Hadoop Assignment 2

Name Tanzeel-ur Rehman

Student Number 648882

Github <https://github.com/htanzeel01/Hadoop->

A screenshot of a computer

Description automatically generated with low confidence

1. Pre-Process Necessary features

A screenshot of a computer program

Description automatically generated with low confidence

Loading the data set. Splitting the data and applying min Max scaler to preprocess the data

A screenshot of a computer code

Description automatically generated with low confidence

Also part of preprocessing the data.

**Predict the cultivator (3 options) of each wine measured**

I used KNN decision tree classifier and random forest classifier. I used these algorithms because they are classification algorithms and would be better to measure the prediction of the cultivators.

**Find top 3 predictive features using 3 different methods of measuring predictiveness.**

**A screenshot of a computer code

Description automatically generated with low confidence**

**A screenshot of a computer program

Description automatically generated with medium confidence**

**A screen shot of a computer code

Description automatically generated with low confidence**

**Report score/accuracy in at least 2 different formats**

**A screenshot of a computer code

Description automatically generated with medium confidence**