision rules and the leaf nodes represent the outcome. It simply asks a question and based on the answer(yes/no), it splits the tree into subtrees.



* **Support Vector Classification: -**  SVC is a supervised learning algorithm used for classification tasks. It aims to find the best hyperplane that separates different classes of data points in an n-dimensional space. The best hyperplane is the one that has the largest margin, which is the distance between the hyperplane and the closest data points of each class, also known as support vectors.



* **LightGBM:-** Light Gradient Boosting Machine is a gradient boosting framework that is particularly used for tasks like classification and regression. It is based on decision trees. Its aim is to increase the efficiency of the model and reduce the memory usage.
* **ExtraTrees Classifier:**- The extra trees classifier is a powerful tool particularly used for classification tasks. It randomly considers a bunch of potential splits at each decision point.



**Chapter 3**

**Methodology**

* **How the Model Works?**
* **Importing Modules: -**  The code starts by importing necessary Python modules and setting up the environment for data analysis and visualization.
* **Data Gathering and Displaying:** Two datasets (df1 and df2) are read using the Pandas library. Concatenation is then performed to combine these datasets into one. Duplicates are then removed from the combined dataset.
* **Exploratory Data Analysis (EDA):** The code performs EDA on the dataset. It includes checking the dataset’s shape. Columns, duplicated values, and basic statistics. Box plots are created to visualize the distribution of key features like nitrogen (N), phosphorous (P), potassium (K), temperature, humidity, pH, and rainfall.
* **Preprocessing:** Outliers in the 'rainfall' feature are identified and removed using the Interquartile Range (IQR) method. Correlation between different features is visualized using a heatmap.
* **Data Visualization:** Bar plots are created to visualize the distribution of nitrogen (N), phosphorous (P), and potassium (K) for each crop.
* **Model Training:** The dataset is split into training and testing sets. Several machine learning models are trained using the training set. The models used include LightGBM, Decision Tree, Random Forest, Support Vector Machine (SVM), and Extra Trees. Each model's accuracy is evaluated using the testing set.
* **Model Evaluation:** Confusion matrices and classification reports are generated to evaluate the performance of each machine learning model. The metrics include accuracy, precision, recall, and F1-score.
* **Model Testing**: The trained Random Forest model is tested with several sample inputs (test,test1, etc.) to make crop predictions based on input feature values.
* **Model Comparison:** The code compares the accuracy of different machine learning models (LightGBM, Decision Tree, Random Forest, SVM, and Extra Trees) using the testing set.
* **Model Saving:** The Support Vector Machine (SVM) model is saved using the pickle library (pickle.dump) for future use.
* **How the GUI works?**
* **Importing Libraries:** The code starts by importing necessary libraries like streamlit, joblib, numpy, pickle, and warnings.
* **Loading the Machine Learning Model:** The code loads a machine learning model(‘modelSVC.pkl’) using the ‘pickle’ library.
* **User Input:** The code uses Streamlit sliders to allow the users to input values for Nitrogen, Phosphorous, Potassium, Temperature, Humidity, pH, and Rainfall.
* **Making Predictions:** The user inputs are combined into a list ‘p’, which is then used to make predictions using the loaded machine learning model.
* **Displaying Results:** The recommended crop is displayed using ‘st.write’.
* **Providing Recommendations and Information:**  Based on the predicted crop, the code displays specific recommendations and information about the do’s and don’ts for cultivating that crop.

**Chapter 4**

**Code**

1. **Collecting and displaying the data**

**A screenshot of a computer

Description automatically generated**

1. **EDA**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

A screenshot of a computer screen

Description automatically generated

1. **Preprocessing The Data**

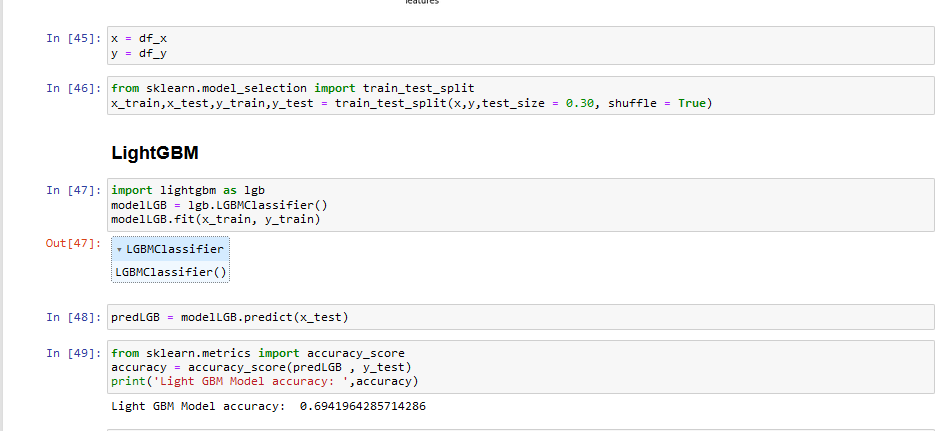
**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

1. **Building various Models**

****

**A screenshot of a computer program

Description automatically generated**

**A screenshot of a computer program

Description automatically generated**

1. **Comparing all the Models**

**A computer code with text

Description automatically generated**

1. **Saving the model with the best accuracy**

**A screenshot of a computer

Description automatically generated**

1. **Creating GUI using Streamlit**

A screen shot of a computer program

Description automatically generated

A screen shot of a computer code

Description automatically generated

A screen shot of a computer program

Description automatically generated

**Chapter 5**

**Need of Crop Recommendation System**

* **Increased yields:** By planting the right crops in the right conditions, farmers can increase their yields.
* **Improved profitability:** Farmers can make more money by planting crops that are in high demand and that are likely to sell for a good price.
* **Reduced risk:** Crop recommendation systems can help farmers to avoid planting crops that are not suited to the area, which can reduce the risk of crop failure.
* **Improved resource management:** Crop recommendation systems can help farmers to make the most of their land, water, and other resources.

**Chapter 6**

**Applications of Crop Recommendation System**

* **Enhanced Crop Selection :-**  The system analyses various factors and then recommends the most suitable crop. This data-driven approach helps the farmers to avoid unsuitable choices and maximize their yield.
* **Increased Profitability :-** By choosing the right crops and managing resources efficiently, farmers can potentially boost their income and improve their livelihoods.
* **Promoting Sustainable Agriculture :-** The system can be integrated with other models to guide farmers towards choosing crops and practices that are less taxing on the environment. This includes recommending crops with lower water requirements, encouraging soil conservation practices, and suggesting integrated pest management techniques.
* **Reduced Risk of Crop Failure :-** Early warnings about potential pest outbreaks, diseases, or unfavourable weather conditions can be integrated into CRS, allowing farmers to take preventative measures and minimize crop losses.

**Chapter 7**

**Conclusion**

This project has delved into the fertile ground of precision agriculture, exploring the potential of data-driven decision-making for optimal crop selection. We recognized the limitations of traditional methods, often reliant on intuition and local knowledge, that can fall short in the face of climatic uncertainties and varying soil conditions.

Our response was to cultivate a personalized crop recommendation system, empowered by analysis of five key environmental factors: Nitrogen, Phosphorus, Potassium (NPK) levels, temperature, pH, rainfall, and humidity. By feeding this rich data into five diverse algorithms, we embarked on a rigorous selection process, ultimately identifying the [mention the best performing algorithm by name] as the champion of accurate and reliable recommendations.