

## Model Optimization and Tuning Phase Report

Date	15 July 2024
Team ID	739706
Project Title	One year life expectancy post on Thoracic Surgery using machine learning
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

Random Forest	<pre># Initialize Random Forest classifier rf = RandomForestClassifier(random_state=42)  # Define the parameter grid for hyperparameter tuning with more values param_grid = {     'n_estimators': [100, 200, 300, 400, 500],     'max_features': ['auto', 'sqrt', 'log2'],     'max_depth': [None, 10, 20, 30, 40, 50],     'min_samples_split': [2, 5, 10, 15],     'min_samples_leaf': [1, 2, 4, 8],     'bootstrap': [True, False] }</pre>	<pre>print(f'Accuracy: {accuracy}') print(f'F1 Score: {f1}') print('Classification Report:') print(classification_report(y_test, y_pred)) print('Confusion Matrix:') print(cm)</pre>
---------------	--	--

Performance Metrics Comparison Report (2 Marks):

Model	Optimal Metric
Random Forest	<pre>Fitting 5 folds for each of 2000 candidates, totalling 10000 fits Best Parameters: {'bootstrap': True, 'max_depth': None, 'max_features': 'sqrt', 'min_samples_leaf': 1, 'min_samples_split': 5, 'n_estimators': 200} Best Score: 0.821553877625521 Scorecv: 0.87104831640019 F1 Score: 0.760158820611074 Classification Report:               precision    recall  f1-score   support       0       0.84      1.00      0.91         70      1       0.80      0.80      0.80         35   accuracy: 0.88  macro avg: 0.82   0.90   0.86   95  weighted avg: 0.79   0.84   0.79   95  Confusion Matrix: [[70  0]  [ 0 35]]</pre>

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

Final Model	Reasoning
<b>Best random forest</b>	The Random Forest model was selected for its superior performance, exhibiting high accuracy during hyperparameter tuning. Its ability to handle complex relationships, minimize overfitting, and optimize predictive accuracy aligns with project objectives, justifying its selection as the final model."