Name: Sambandh Bhusan Dhal
UIN : 726006080

Challenge Name:
To build a binary classifier based on SVM (Support Vector Machine) and submit the predictions on the testing data.

Specifications:
Both the training and testing datasets have 200 rows and 2 columns each. The training dataset has the corresponding labels attached with it whether they belong to class 0 or 1.

Demonstration:
Hyperparameter tuning and Kernel Selection:
I used K-fold cross-validation on the training data (K =10) and tried out three different kernels (linear, polynomial and radial) SVM with degree of the kernel ranging from 3 to 5 and different values of gamma ranging from 0.5 to 10 with an interval of 0.5 .
Here, 0.5 is the value of 'auto' where auto = 1/ number of features in the training data = ½ = 0.5
I even tried large values of gamma on the training data like 50 and 100.
Large values of gamma go on to overfit the model and we go on to achieve an accuracy of nearly 99% on the training data.
I used 10-fold cross-validation technique so as to prevent any overfitting on the training dataset as the size of the dataset is very small i.e. just 200 observations.
I have reported the different values of accuracy which I got for the 3 kernels below (gamma = 'auto' and degree = 3):
Linear SVM:

```
[ 0.75  0.85  0.75  0.75  0.85  0.85  0.8   0.9   0.85  0.85]
```

Polynomial SVM:(Degree =3 and Gamma = 0.5)

```
[ 0.75  0.65  0.8   0.85  0.95  0.8   0.65  0.95  0.85  0.75]
```

Radial SVM: (Degree = 3 and Gamma = 0.5)

```
[ 0.85  0.85  0.9   0.85  0.85  0.85  0.75  1.    0.85  0.9 ]
```

Using K-fold cross-validation technique, we can go on to conclude that Radial SVM goes on to give the highest accuracy as we report all the values of cross-validation accuracy which are greater than or equal to 85%.

Conclusion:
Therefore, I decided to train my model with **radial kernel SVM (gamma = 'auto') and degree of polynomial =3** and went on to achieve an accuracy of **88.5%** on the entire training dataset.

I used this kernel with all the same specifications to report the results on the testing data.
I have plotted the kernels (each for linear SVM, polynomial SVM and radial SVM) juxtaposed with the training data in the Jupyter notebook which I have submitted .

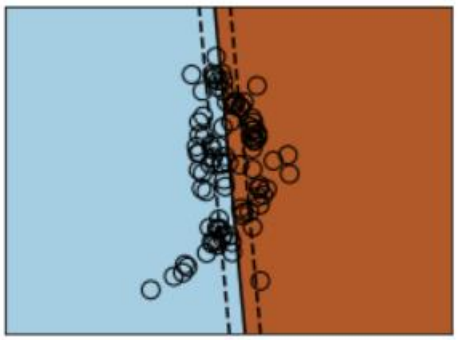

Fig 1: Linear SVM juxtaposed with the training set

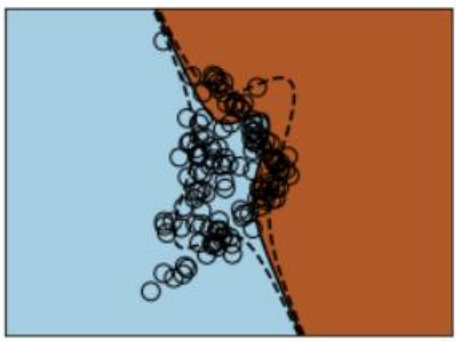

Fig 2: Polynomial SVM juxtaposed with the training set

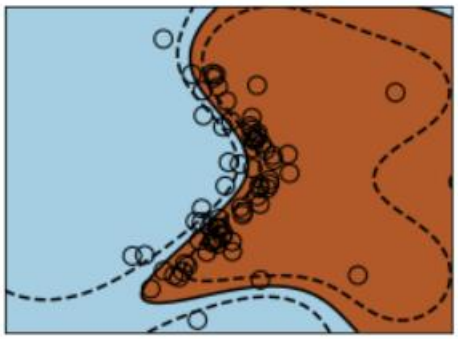

Fig 3: Radial SVM juxtaposed with the training set (Optimal Classifier which went on to achieve 88.5% accuracy on the training data)