Model Optimization and Tuning Phase Template

|  |  |
| --- | --- |
| Date | 07 July 2024 |
| Team ID | SWTID1720451040 |
| Project Title | ECommerce Shipping Prediction Using Machine Learning |
| Maximum Marks | 10 Marks |

# Model Optimization and Tuning Phase

In order to attain optimal performance, machine learning models undergo optimization and tuning at this phase. It entails fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection in order to increase anticipated accuracy and efficiency.

# Hyperparameter Tuning Documentation (6 Marks):

|  |  |  |
| --- | --- | --- |
| **Model** | **Tuned Hyperparameters** | **Optimal Values** |
| SVM | c, kernel, gamma | 1.0, rbf, 0.01 |
| random forest | n\_Estimators, Criterion, Max\_Depth, Max\_features | none,1e-9 |
| KNN | n\_neighbors, weights, algorithm, p | 25, uniform, auto, 2 |
| XGBoost | booster | gbtree |
|  |  |  |

|  |  |  |
| --- | --- | --- |
| ANN | Units, kernel\_initialiser, activation | Input layer: 16, ‘random\_uniform’, ‘relu’  First Hidden Layer:  16, ‘random\_uniform’, ‘relu’  Second Hidden Layer: 8, ‘random\_uniform’, ‘relu’  Output layer: 1, ‘random\_uniform’, ‘relu’ |

**Performance Metrics Comparison Report (2 Marks):**

|  |  |  |
| --- | --- | --- |
| **Model** | **Baseline Metric** | **Optimized Metric** |
| SVM | Accuracy,F1 Score:    : | Accuracy,F1 Score: |

|  |  |  |
| --- | --- | --- |
| random forest | Accuracy,F1 Score: | Accuracy,F1 Score: |
| KNN | Accuracy,F1 Score: | Accuracy,F1 Score: |
| XGBoost | Accuracy,F1 Score: | Accuracy,F1 Score: |

|  |  |  |
| --- | --- | --- |
| ANN | Accuracy,F1 Score: | Accuracy,F1 Score: |

# Final Model Selection Justification (2 Marks):

|  |  |
| --- | --- |
| **Final Model** | **Reasoning** |
| **Random Forest** | The Random Forest model was chosen as the most optimized model due to its superior performance metrics. Based on its best accuracy of 68.42%, it was effective in producing accurate forecasts  .  Additionally, it showed a high precision score of 93.00%, proving its dependability in correctly identifying genuine positives. Random Forest's ensemble approach minimizes overfitting and improves generalization to new data. In keeping with the objectives of the project, Random Forest is the ideal choice for enhancing delivery time projections because of these characteristics. |