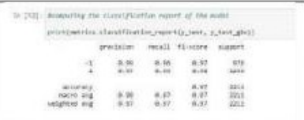
Project Development Phase
Model Performance Test

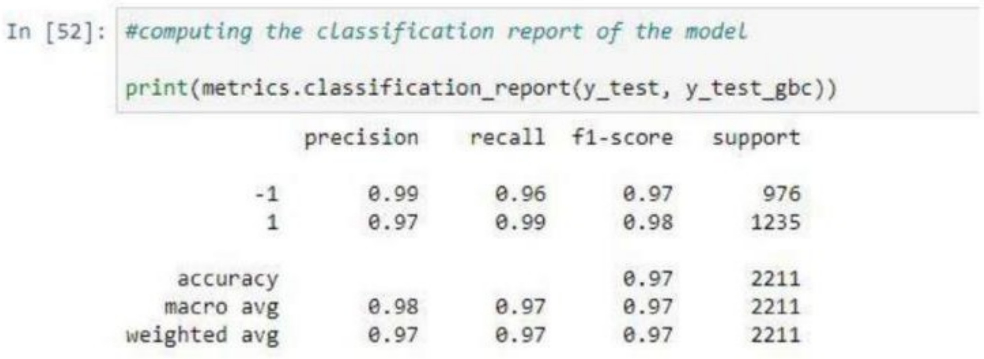
Date	10 November 2022
Team ID	PNT2022TMID06282
Project Name	Web Phishing Detection
Maximum Marks	10 Marks

Model Performance Testing:

Project team shall fill the following information in model performance testing template.

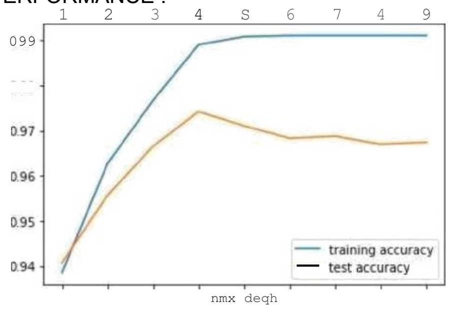
S.No.	Parameter	Values	Screenshot
1.	Metrics	**Classification Model:** **Gradient Boosting Classification** Accuray Score- 97.4%	
2.	Tune the Model	Hyperparameter Tuning - 97% Validation Method – KFOLD & Cross Validation Method	Wilcoxon signed-rank test [illegible]

1. METRICS:
CLASSIFICATION REPORT:



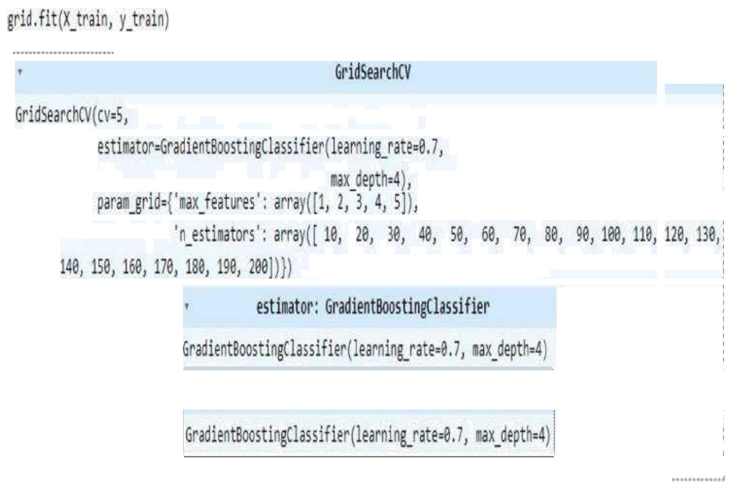
```
In [52]: #computing the classification report of the model

         print(metrics.classification_report(y_test, y_test_gbc))
                       precision    recall  f1-score   support

                  -1       0.99      0.96      0.97       976
                   1       0.97      0.99      0.98      1235

            accuracy                           0.97      2211
           macro avg       0.98      0.97      0.97      2211
        weighted avg       0.97      0.97      0.97      2211
```

PERFORMANCE :



	ML Model	Accuracy	fJ_score	Recall	Precision
0	Gradient Boosting C tassifier	0.974	0.977	0.994	0.986
1	CatBoost Classif er	0.972	0.975	0.994	0.989
2	Random Farest	0.969	0.972	0.992	0.991
3	Support Vector Machine	0.964	0.968	0.980	0.96S
4	Dec's on Tree	0.938	0.962	0.991	0.993
5	K-NeanxtNéghboc	0.956	0.9d1	0.991	0.989
6	Log st c Regression	0.934	0.941	0.943	0.927
	Naive Bayes Classified	0.605	0.454	0.292	0.997
8	XGBoost C tassifier	0.548	0.348	0.993	0.984
g	Multi-layer Perceptron	0.V3	0T43	0.989	0.983

2. TUNE THE MODEL - HYPERPARAMETER TUNING



```
grid.fit(X_train, y_train)

GridSearchCV

GridSearchCV(cv=5,
             estimator=GradientBoostingClassifier(learning_rate=0.7,
                                                  max_depth=4),
             param_grid={'max_features': array([1, 2, 3, 4, 5]),
                         'n_estimators': array([ 10,  20,  30,  40,  50,  60,  70,  80,  90, 100, 110, 120, 130,
       140, 150, 160, 170, 180, 190, 200])})

estimator: GradientBoostingClassifier
GradientBoostingClassifier(learning_rate=0.7, max_depth=4)

GradientBoostingClassifier(learning_rate=0.7, max_depth=4)
```

```
In [59]: print("The best parameters are %s with a score of %0.2f"
          | (jrul.hst , yr\J.kst s‹o'e ))
```

be oest yaraeters are (‘tax features “. 5, ’n estimators “. 280) Pitt z score of
0.97

VALIDATION METHODS: XFOLD & Cross Folding

Nlcoxon signed-rank test

```
In [78] #KFOLD and Cross Validation Model

        Tram sclpy .stats irqso›“t ›vilcoxsn
        4 in sklea rn . datasets Spot t load ie is
        T mm s klearn .ensemb le rpo›-l Grad1entBoost 1ngC 1a s s iller
        f row xgboosl 1iopor-l XG8C1ass1f1er
        T i am sklearn .mode 1 seT ec11 pn mpo r 4 c ross va} score, KFpld


        X    1c›ad iris( ) .data
        y = 1oad inxs ( } . target

        # Prepare models and select your CV method
        codell      GradientBoostingC1assl fl en {n  estimators - IN)
        model2 = JGBClassifier(n estimators=100)
        kf = KFod(n aplits2G, rundom_stetc=Nbne)
        # Extract results for each model on the same folds
        re su1ts  oode11  =  cros  s  va1  scsre(«odell,  X,  y,  cv=kf)
        resu1ts  uode12  -  cross  oval  score {‹sodel 2,   X,  y, cv=kf)
        stat,  p   =  ›vi1coxsn  result s  ncde 11,  result s       e12, zero method= ' z sp } lt ' );
        stat

Out[78] 95.B
```

5›‹2CV combined F test

```
In [89]: fron  mlxtend .ex'a1uate lnpoi l  coa+6ined_ftest_5x2c z
         f roo sk1earn.tree 1mpoi-t Oec1 s lonTreeC1as s *f 1er, E xt r a T ceeC1 as s Oficr
         f 'oo sk1earn .ensetrd1e impoi t Gradient oo stlngClas s iller
         I i am m1xtend . data lmpoi l iris data


         X, y = 1r1s_ data ()
         clll    -  GradientBoost1ngC1assLf 1‘er /' ,*
         c1f2 = Oec 1 s i on T reeC1 ass After ( )

         # Calculate p-value
         I, p - comb tried ftesi 5x2cv (est inotorl = c111,
                                        est1mator 2-c1f2,
                                        X=X, y=y,
                                        random seed•1)

         print('f-value:', f)
         p r iot ( ' y - va i ue : ' , p )

         I-x'a1ue: 1.727272727272733
         p ,alue: 0.2840z39734231?82
```