# Model Optimization and Tuning Phase

| | |
|---|---|
| Date | 9th July 2024 |
| Team ID | SWTID1720017249 |
| Project Title | Panic Disorder Detection |
| 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 |
|---|---|---|
| Gradient Boosting | n_estimators, learning_rate, max_depth, min_samples_split, min_samples_leaf, subsample | n_estimators: 200, learning_rate: 0.1, max_depth: 4, min_samples_split: 5, min_samples_leaf: 2, subsample: 0.9 |
| XGBoost | n_estimators, learning_rate, max_depth, subsample, colsample_bytree, gamma | n_estimators: 300, learning_rate: 0.1, max_depth: 4, subsample: 0.8, colsample_bytree: 0.9, gamma: 0.1 |
| Decision Tree | criterion, splitter, max_depth, min_samples_split, min_samples_leaf, max_features | criterion: 'gini', splitter: 'best', max_depth: 30, min_samples_split: 2, min_samples_leaf: 1, max_features: None |

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

| Model | Baseline Metric (Accuracy) | Optimized Metric (Accuracy) |
|---|---|---|
| Decision Tree | 90.45 | 90.34 |
| Gradient Boosting | 90.13 | 91.54 |
| XGBoost | 91.12 | 91.51 |

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

| Final Model | Reasoning |
|---|---|
| Gradient Boosting | Gradient Boosting achieved the highest optimized accuracy of 91.54%, showing a significant improvement over its baseline accuracy of 90.13%. This indicates that the model benefits greatly from hyperparameter tuning. Although XGBoost also performed well with an optimized accuracy of 91.51%, the slight edge in accuracy and the stability of Gradient Boosting make it the preferred choice. Furthermore, Gradient Boosting has shown to be less prone to overfitting in this context compared to Decision Tree, whose optimized accuracy decreased. Therefore, Gradient Boosting is selected for its superior performance and robustness after optimization. |