**Google Collab link for the code is** :

<https://colab.research.google.com/drive/16jDEWsSptcZf1XIkODVfCZH2Css6dFnQ>

**Dataset used:**

Data set used for the lab is Brest cancer data from:

<https://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/>

## Information:

* Number of Instances: 699 (as of 15 July 1992)
* Number of Attributes: 10 plus the class attribute
* Attribute Information: (class attribute has been moved to last column)
* # Attribute Domain
  + 1. Sample code number id number
  + 2. Clump Thickness 1 - 10
  + 3. Uniformity of Cell Size 1 - 10
  + 4. Uniformity of Cell Shape 1 - 10
  + 5. Marginal Adhesion 1 - 10
  + 6. Single Epithelial Cell Size 1 - 10
  + 7. Bare Nuclei 1 - 10
  + 8. Bland Chromatin 1 - 10
  + 9. Normal Nucleoli 1 - 10
  + 10. Mitoses 1 - 10
  + 11. Class: (2 for benign, 4 for malignant)
* Missing attribute values: 16
* There are 16 instances in Groups 1 to 6 that contain a single missing
* (i.e., unavailable) attribute value, now denoted by "?".
* Class distribution:
* Benign: 458 (65.5%)
* Malignant: 241 (34.5%)

## Pre processing on data:

* Data is cleaned for any blank attribute value and with value”?”.
* Removed id number from features.
* Class 0 is given to “Benign” and Calss 1 is given to “Malignant”.
* Data is normalized using sklearn preprocessing module.

## Results:



Output.txt is submitted with the result

## Summary:

Worst performer: Decision Tree with accuracy of 79.5 %.

Best Performer: Bagging Classifier with 91.2%.

## References:

1. http://scikit-learn.org/
2. <https://archive.ics.uci.edu/ml/datasets>