

Name: Darshan Rajopadhye

Roll No.: 17CO045

Subject: Machine Learning

Class: BE Computer 1

Assignment 3:

Problem Statement:

Search a COVID related dataset

1. Download the dataset
2. Perform Pre-processing on the dataset
3. Use this dataset to build a Naïve Bayes Classifier
4. Use this dataset to build a Decision Tree Classifier
5. Apply Ensemble to improve the accuracy.

Tool used:

Jupyter Notebook(Python).

Dataset:

The dataset consists of the symptoms of over 2 lakh people from all over the world classified as positive or negative.

Implementation:

Below attached is the converted notebook file with output.

Conclusion:

Test data set performance of both the models are:

1. Naive Bayes: 83.0%
2. Decision Tree:83.3%
3. Gradient Boosting: 86.27%
4. Random Forest:83.58%