NITTE MEENAKSHI INSTITUTE OF TECHNOLOGY

COURSE LA2 PROPOSAL

On

Crop Recommendation Prediction

Submitted in partial fulfilment of the requirement for the award of

Degree of Bachelor of Engineering

in

Computer Science and Engineering

Submitted by:

Aabhash Manandhar	1NT19CS004
Chirag Jung Thapa	1NT19CS060
Prajesh Shrestha	1NT19CS139
Utsav Sapkota	1NT19CS206

Department of Computer Science and Engineering

2021-22

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the Course Research Paper titled "Crop Recommendation Prediction" is an authentic work carried out by Aabhash Manandhar(1NT19CS004), Chirag Jung Thapa(1NT19CS060),Prajesh Shrestha (1NT19CS139) and Utsav Sapkota(1NT19CS206), Bonafide students of Nitte Meenakshi Institute of Technology, Bangalore in partial fulfilment for the award of the degree of Bachelor of Engineering in COMPUTER SCIENCE AND ENGINEERING of Visvesvaraya Technological University, Belagavi during the academic year 2021-22.

Name and Signature of the Faculty In charge   Name and Signature of the HOD

ACKNOWLEDGEMENT

Table of Contents

ABSTRACT

In this project, we help the farmers to get informed decision about the farming strategy. The algorithms used for data mining are Random Forests, Decision Tree Algorithms. We train our model using the data set provided from Kaggle which in turn helps us build predictions that are close to the accuracy that the experts make. After making a prediction we compare the predicted answer with the actual value and adjust the changes.

INTRODUCTION

The goal of this project is to allow a farmer to get an idea of which crop to plant during which season of the month.

Machine learning is an important decision support tool for crop yield prediction, including supporting decisions on what crops to grow and what to do during the growing season of the crops.

Machine learning is everywhere throughout the whole growing and harvesting cycle. It begins with a crop prediction — from the soil preparation, seeds breeding and water feed measurement.

The model that we are preparing helps us to determine the right crop corresponding to the different attributes from the dataset.

DATASET

We chose the dataset from Kaggle and it has the following column:

- N - ratio of Nitrogen content in soil
- P - ratio of Phosphorous content in soil
- K - ratio of Potassium content in soil
- temperature - temperature in degree Celsius
- humidity - relative humidity in %
- ph - ph value of the soil
- rainfall - rainfall in mm

This dataset was build by augmenting datasets of rainfall, climate and fertilizer data available for India. The dataset is suited for data mining.

METHODS AND MODELS

We are trying to find out the best model by comparing their output and their performance in both the train and test instances. We will use various variations of decision tree models to find out the optimal algorithm.

We use decision tree algorithm to predict various crops that are suitable for growing based on the conditional attributes. The conditional attributes are humidity, temperature, rainfall, etc.

One of the next algorithm we will be using is the Random Forest Algorithm. Random Forest is a mainstream AI algorithm that has a place with the regulated learning strategy. It might be used for both Classification and Regression issues in ML. It depends on the idea of ensemble

learning, which is a cycle of joining numerous classifiers to tackle an intricate issue and to improve the presentation of the model.

Random Forest is a method of combining multiple Random Trees (thus - Forest) into one big classifier using even more randomization (selection of random samples with replacement for training each tree plus random selection of features which tree can use to perform split).

PRESENTATION AND VISUALIZATION

Data can be represented using graphs and other statistical methods. The agricultural data can be represented in charts, maps, and graphs to help one better understand the data. It can be done with the help of various data visualization tools.

With the visualization of Data we can understand the data in a better way. Technology and mobile devices that people interact with daily create tons of information. That's a lot of complex data to analyze and track. But, with automated tools, machine learning and AI-powered data integration, you can organize all your data and make sense of it. That's where data visualization comes in.

ROLES

Team Member	Task
Aabhash Manandhar	Decision Tree Classifiers
Chirag Jung Thapa	Decision Tree Classifiers
Prajesh Shrestha	Random Forest
Utsav Sapkota	Random Forest

SCHEDULE

Date	Task to be Completed
29/12/2021	Choose dataset and write proposal
10/01/2022	Data pre- processing to be completed
13/01/2022	Implement algorithms
15/01/2022	Completion of Project and the Report
17/01/2022	Presentation

REFERENCES

Project Proposal: https://www.fool.com/the-blueprint/project-proposal/

Random Forest: https://www.upgrad.com/blog/introduction-random-forest-algorithm/

Dataset: https://www.kaggle.com/atharvaingle/crop-recommendation-dataset

Decision Tree Classifiers: https://towardsdatascience.com/decision-tree-classifier-explainedin-real-life-picking-a-vacation-destination-6226b2b60575

Visualization: https://www.jigsawacademy.com/blogs/data-science/what-is-datavisualization/

https://www.geeksforgeeks.org/visualization-and-prediction-of-crop-production-data-using-python/