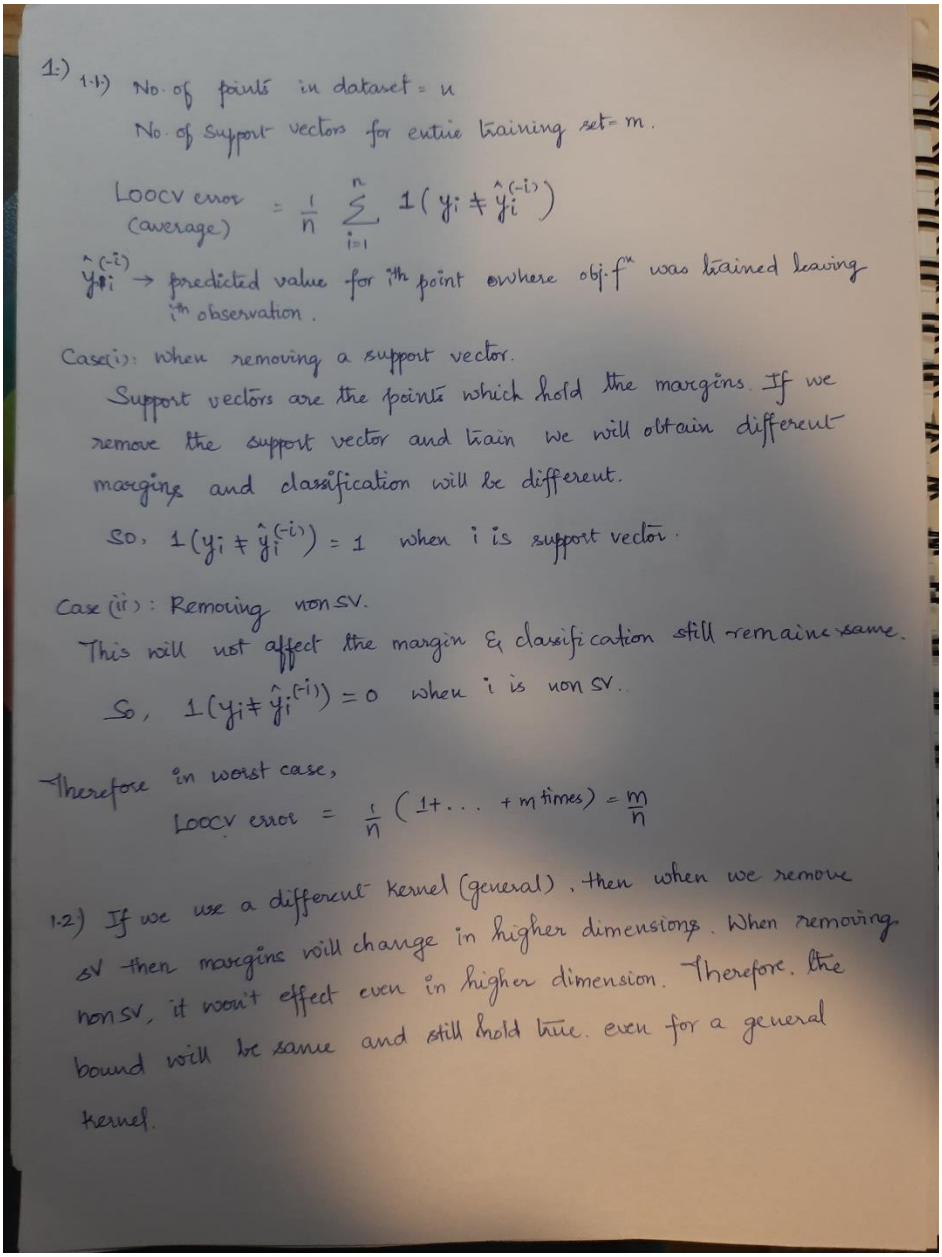
CSE 512 Spring 2021 - Machine Learning - Homework 4

Your Name: Irfan Ahmed
Solar ID : 113166464
NetID Email: irfan.ahmed@stonybrook.edu



1.) 1.1.) No. of points in dataset = n
No. of support vectors for entire training set = m.

$$\text{LOOCV error (average)} = \frac{1}{n}\sum_{i=1}^{n} 1(y_i \neq \hat{y}_i^{(-i)})$$

$\hat{y}_i^{(-i)}$ → predicted value for i^{th} point where obj. f^n was trained leaving i^{th} observation.

Case(i): When removing a support vector.
Support vectors are the points which hold the margins. If we remove the support vector and train we will obtain different margins and classification will be different.
So, $1(y_i \neq \hat{y}_i^{(-i)}) = 1$ when i is support vector.

Case (ii): Removing non SV.
This will not affect the margin & classification still remains same.
So, $1(y_i \neq \hat{y}_i^{(-i)}) = 0$ when i is non SV.

Therefore in worst case,
LOOCV error = $\frac{1}{n}(1 + \ldots + m \text{ times}) = \frac{m}{n}$

1.2) If we use a different kernel (general), then when we remove sv then margins will change in higher dimensions. When removing non sv, it won't effect even in higher dimension. Therefore, the bound will be same and still hold true. even for a general kernel.

Q2 – Colab link - (Open with SBU email)
https://colab.research.google.com/drive/1VkqQgnHuFRAZsgk75zU83s7rMq4n5xye?usp=sharing

Running k-fold for max_depth hyperparameter gives best accuracy for max_depth = 9

```
For max_depth = 3 - mean : 0.8624123226556447 , variance : 2.9060335094163053e-05
For max_depth = 4 - mean : 0.8684624366041315 , variance : 2.938011737040485e-05
For max_depth = 5 - mean : 0.8716257706893261 , variance : 3.796171406888943e-05
For max_depth = 6 - mean : 0.8724549712997671 , variance : 3.035801285392901e-05
For max_depth = 7 - mean : 0.8725777554147219 , variance : 3.8015468434759706e-05
For max_depth = 8 - mean : 0.8723935179492441 , variance : 3.5444160081458636e-05
For max_depth = 9 - mean : 0.8730385282427039 , variance : 2.7687810529702096e-05
For max_depth = 10 - mean : 0.8715336236674892 , variance : 2.637496412355618e-05
```

Running RandomSearchCV gives best results for :

```
search.best_params_
```

```
{'colsample_bytree': 0.6,
 'gamma': 0.4,
 'learning_rate': 0.15,
 'max_depth': 7,
 'min_child_weight': 4,
 'reg_alpha': 0.001,
 'subsample': 0.8}
```

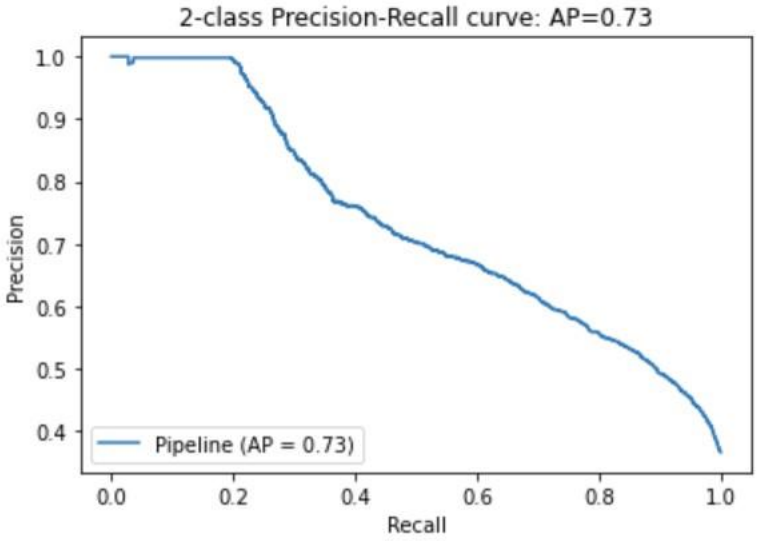
```
model = XGBClassifier(max_depth=7,learning_rate=0.15,ga
model.fit(X_train,y_train)
y_pred = model.predict(X_test)
print_scores(y_test,y_pred)
```

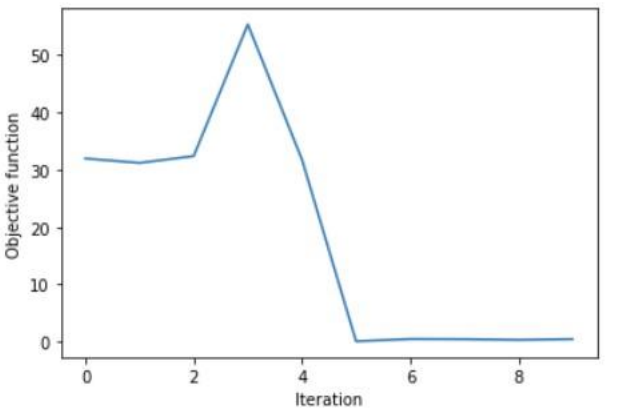
```
Accuracy : 0.8726122474049506

Confusion matrix :
[[11698   737]
[ 1337  2509]]

Average Precision Score : 0.5863674999321782
```

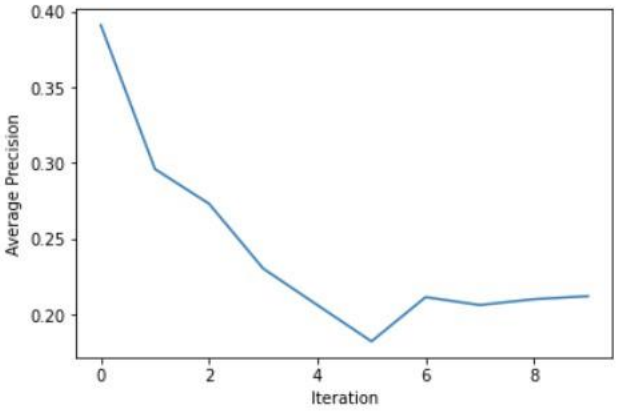
Q3 – Google Colab Notebook – (Open with SBU email)
https://colab.research.google.com/drive/1fErD7SKCmJguwgMcULBfiKRCWPk4_w8b?usp=sharing

a) Running Compute_map => `mAP: 7.703187293373048e-05`

```
Text(0.5, 1.0, '2-class Precision-Recall curve: AP=0.73')
```







c)

```
Objective function values : [31.937036569657288, 31.168152136733326,
32.354270444914505, 55.29917853638889, 31.76831902227367,
0.10988912467697536, 0.4987308974726487, 0.46639354913470965,
0.3530390225399158, 0.4786233359024562]

Average Precision values : [0.39090618555967294, 0.2960211807046842,
0.27313317288780287, 0.23042561514850513, 0.20636851670226034,
0.1824233263378137, 0.21160911953766776, 0.2063933478104194,
0.2102226018938276, 0.21224338556728095]
```

d) AP = 0.0000474