Avnish Singh jaswal  
00113207218

CSE- 1

MACHINE LEARNING

LAB PROGRAM 2

**EXPERIMENT-2**

**Problem Statement**

Estimate the accuracy of decision classifier on breast cancer dataset

**Algorithm**

**Step-1:** Begin the tree with the root node, says S, which contains the complete dataset.

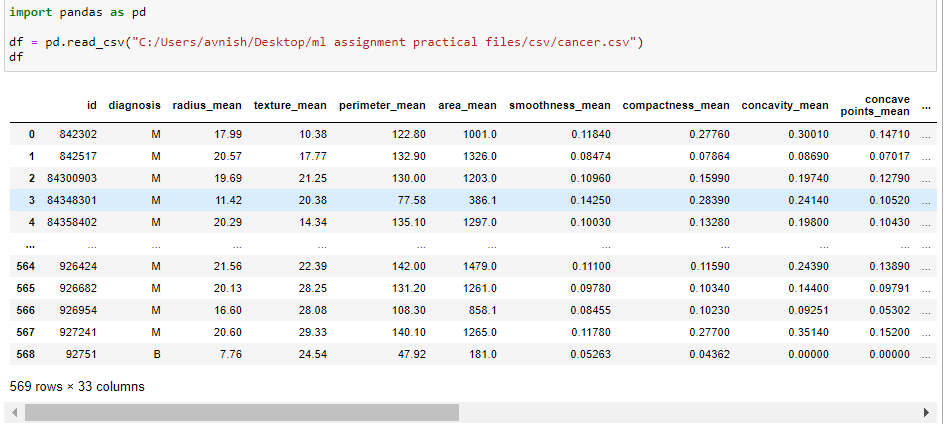
**Step-2:** Find the best attribute in the dataset using Attribute Selection Measure (ASM).

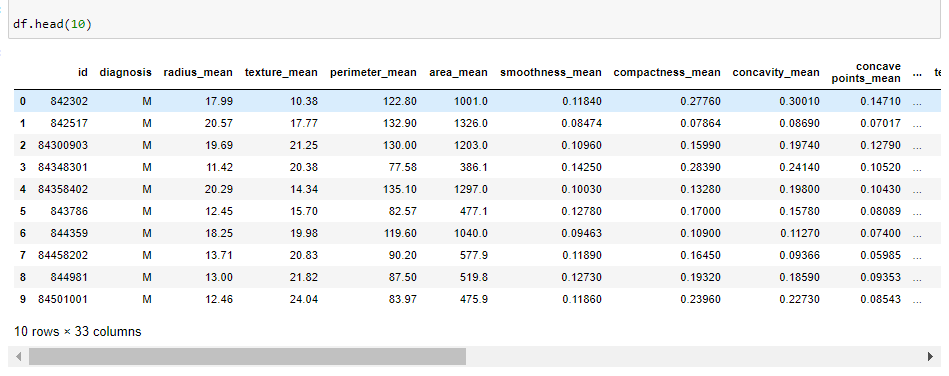
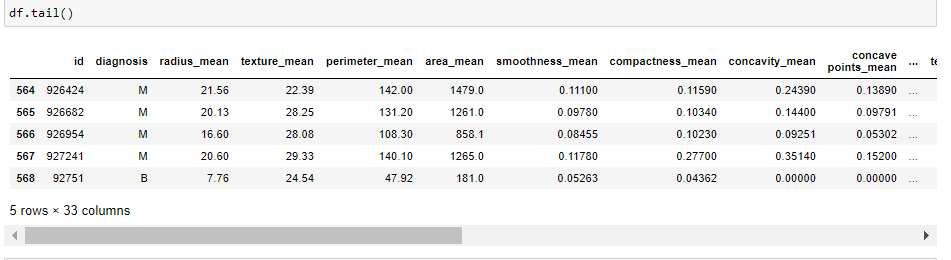
**Step-3**: Divide the S into subsets that contains possible values for the best attributes.

**Step-4:** Generate the decision tree node, which contains the best attribute**.**

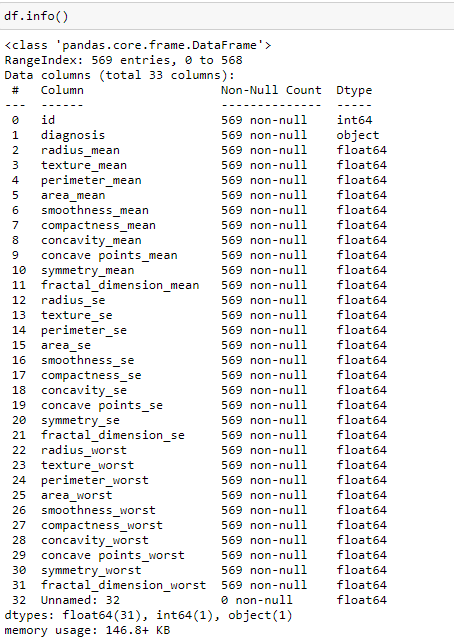
**Step-5:** Recursively make new decision trees using the subsets of the dataset created in step -3.Continue this process until a stage is reached where you cannot further classify the nodes and called the final node as a leaf node.

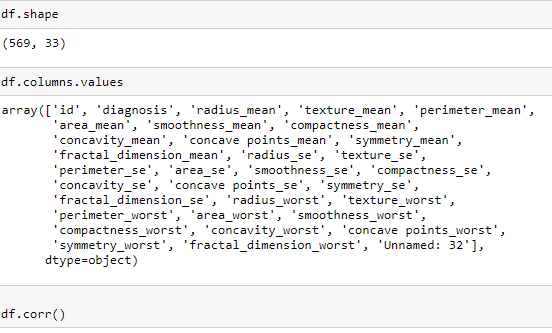
**Program Code Snippet**

**Loading Dataset** 

**Info About the Data**



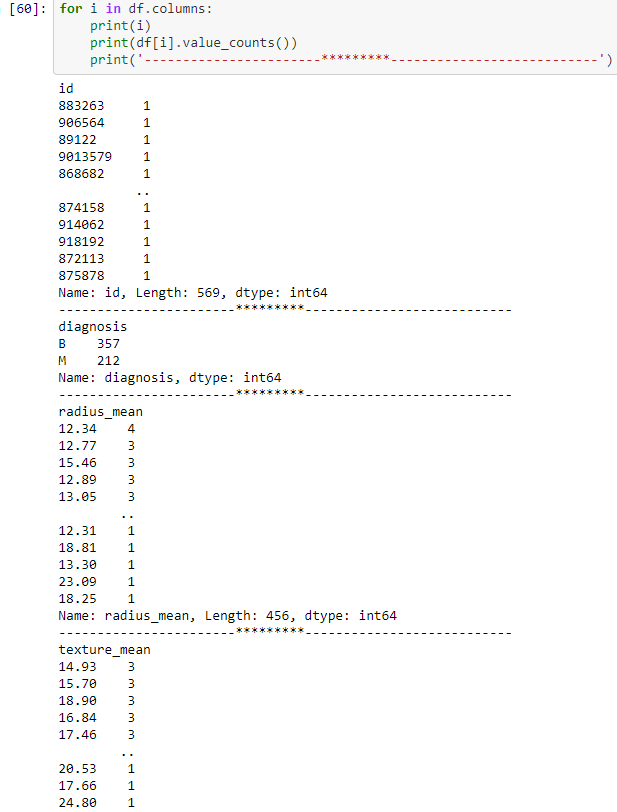


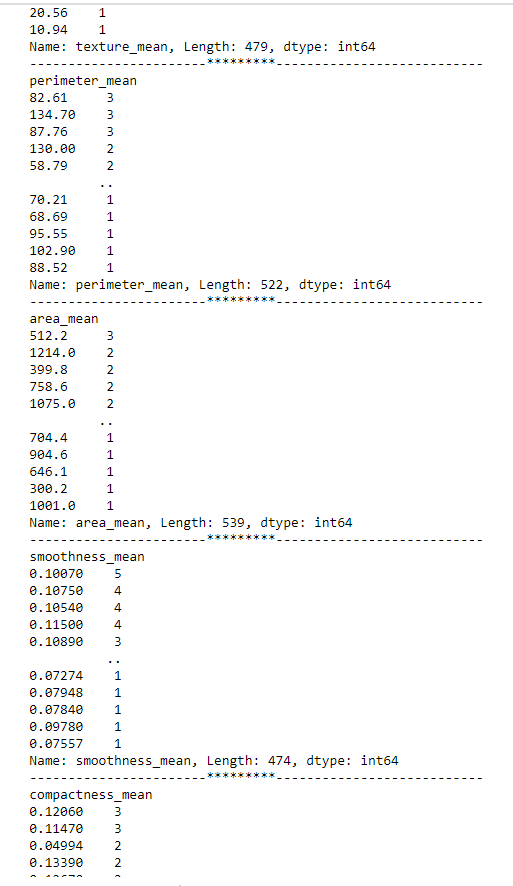
**Preprocessing/Cleaning of dataset.**

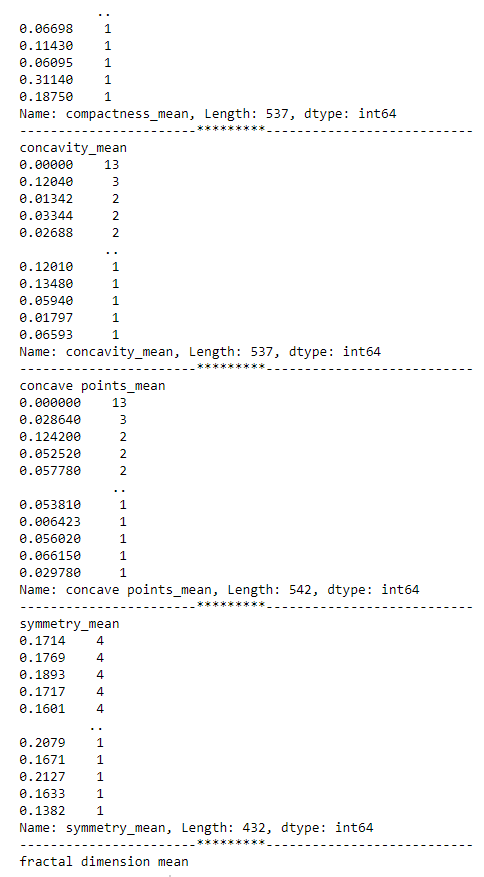


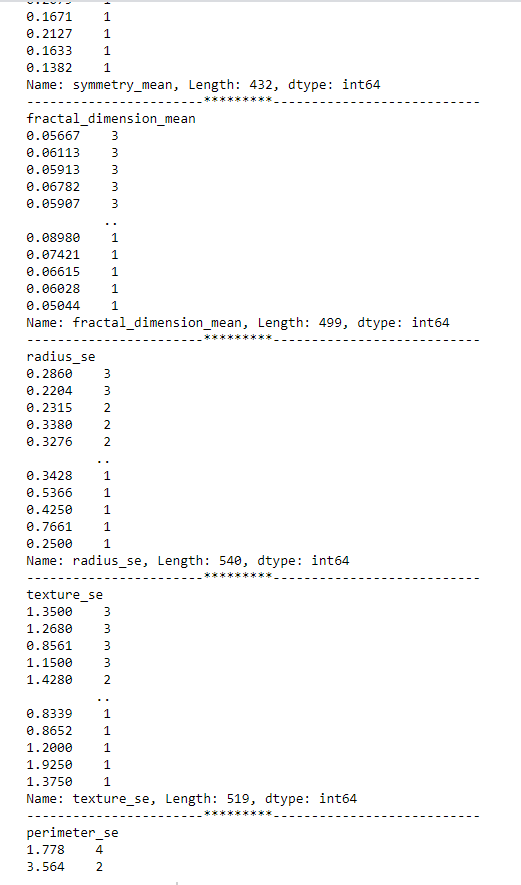


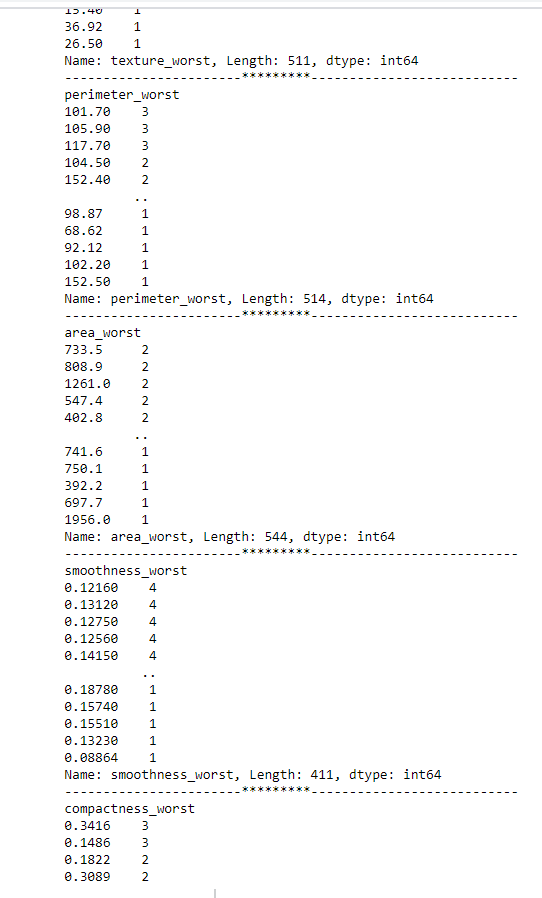
**Data Cleaning**

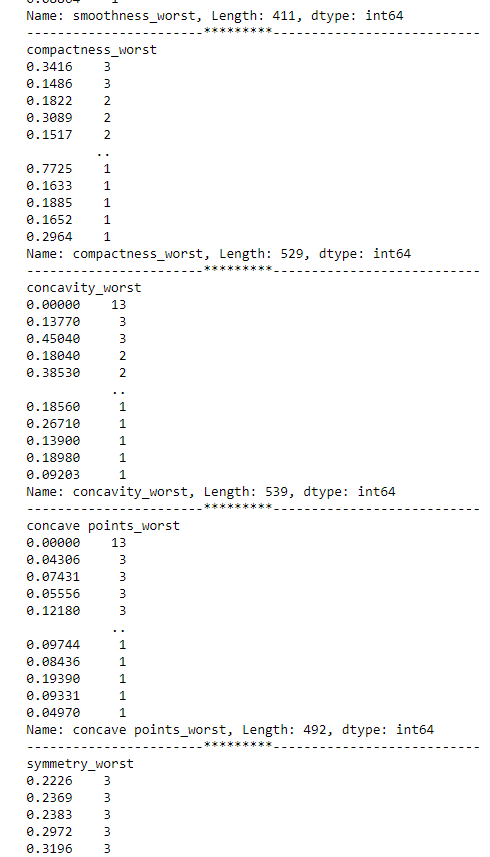


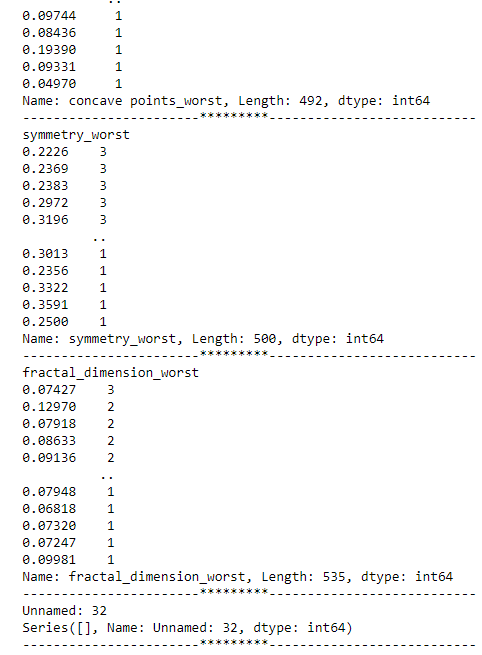


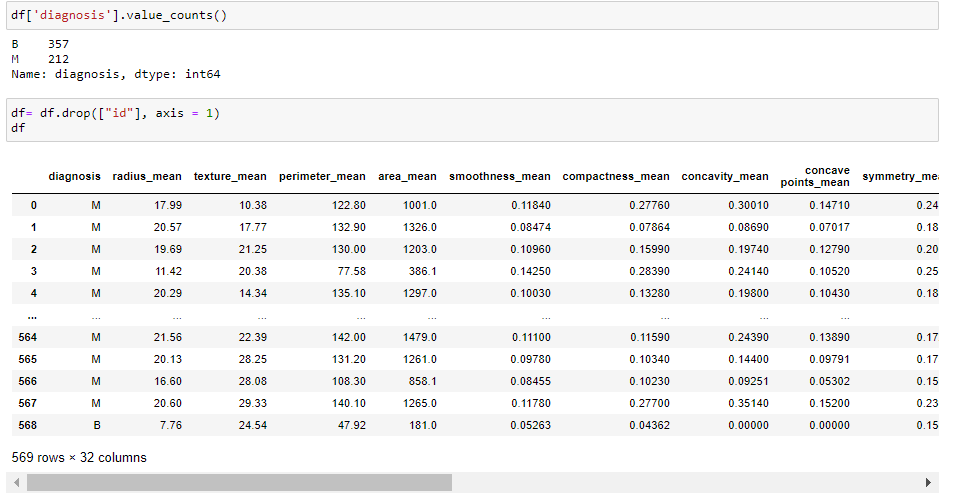




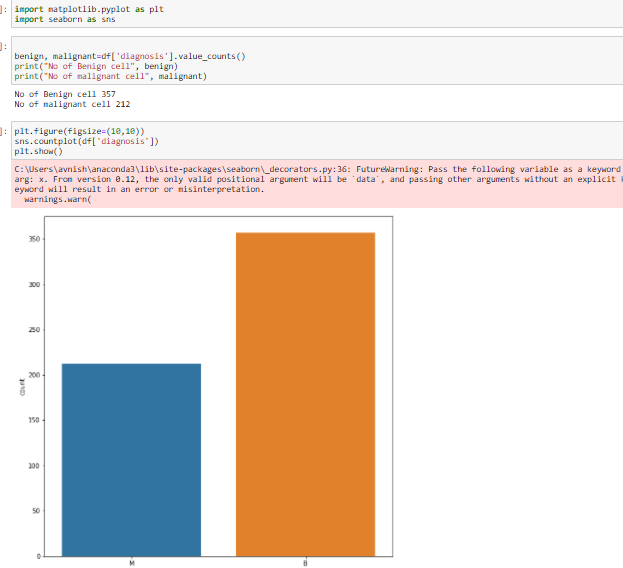




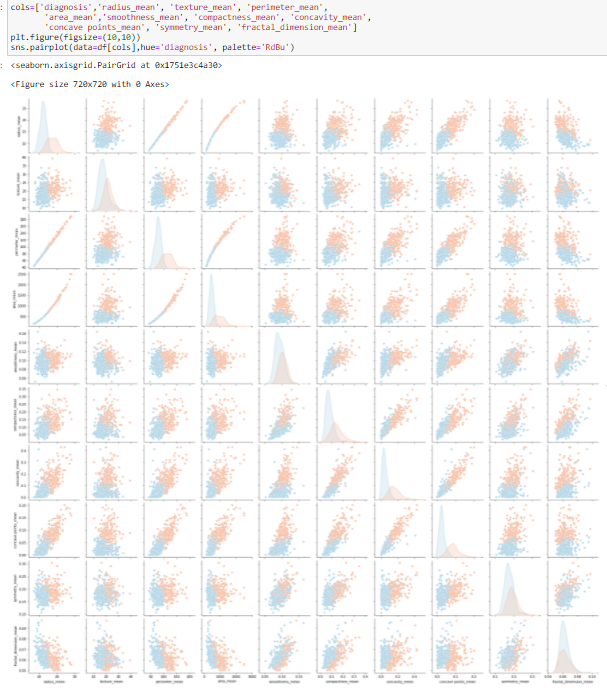


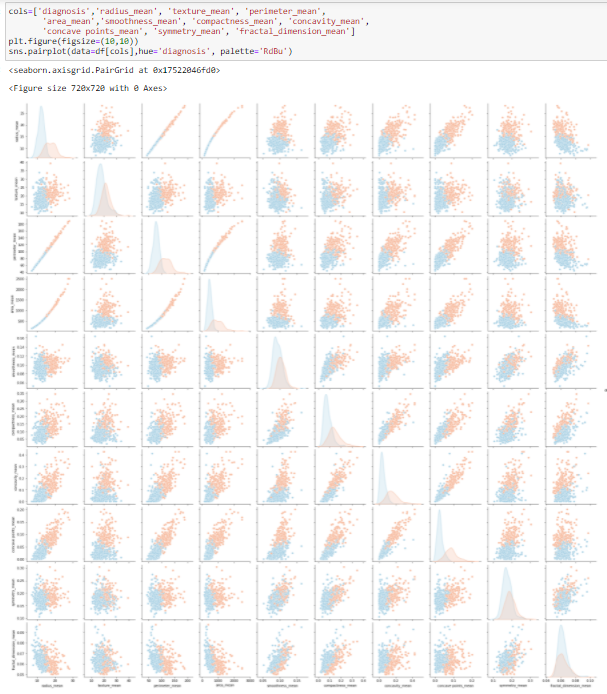


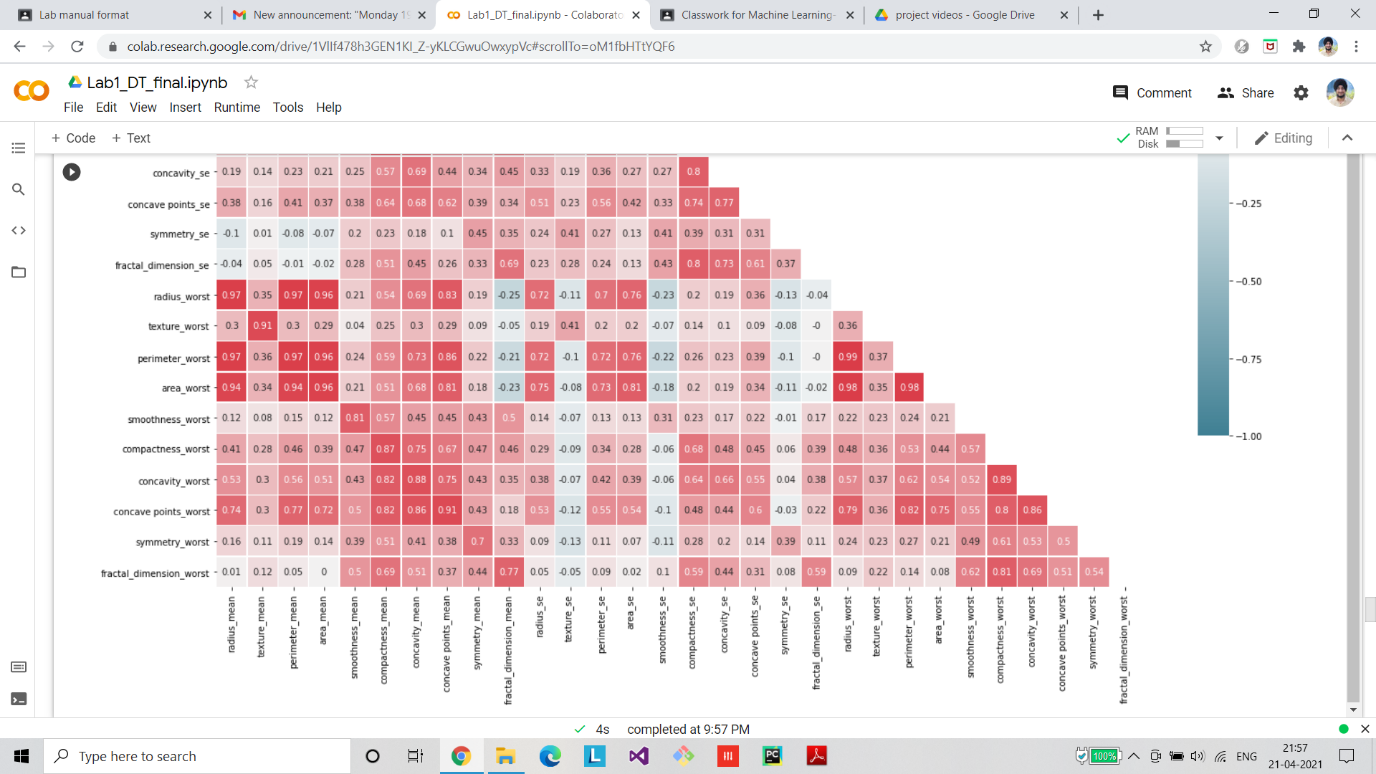
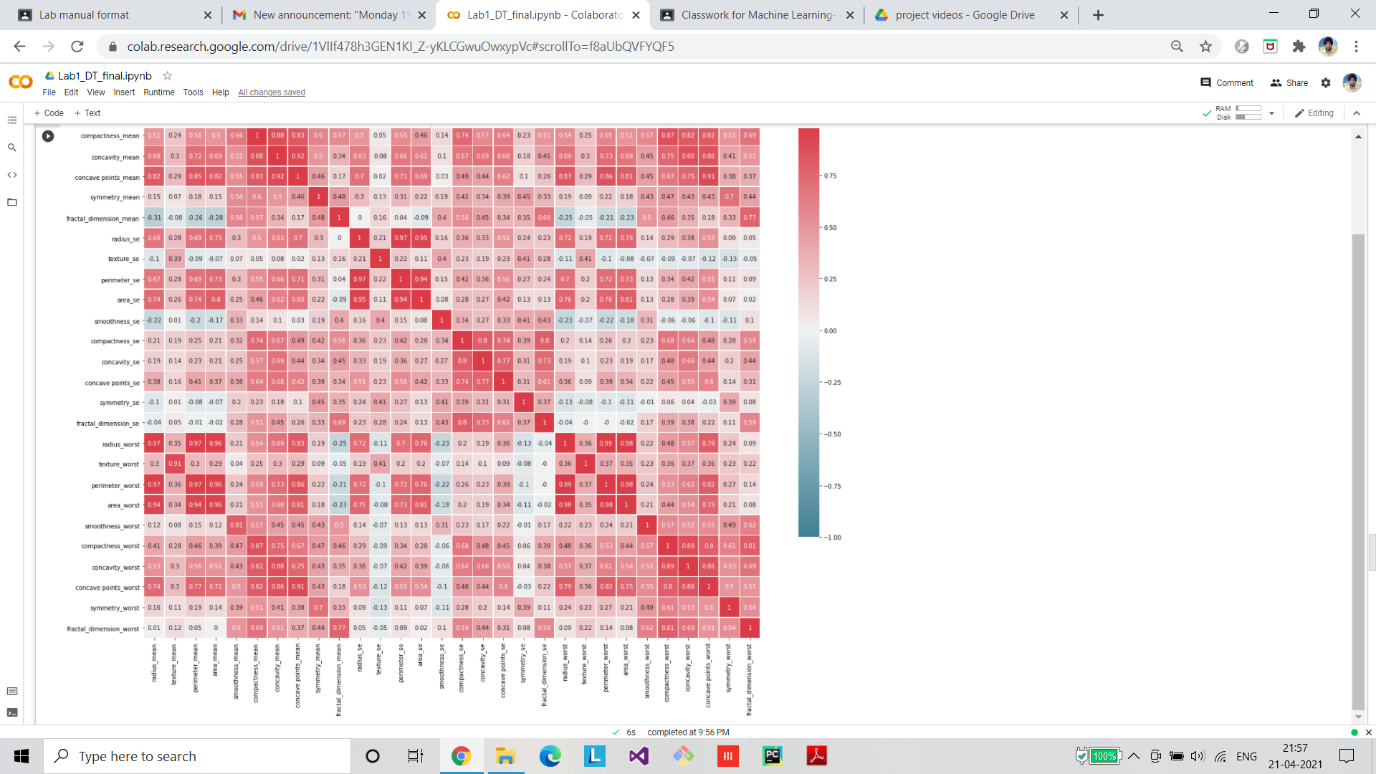
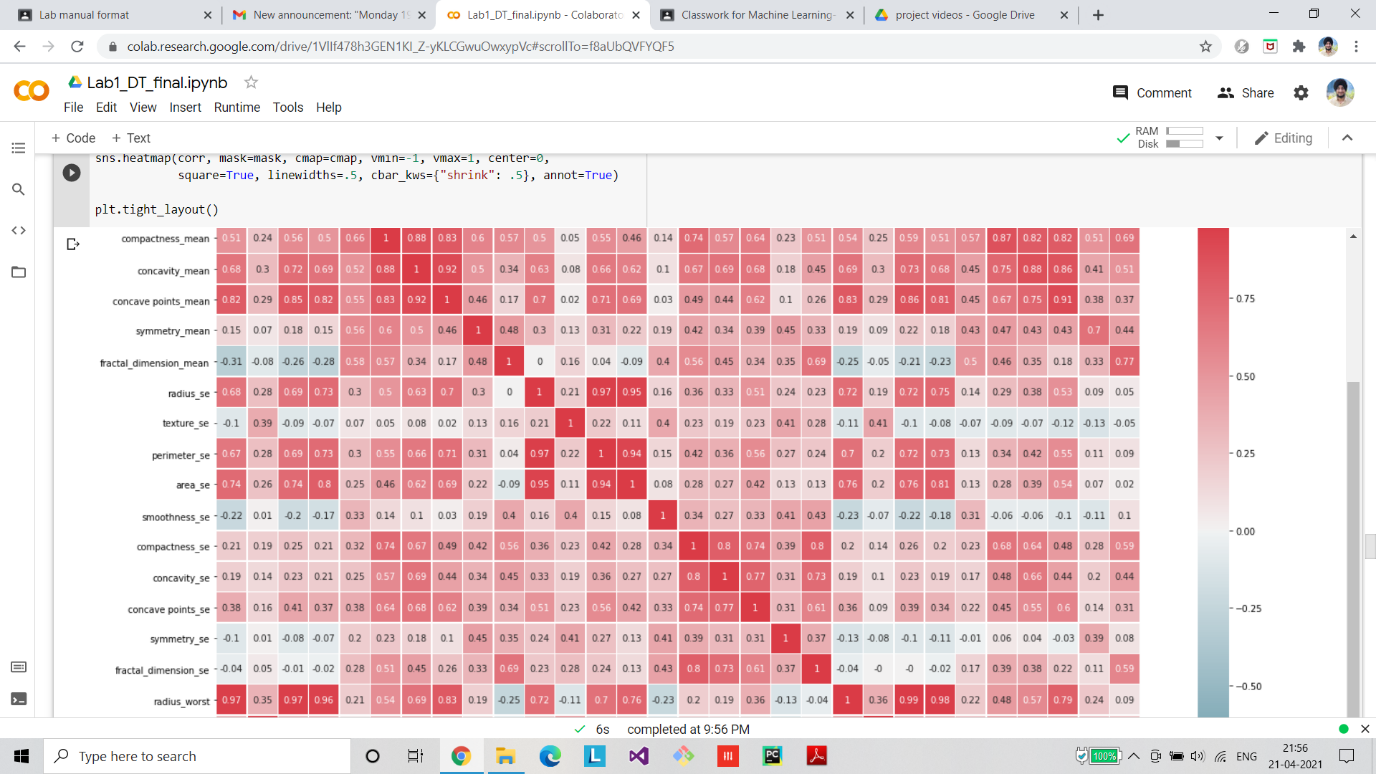
**Visualization**



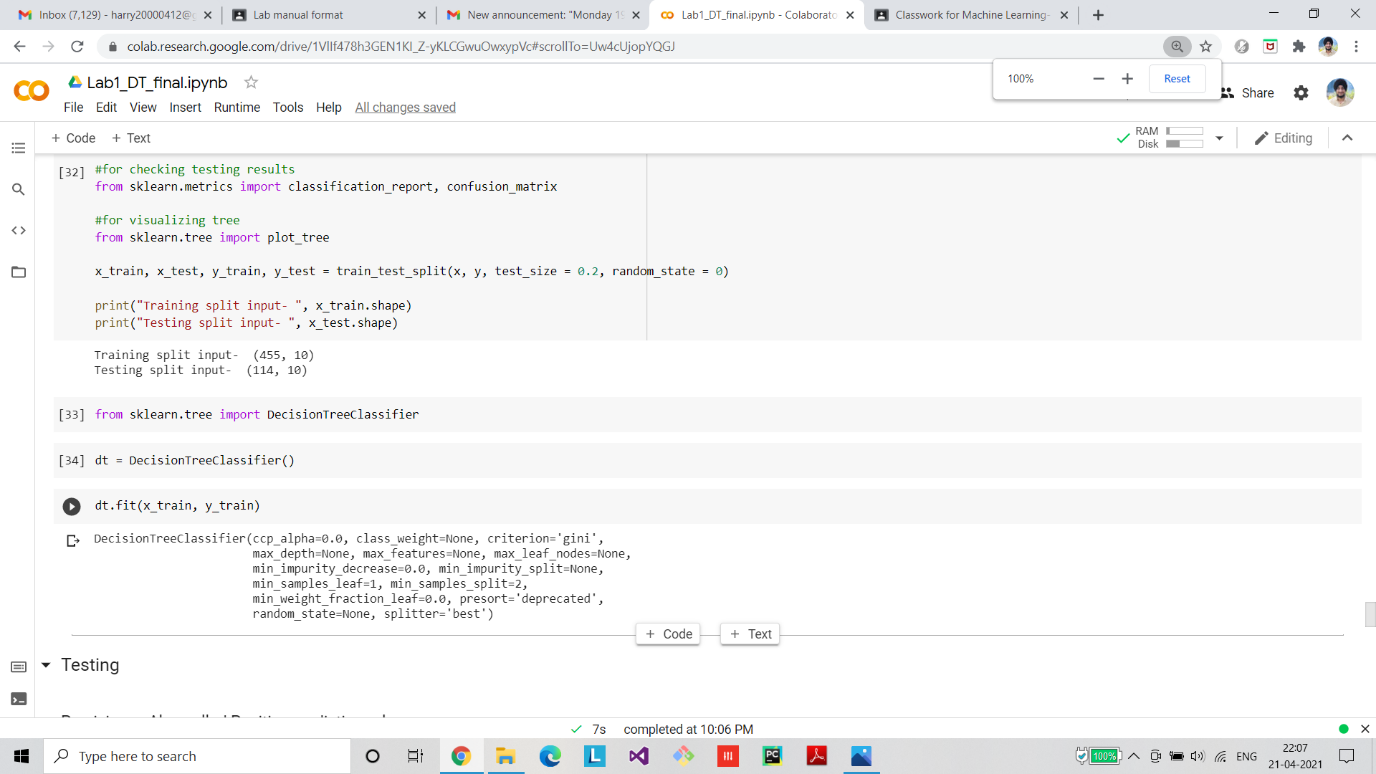
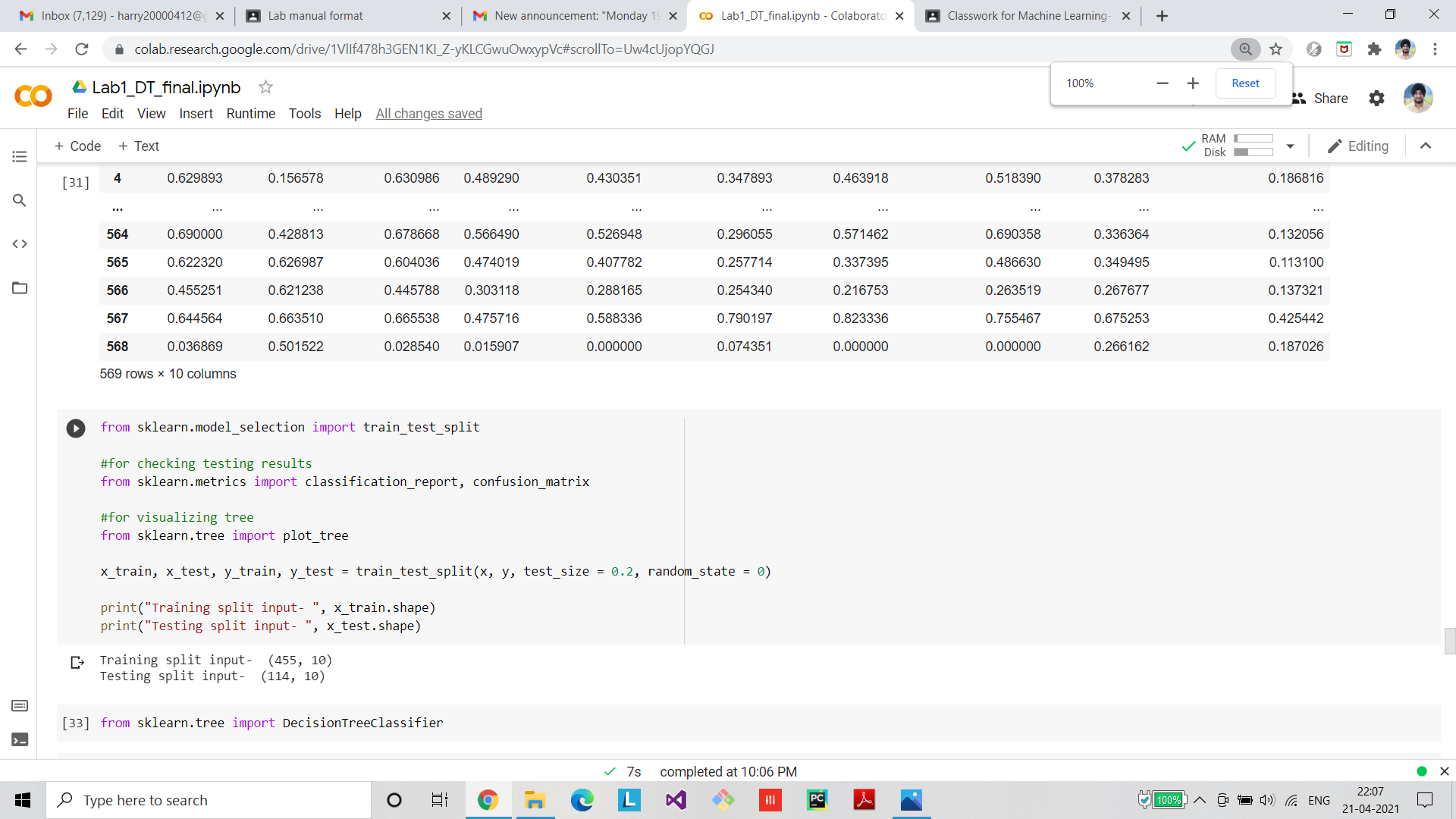
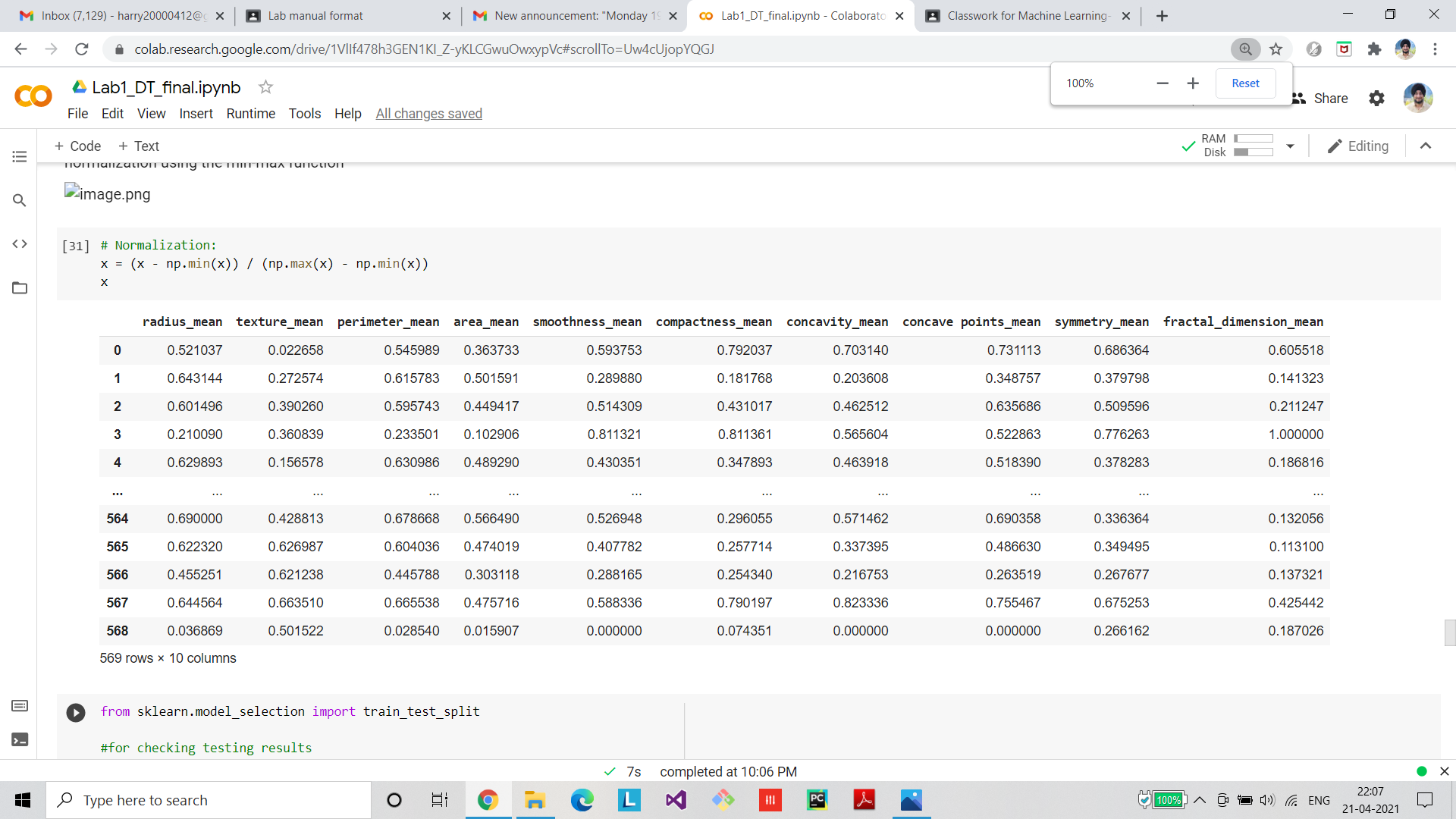
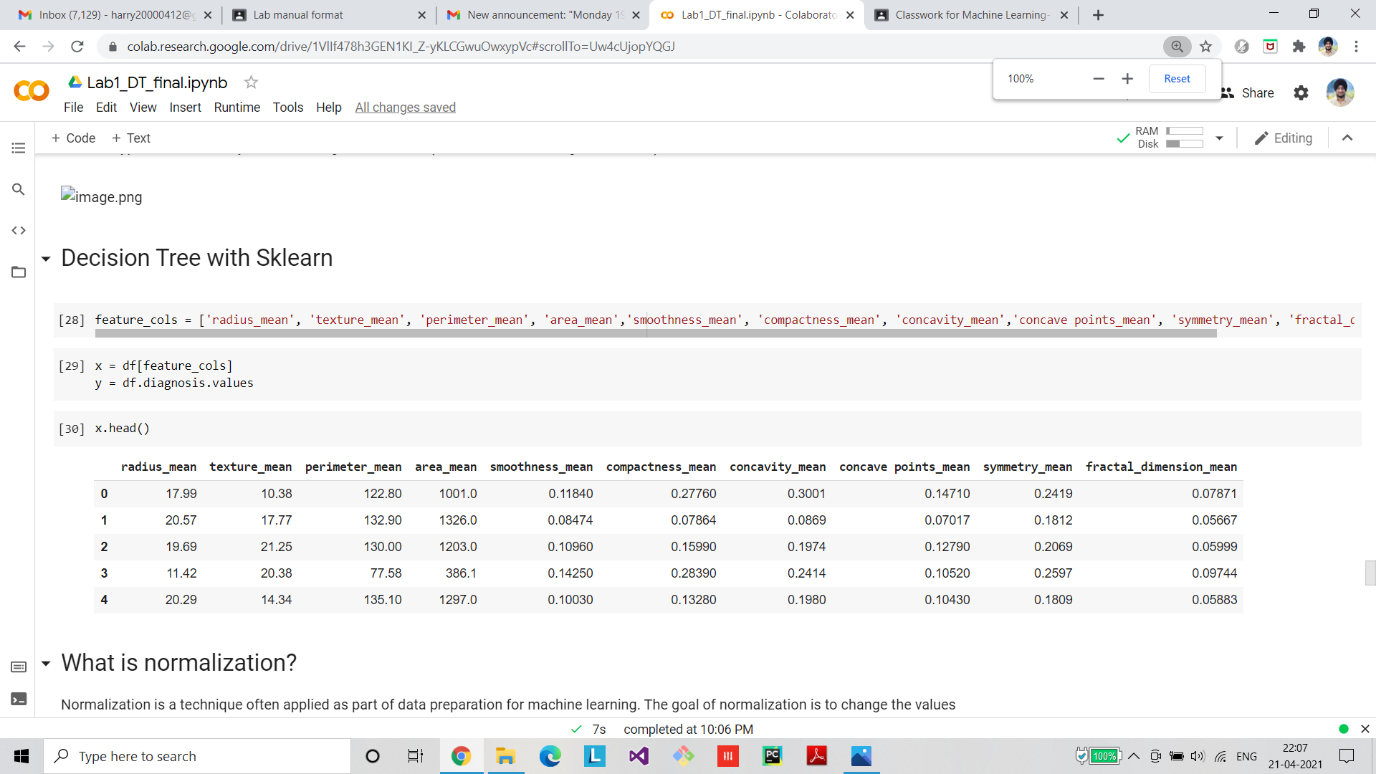




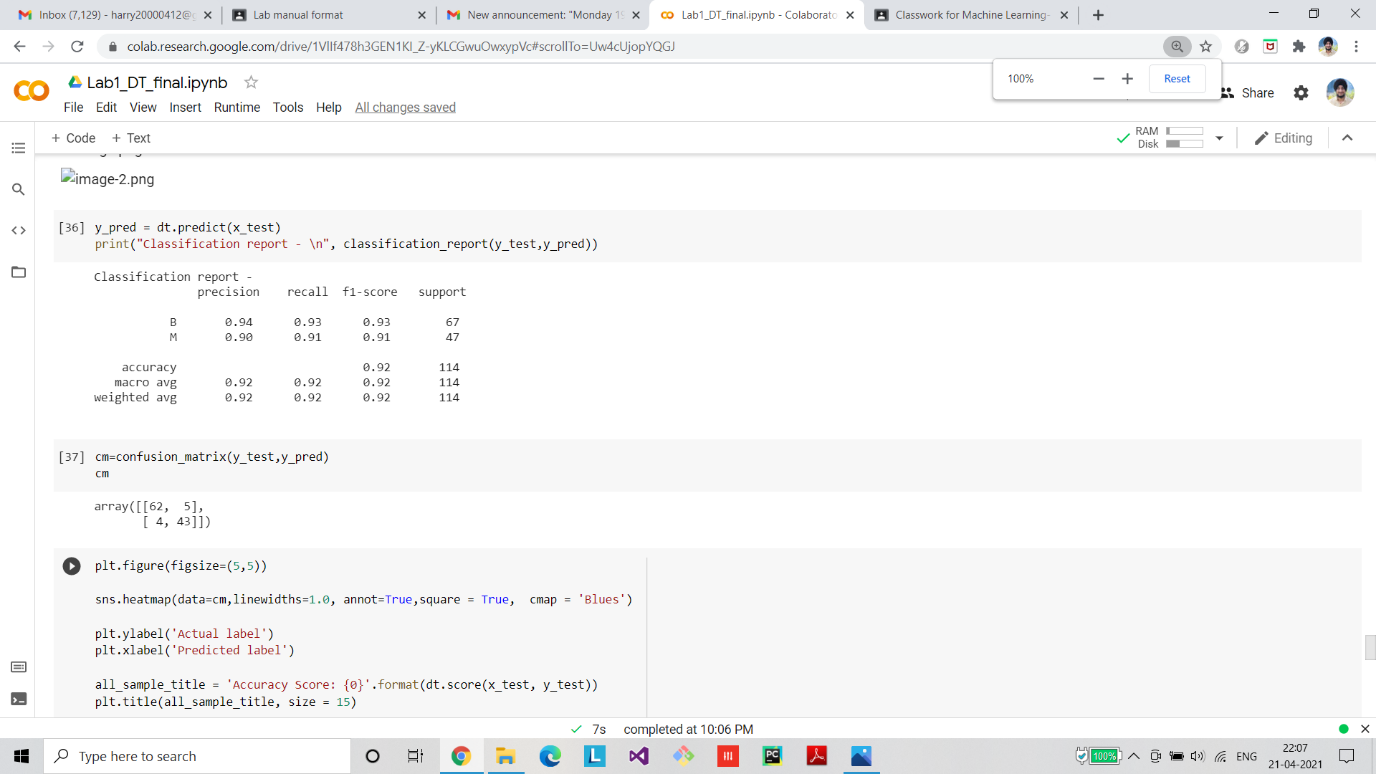


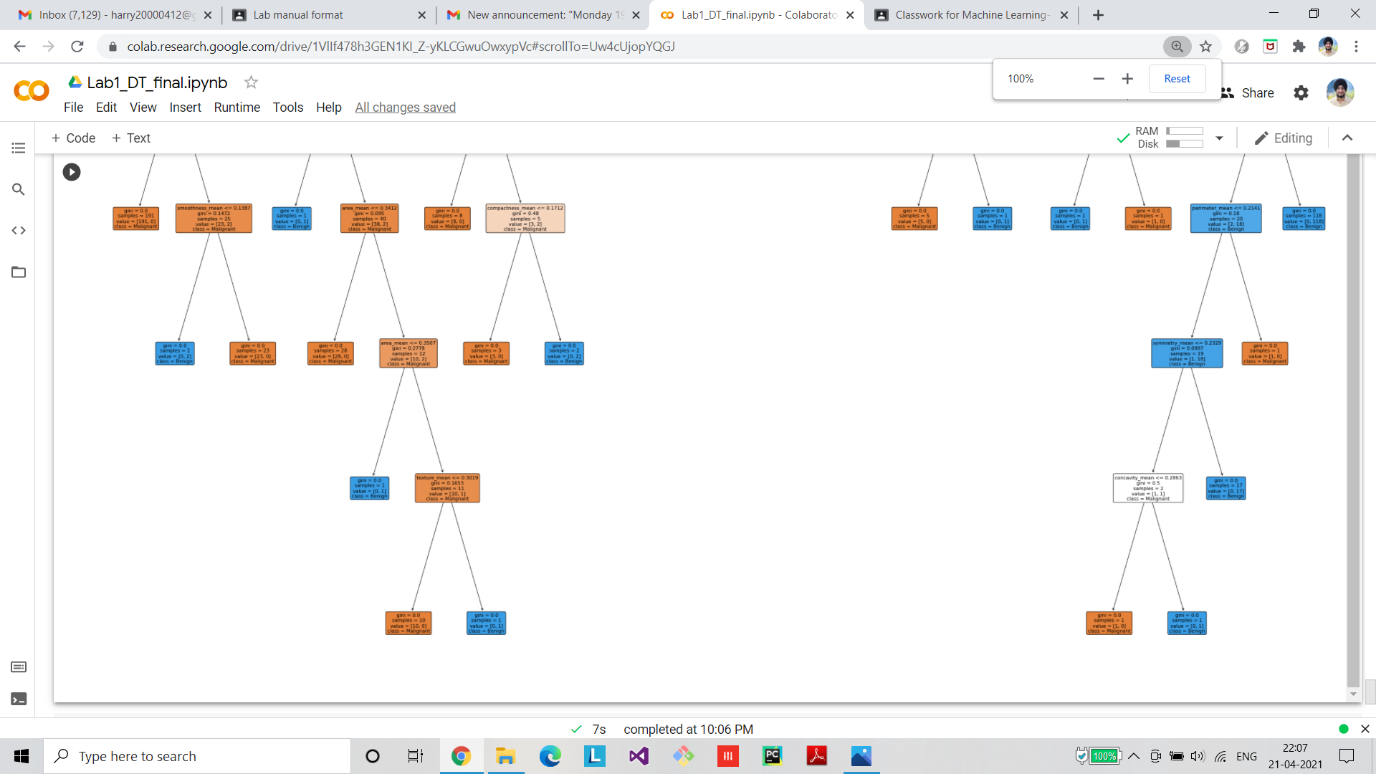
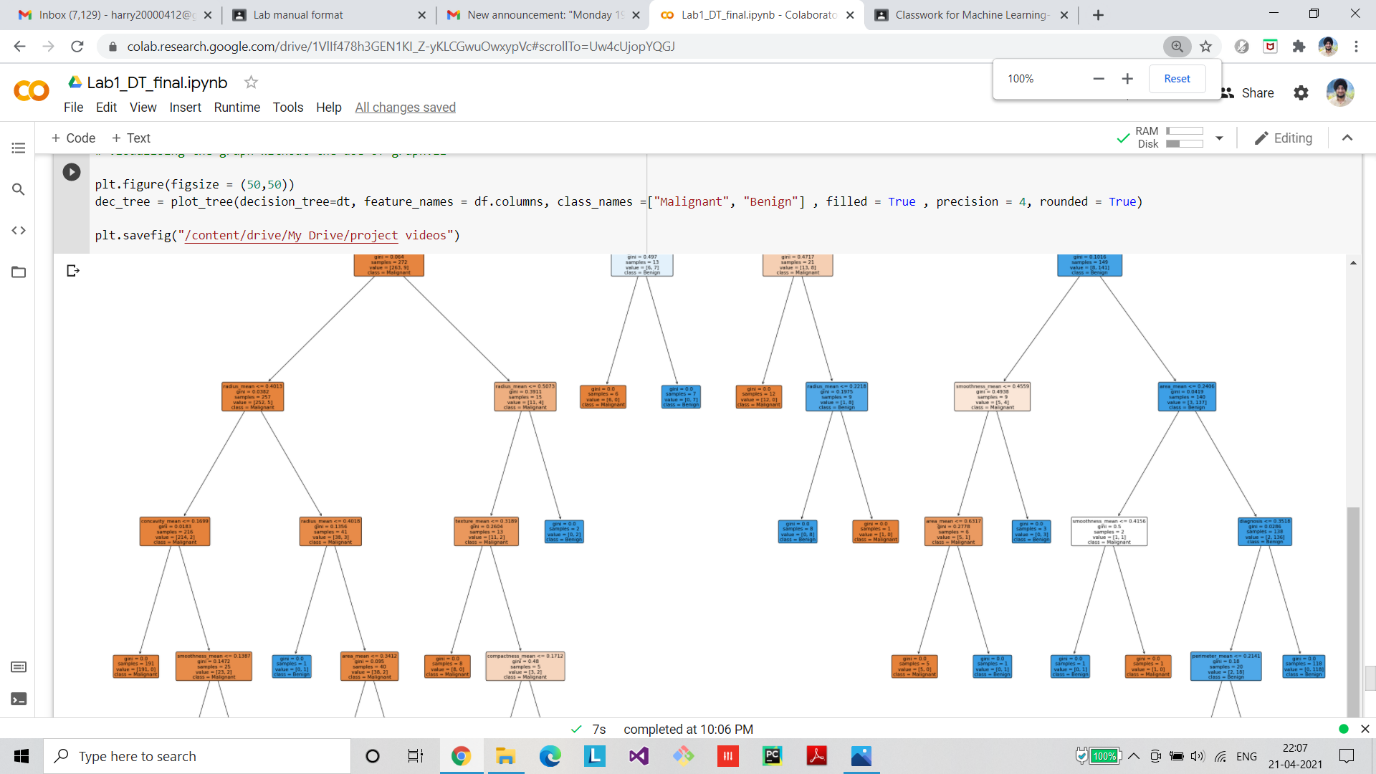
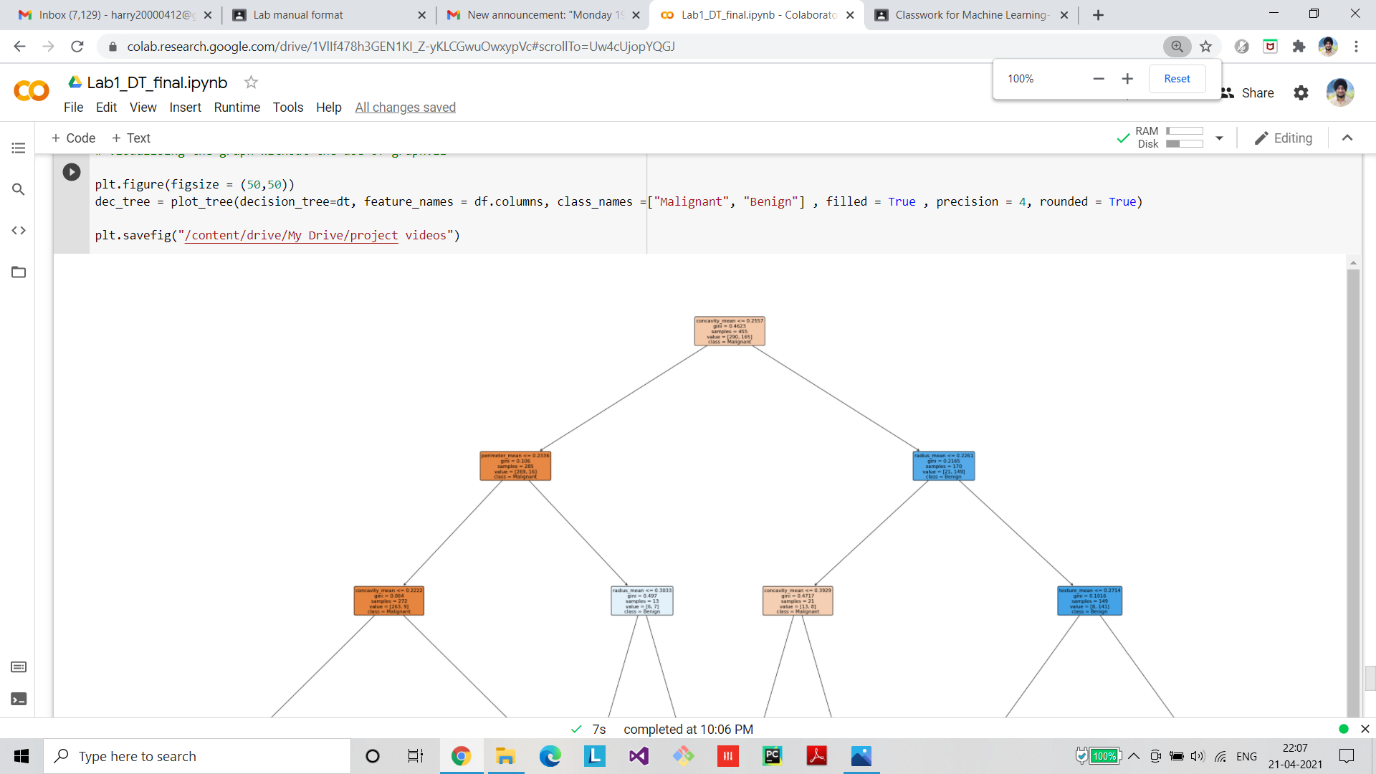
****

**Algorithm Implementation**

****

**Final Graph/ROC/Confusion Matrix**

**** 

****

**Github Link**

https://github.com/avnish9898/Ml-Experiment/blob/main/exp2.ipynb