

## Model Optimization and Tuning Phase Template

Date	12 July 2024
Team ID	xxxxxx
Project Title	Human Resource Management Predicting Employee Promotions 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
DecisionTreeClassifier	<pre>param_grid = {     'criterion': ['gini', 'entropy'],     'max_depth': [None, 10, 20, 30, 40, 50],     'min_samples_split': [2, 5, 10],     'min_samples_leaf': [1, 2, 4] }</pre>	<pre>{'criterion': 'gini',  'max_depth': None,  'min_samples_leaf': 1,  'min_samples_split': 5}</pre> <p>Accuracy Score:</p> <p>0.941</p>

RandomForest Classifier	<pre>param_grid = {     'n_estimators': [10, 50, 100, 200],     'criterion': ['gini', 'entropy'],     'max_depth': [None, 10, 20, 30, 40, 50],     'min_samples_split': [2, 5, 10],     'min_samples_leaf': [1, 2, 4] }</pre>	<pre>{'criterion': 'entropy',  'max_depth': 40,  'min_samples_leaf': 1,  'min_samples_split': 5,  'n_estimators': 200}</pre> <p>Accuracy Score:</p> <p>0.958</p>
KNeighborsClassifier	<pre>param_grid = {     'n_neighbors': [3, 5, 7, 9, 11],     'weights': ['uniform', 'distance'],     'algorithm': ['auto', 'ball_tree', 'kd_tree', 'brute'],     'p': [1, 2] }</pre>	<pre>{'algorithm': 'ball_tree',  'n_neighbors': 3, 'p': 1,  'weights': 'distance'}</pre> <p>Accuracy Score:</p> <p>0.928</p>
GradientBoostingClassifier	<pre>param_grid = {     'n_estimators': [50, 100, 200],     'max_depth': [3, 6, 9],     'learning_rate': [0.01, 0.1, 0.2],     'subsample': [0.8, 1.0],     'colsample_bytree': [0.8, 1.0] }</pre>	<pre>{'colsample_bytree': 1.0,  'learning_rate': 0.2,  'max_depth': 9,  'n_estimators': 200,  'subsample': 0.8}</pre> <p>Accuracy Score:</p> <p>0.945</p>

## Performance Metrics Comparison Report (2 Marks):

Model	Baseline Metric	Optimized Metric
DecisionTreeClassifier	<pre> precision    recall  f1-score   support        0       0.95      0.93      0.94      15065       1       0.93      0.95      0.94      15019   accuracy          0.94      30084  macro avg          0.94      30084 weighted avg          0.94      30084 </pre>	<pre> Classification Report: precision    recall  f1-score   support        0       0.95      0.94      0.94      15065       1       0.94      0.95      0.94      15019   accuracy          0.94      30084  macro avg          0.94      30084 weighted avg          0.94      30084  Accuracy Score: 0.9413974205557771 </pre>
RandomForestClassifier	<pre> precision    recall  f1-score   support        0       0.96      0.95      0.96      15065       1       0.95      0.96      0.96      15019   accuracy          0.96      30084  macro avg          0.96      30084 weighted avg          0.96      30084 </pre>	<pre> Classification Report: precision    recall  f1-score   support        0       0.96      0.96      0.96      15065       1       0.96      0.96      0.96      15019   accuracy          0.96      30084  macro avg          0.96      30084 weighted avg          0.96      30084  Accuracy Score: 0.9580507911182023 </pre>
KNeighborsClassifier	<pre> precision    recall  f1-score   support        0       0.98      0.83      0.90      15065       1       0.85      0.98      0.91      15019   accuracy          0.91      30084  macro avg          0.91      30084 weighted avg          0.91      30084 </pre>	<pre> Classification Report: precision    recall  f1-score   support        0       0.97      0.88      0.93      15065       1       0.89      0.97      0.93      15019   accuracy          0.93      30084  macro avg          0.93      30084 weighted avg          0.93      30084  Accuracy Score: 0.9287661215263928 </pre>
GradientBoostingClassifier	<pre> precision    recall  f1-score   support        0       0.90      0.81      0.85      15065       1       0.82      0.91      0.87      15019   accuracy          0.86      30084  macro avg          0.86      30084 weighted avg          0.86      30084 </pre>	<pre> Classification Report: precision    recall  f1-score   support        0       0.95      0.94      0.95      15065       1       0.94      0.95      0.95      15019   accuracy          0.95      30084  macro avg          0.95      30084 weighted avg          0.95      30084  Accuracy Score: 0.9457186544342507 </pre>

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

Final Model	Reasoning
RandomForestClassifier	<p>The RandomForestClassifier was chosen as the final optimized model due to its impressive performance, achieving 95.8% accuracy. Its robustness to overfitting, ability to handle large datasets, and high accuracy make it ideal for the task. The ensemble approach, which involves averaging multiple decision trees, enhances stability and generalization to unseen data. Additionally, random forests efficiently manage datasets with numerous features and provide insights into feature importance. This model's balance of accuracy, robustness, and interpretability, along with its versatility, confirms its suitability for the problem at hand.</p>