Avnish Singh jaswal  
00113207218

CSE- 1

MACHINE LEARNING

LAB PROGRAM

**EXPERIMENT-3**

**Problem Statement**

Estimate the accuracy of decision tree classifier on breast cancer dataset using 5 fold cross validation

**Algorithm**

In a decision tree, for predicting the class of the given dataset, the algorithm starts from the root node of the tree.

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

This algorithm compares the values of root attribute with the record (real dataset) attribute and, based on the comparison, follows the branch and jumps to the next node.

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

For the next node, the algorithm again compares the attribute value with the other sub-nodes and move further.

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

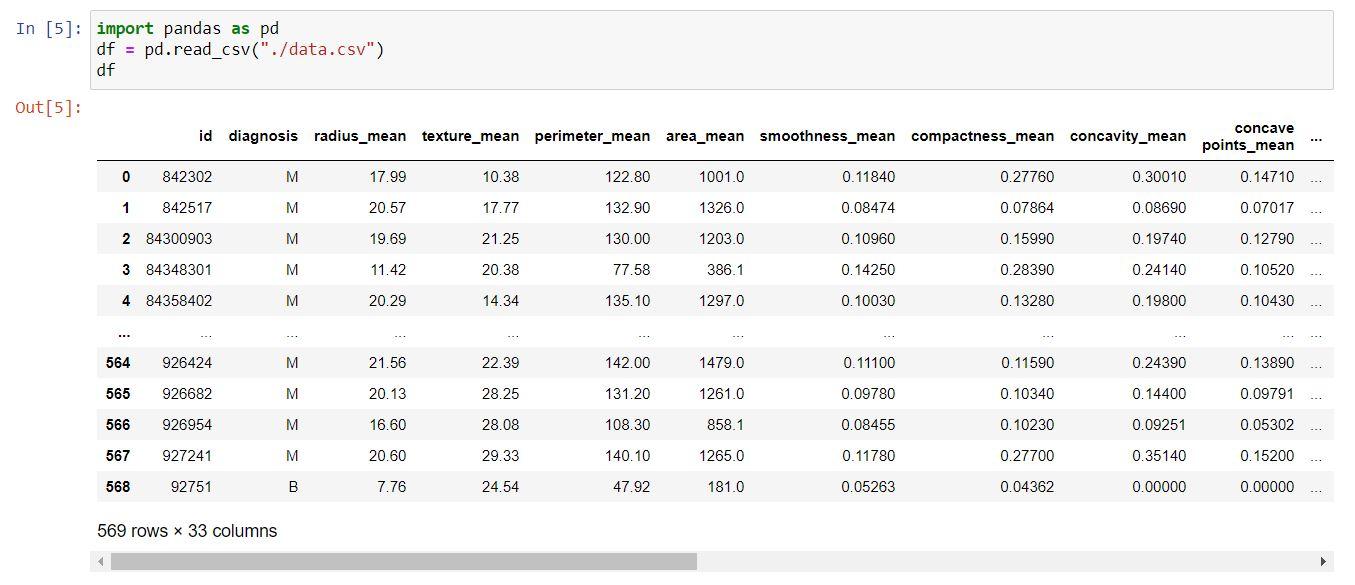
**Step-4:** Generate the decision tree node, which contains the best attribute.

It continues the process until it reaches the leaf node of the tree.

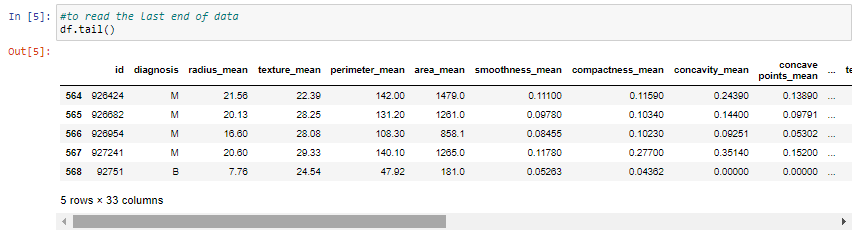
**Step-5:** Recursively make new decision trees using the subsets of the dataset created in step -3. Continue this process until a stage is reached where it cannot further classify the nodes and called the final node as a leaf node.

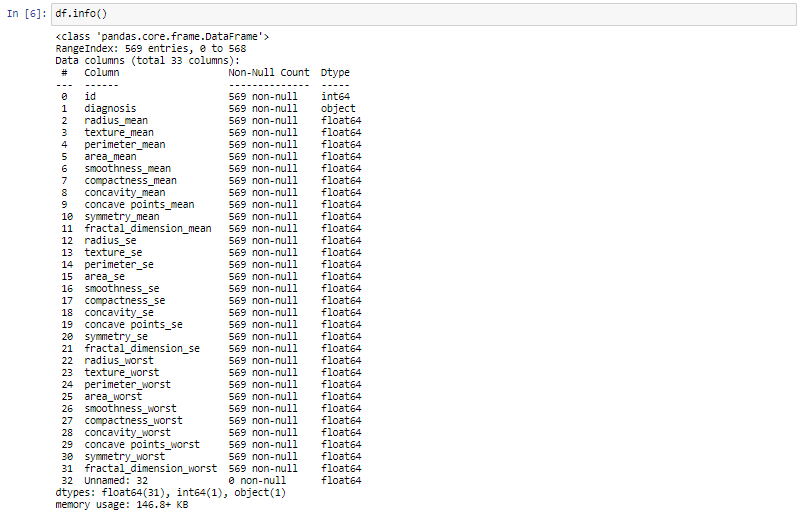
**Program Code Snippet**

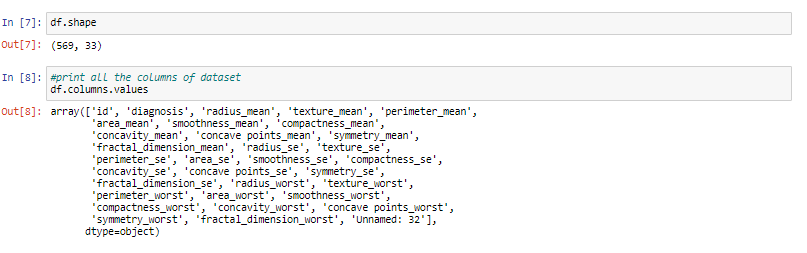
## LOADING DATA SET:

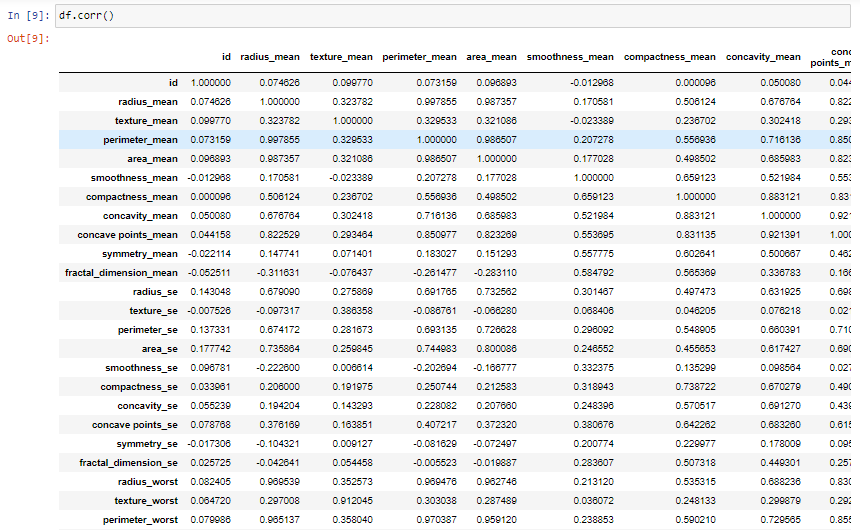


## PREPROCESSING:

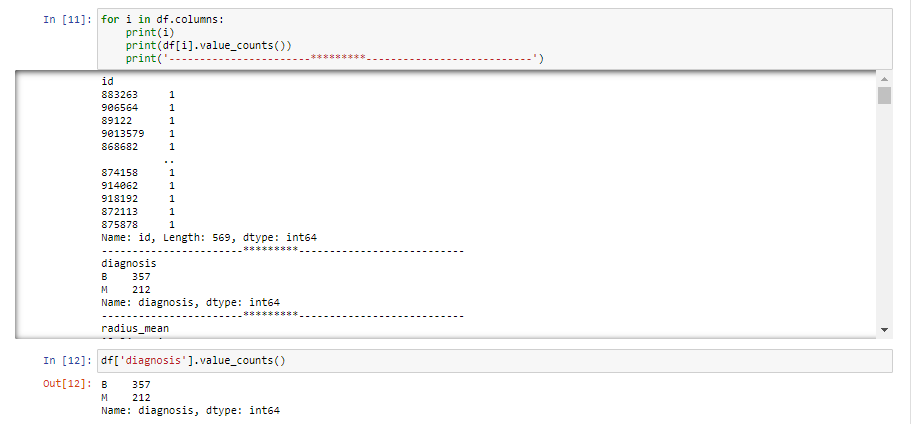


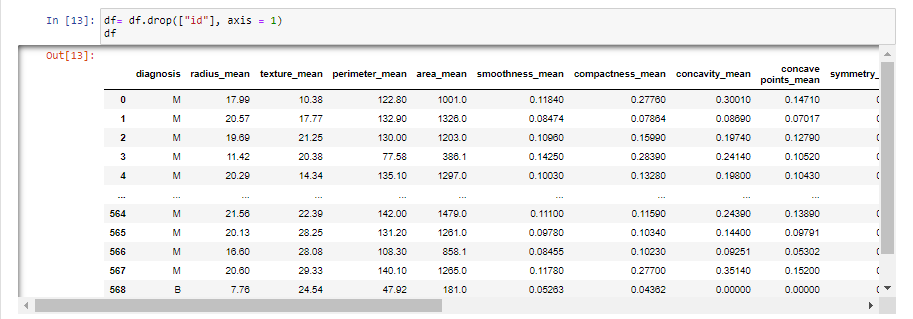


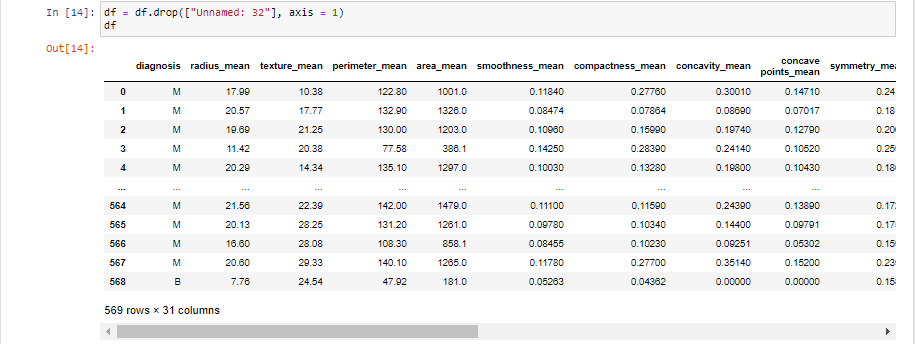






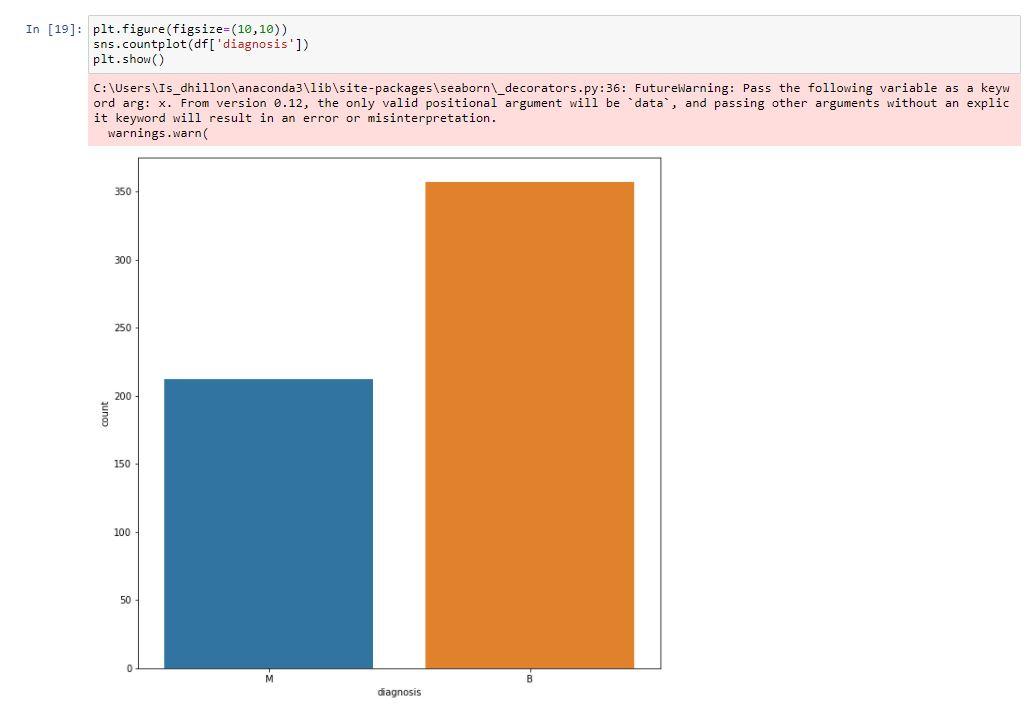


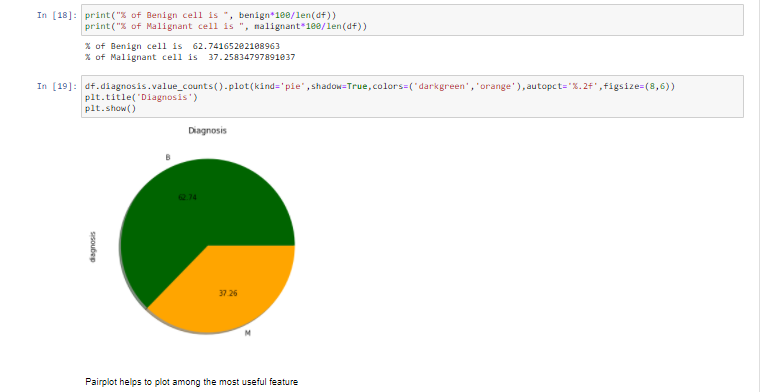


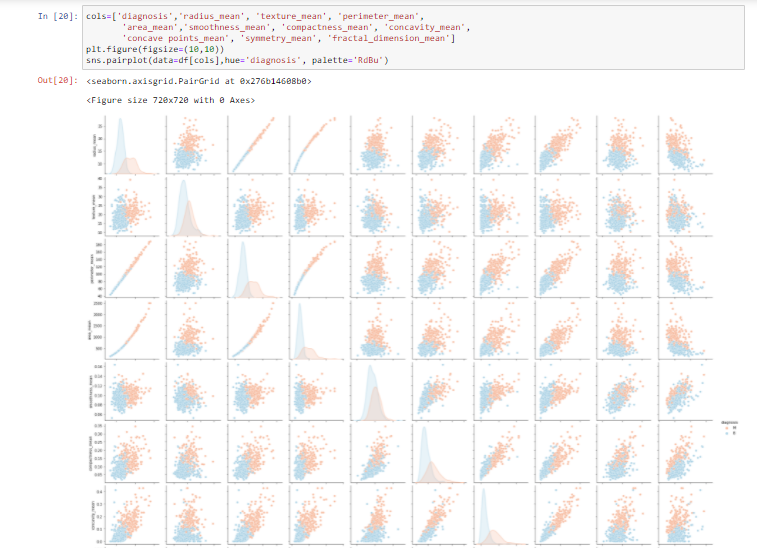


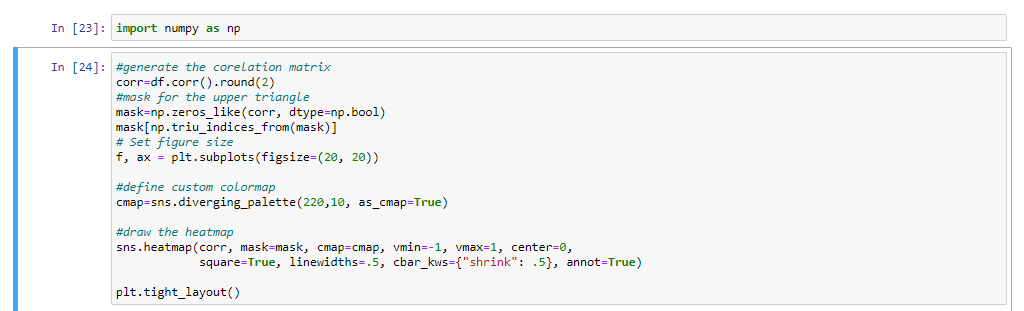
## VISUALIZATION:

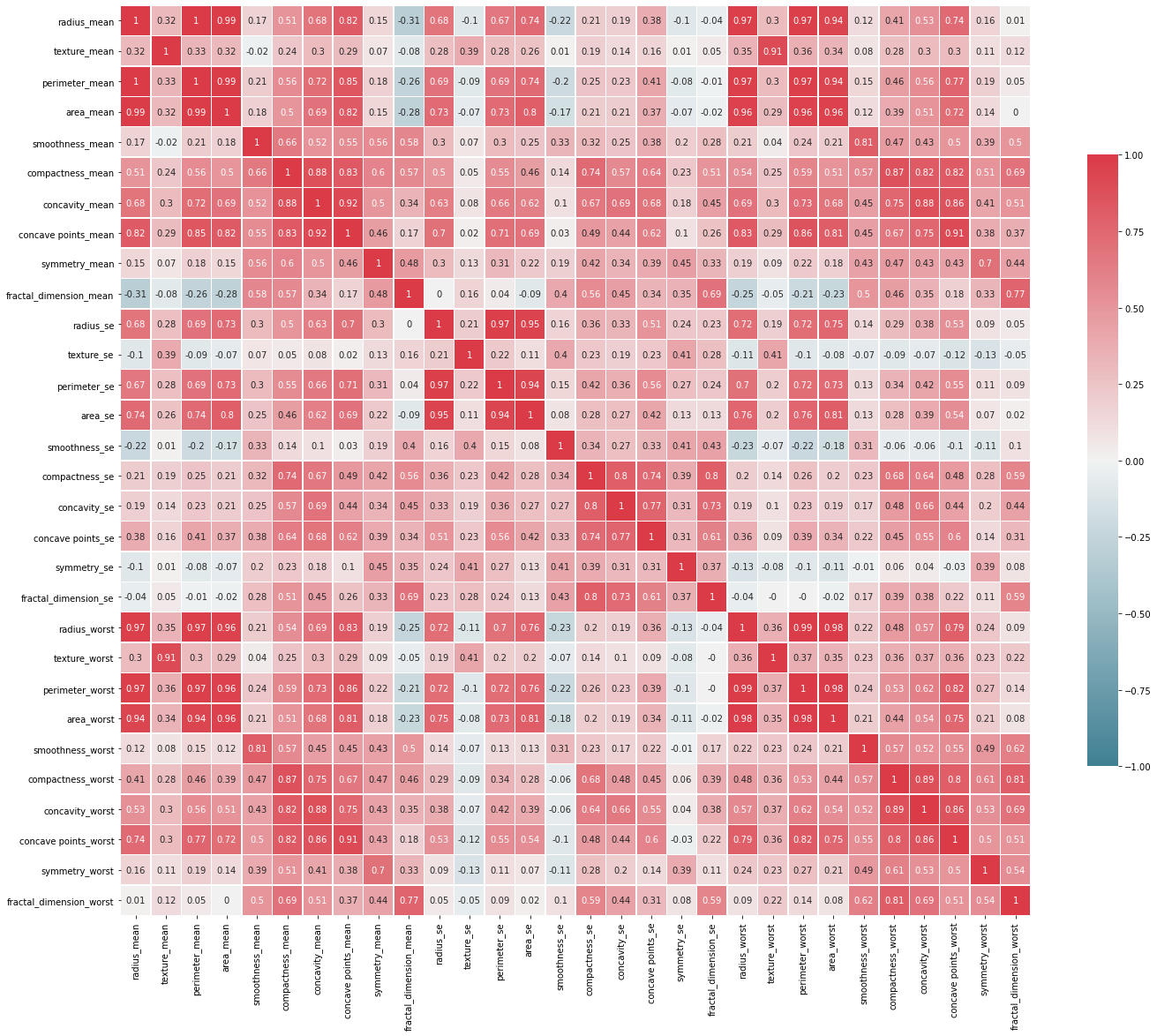


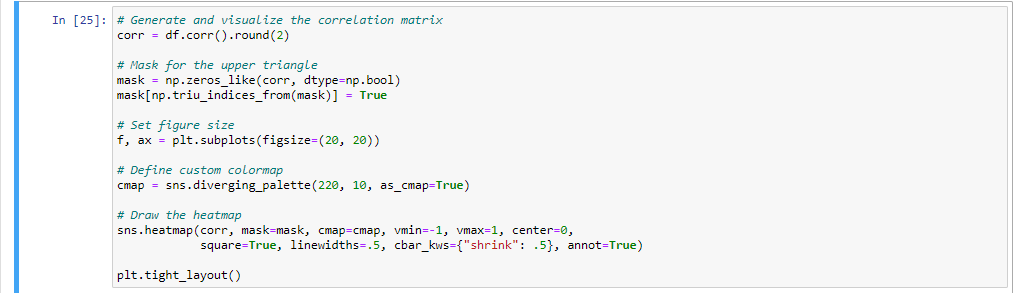


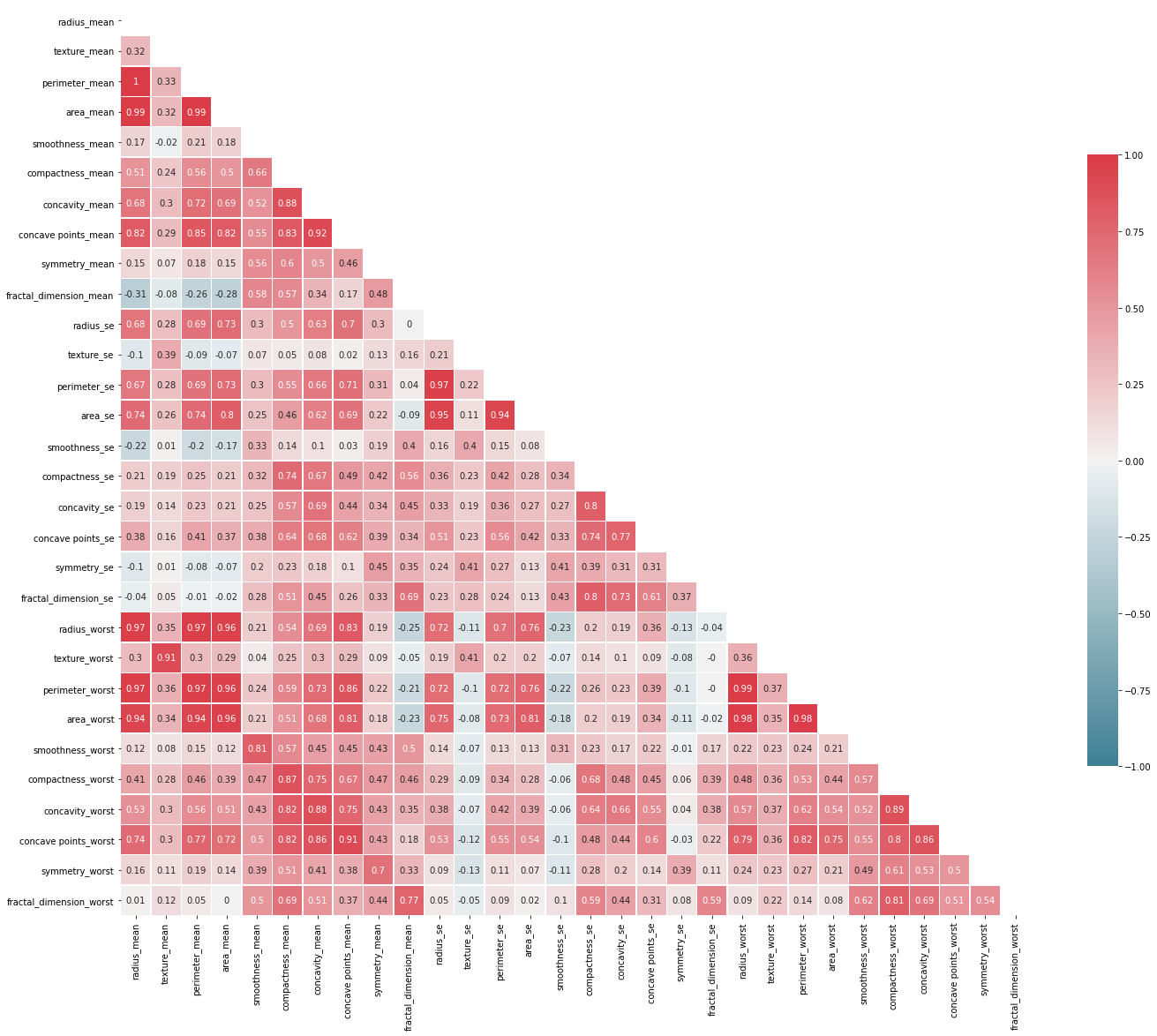


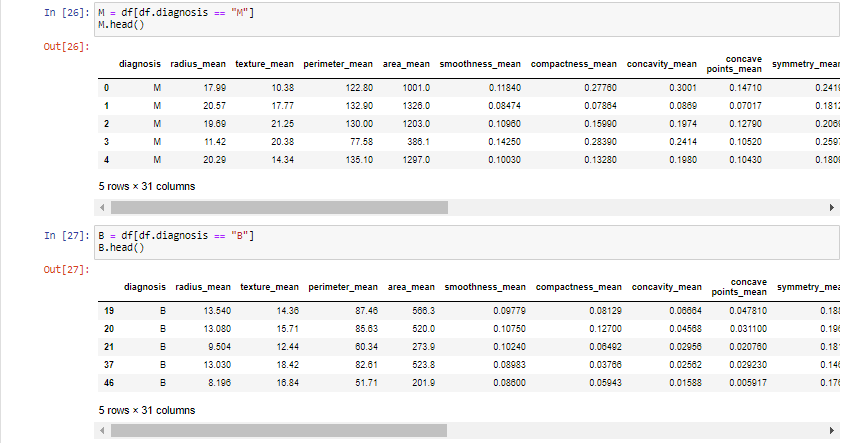


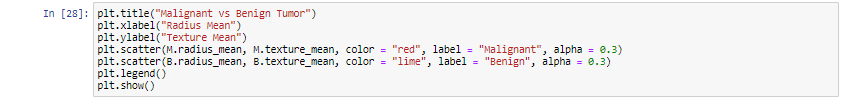


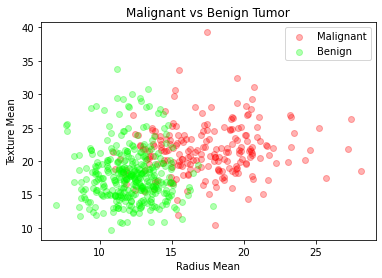




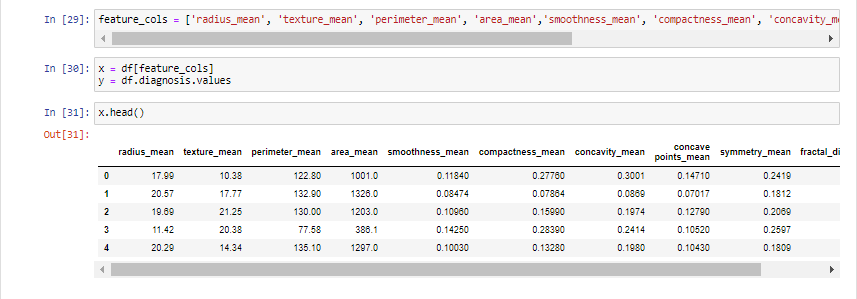


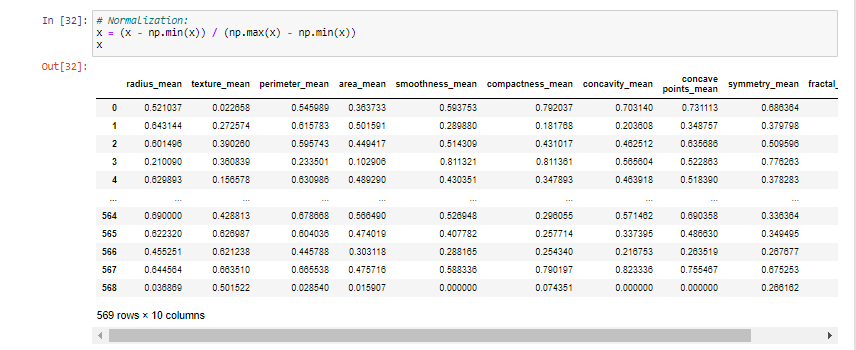


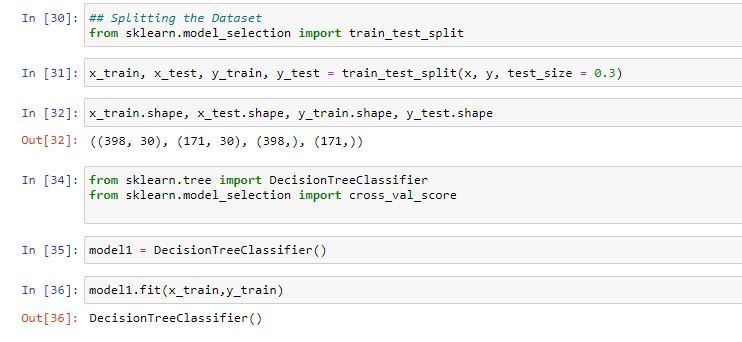


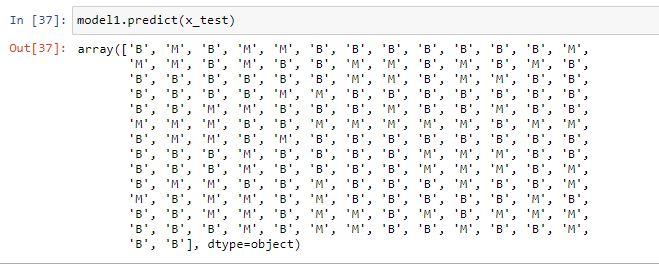


## ML ALGORITHM IMPLEMENTATION:

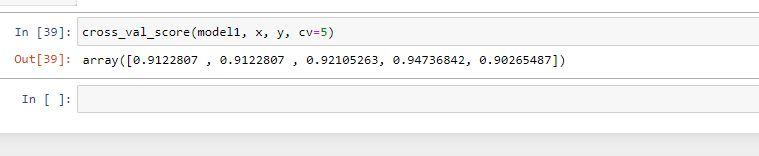








## FINAL RESULT:



**Github Link**

https://github.com/avnish9898/Ml-Experiment/blob/main/exp3.ipynb