**Model Optimization and Tuning Phase Template**

|  |  |
| --- | --- |
| Date | 20 June 2025 |
| Team ID | SWTID1749709635 |
| Project Title | Mental Health Prediction |
| Maximum Marks | 10 Marks |

**Model Optimization and Tuning Phase**

The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

### Hyperparameter Tuning Documentation (6 Marks):

|  |  |  |
| --- | --- | --- |
| **Model** | **Tuned Hyperparameters** | **Optimal Values** |
| Logistic Regression | max\_iter, random\_state | max\_iter = 1000, random\_state = 42 |
| K- Nearest Neighbors | n\_neighbors | n\_neighbors = 5 (default value) |
| Decision Tree | random\_state | random\_state = 42 |
| Random Forest | random\_state | random\_state = 42 |
| Naïve Bayes | None | GaussianNB uses default |
| Support Vector Machines | probability, random\_state | probability = True, random\_state = 42 |
| Extreme Gradient Boosting | use\_label\_encoder, eval\_metric, random\_state | use\_label\_encoder = False, eval\_metric = ‘logloss’, random\_state = 42 |
| Adaptive Boosting | random\_state | random\_state = 42 |
| Gradient Boosting | random\_state | random\_state = 42 |

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

|  |  |  |
| --- | --- | --- |
| **Model** | **Baseline Metric** | **Optimized Metric** |
| Logistic Regression | Accuracy = 0.7035 | Accuracy = 0.7559 |
| K- Nearest Neighbors | Accuracy = 0.6786 | Accuracy = 0.7165 |
| Decision Tree | Accuracy = 0.8087 | Accuracy = 0.8504 |
| Rnadom Forest | Accuracy = 0.8644 | Accuracy = 0.9213 |
| Naïve Bayes | Accuracy = 0.7093 | Accuracy = 0.7205 |
| Support Vector Machines | Accuracy = 0.7416 | Accuracy = 0.7795 |
| Extreme Gradient Boosting | Accuracy = 0.8370 | Accuracy = 0.8701 |
| Adaptive Boosting | Accuracy = 0.7062 | Accuracy = 0.7480 |
| Gradient Boosting | Accuracy = 0.8059 | Accuracy = 0.8386 |

### Final Model Selection Justification (2 Marks):

|  |  |
| --- | --- |
| **Final Model** | **Reasoning** |
| Random Forest | Random Forest was chosen as the final model because it achieved the highest accuracy (92%) after hyperparameter tuning. It also offers robust handling of categorical and numerical features, resistance to overfitting, interpretability through feature importances, and fast prediction performance. |