MACHINE LEARNING ASSIGNMENT

Deadline for submission : 24 June 2018 by 11:59pm

Instructions:

- This assignment is for the session taken on Basics of Machine Learning.
- For submission follow the same instruction as previous assignments.
- Submit a single Jupyter Notebook file for the solution.
- All your assignments should go in Introduction-to-Data-Science/Week-6/ Assignment_5/<your_name>
- Attach a README.MD file which explains the flow of your code and any special function you are using.

Problem:

Build a simple classification model with the python machine learning library Scikit-learn. The dataset you will be using is the Breast Cancer Dataset which is one of the best dataset for binary classification. There are 2 categories as result (or target) which are Benign and Malignant. You need to build the model and try to achieve the best possible accuracy. The splitting of training and testing dataset should be done using the built in function train_test_split() with default parameters (as done in the session). Kindly look into other classifiers that Scikit-learn provide and try to use them to compare the accuracy between these classifiers. There is no strict compulsion to how many classifiers accuracy you are comparing. Try to process data accordingly to increase the accuracy of your model.

TO DOWNLOAD THE DATASET (Scikit-learn provides the dataset):

```
from sklearn import datasets

cancer = datasets.load_breast_cancer()
```

Link to the list of classifiers:

[http://scikit-learn.org/stable/auto_examples/classification/plot_classifier_comparison.html]