Project Development Phase

Model Performance Test

|  |  |
| --- | --- |
| Date | 13 November 2022 |
| Team ID | PNT2022TMID20741 |
| Project Name | Project — 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 |  |

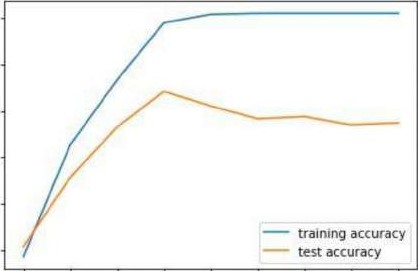
1. METRICS:

CLASSIFICATION REPORT:

In [52] : #computing the classification report of the model print(metrics.classification\_report(y\_test, y\_test\_gbc))

|  |  |  |  |
| --- | --- | --- | --- |
|  | precision | recall fl-score | support |
| -1 | e. 99 | e. 96 e .97 | 976 |
| 1 | e.97 | e.99 0.98 | 1235 |
| accuracy |  | e.97 | 2211 |
| macro avg | e. 98 | e." 0.97 | 2211 |
| weighted avg | e. 97 | e. 97 0.97 | 2211 |

PERFORMANCE :



1

2

3

4

5

6

7

8

9

nux\_depth

097

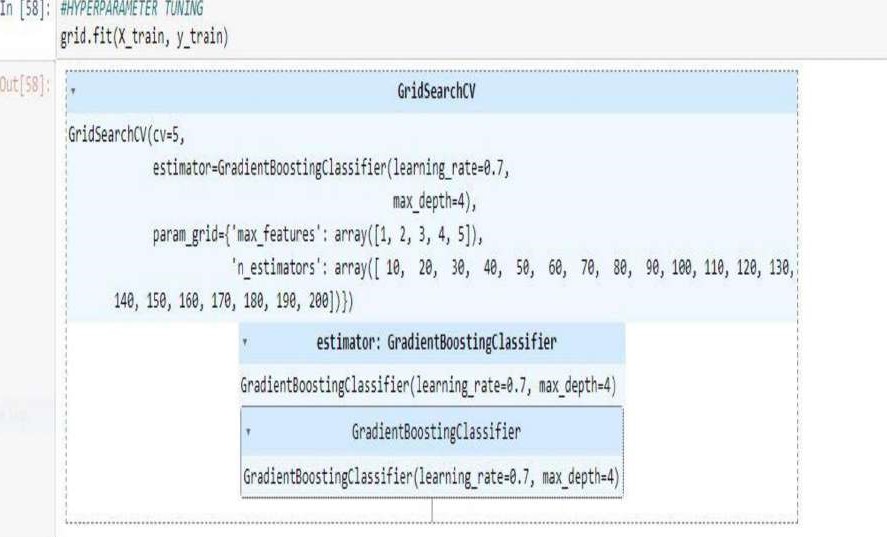
096

095

Out 33 J : ML Model Accuracy fl \_score Recall Precision

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Gradient Boosting Classifier | 0.974 | 0.977 | 0.994 | 0.986 |
| CatBoost Classifier | 0.972 | 0.975 | 0.994 | 0.989 |
| 2 Random Forest | 0.969 | 0.972 | 0.992 | 0.991 |
| 3 Support Vector Machine | 0.964 | 0.968 | 0.980 | 0.965 |
| 4 Decision Tree | 0.958 | 0.962 | 0.991 | 0.993 |
| 5 K-Nearest Neighbors | 0.956 | 0.961 | 0.991 | 0.989 |
| 6 Logistic Regression | 0.934 | 0.941 | 0.943 | 0.927 |
| Naive Bayes Classifier | 0.605 | 0454 | 0292 | 0.997 |
| 8 XGBoost Classifier | 0.548 | 0.548 | 0.993 | 0.984 |
| 9 Multi-layer Perceptron | 0.543 | 0.543 | 0.989 | 0983 |

1. TUNE THE MODEL - HYPERPARAMETER TUNING



|  |
| --- |
| In [59]:  X  (grid.  best\_params\_,  grid.best\_score\_))  The best parneters are {'max\_features': 5, 'n\_estimators': 20) with a score of 0.97 |

VALIDATION METHODS: KFOLD & Cross Folding

Wilcoxon signed-rank test

|  |  |
| --- | --- |
| In [78]: #KFOLD and Cross Validation  from scipy . stats import wilcoxon from sklearn. datasets import load\_iris from sklearn .ensemble import GradientBoostingC1assifier from xgboost import X6BC1assifier from sklearn .model\_selection import  KF01d  # Prepare models and select your CV method  — =   |  | | --- | | modell mode12 kf g  stat, stat |   Extract results for each model on the same folds results\_modell = X, y, cv=kf) results\_mode12 — cross\_va1\_score(mode12, X, y, cv=kf) p = results\_mode12, zsplit• );  outt78J: 9S.ø |

5x2CV combined F test

In [891: from mlxtend. evaluate import combined ftest\_5x2cv from sklearn. tree import DecisionTreeCIassifier, ExtraTreeCIassifier from sklearn.ensemble import GradientBoostingC1as5ifier from mlxtend.data import iris\_data # Prepare data and c Ifs

clfl GradientBoostingC1assifier() clf2 • DecisionTreeCIassifier()

# CaLcuLate p-vaLue f, p cortined estimator2=c1f2,

j print( f) print( •p-value: ' , p)

f-value: 1.727272727272733 p-value: 0.284ø135734291782