



Model Optimization and Tuning Phase Template

Date	15 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
Decision Tree		
	<pre># Function to train and evaluate a Decision Tree model with hyperparameter tuning def decisionTree(X_train, X_test, y_train, y_test): # Define the parameter grid param grid = {</pre>	Best Parameters found by GridSearchOl: ('criterion': 'entropy', 'man_depth': 40, 'min_samples_leaf': 1, 'min_samples_split': 2)
	}	Accuracy: 0.94







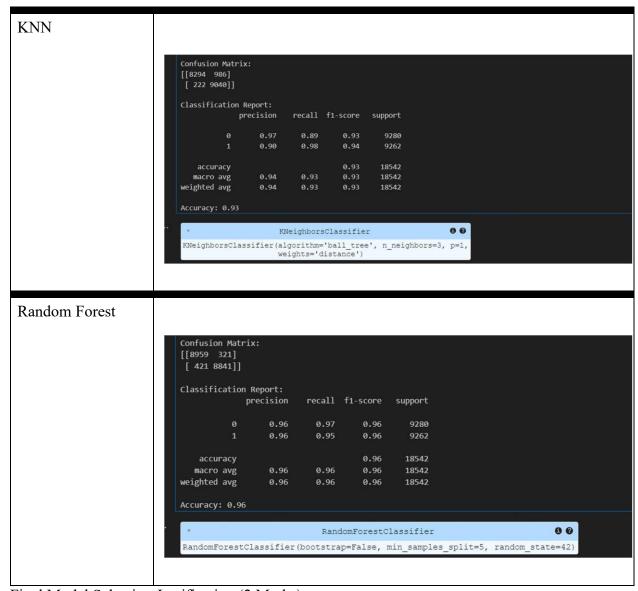


Performance Metrics Comparison Report (2 Marks):

	Optimized Metric					
Decision Tree						
	Confusion Matrix: [[8668 612] [453 8809]]					
	Classification Rep	ort:				
			recall	f1-score	support	
	Ø	0.95	0.93	0.94	9280	
	1	0.94	0.95	0.94	9262	
	accuracy			0.94	18542	
	macro avg weighted avg	0.94 0.94	0.94 0.94	0.94 0.94	18542 18542	
		0.94	0.94	0.94	16342	
	Accuracy: 0.94					
			Decisi	onTreeCla	ssifier	9 9
	DecisionTreeClas	sifier(c	riterion	n='entropy	', max depth=40,	random_state=42)
	1				_	
Gradient Boost						
Gradient Boost	Confusion Matrix					
Gradient Boost	Confusion Matrix [[8678 602]	:				
Gradient Boost		:				
Gradient Boost	[[8678 602] [695 8567]]					
Gradient Boost	[[8678 602] [695 8567]] Classification Ro		recall	f1-score	support	
Gradient Boost	[[8678 602] [695 8567]] Classification Re	eport: ecision		f1-score	support	
Gradient Boost	[[8678 602] [695 8567]] Classification Ro	eport:	recall 0.94 0.92	f1-score 0.93		
Gradient Boost	[[8678 602] [695 8567]] Classification Re pro 0 1	eport: ecision 0.93	0.94	f1-score 0.93 0.93	support 9280 9262	
Gradient Boost	[[8678 602] [695 8567]] Classification Re pro	eport: ecision 0.93	0.94	f1-score 0.93	support 9280	
Gradient Boost	[[8678 602] [695 8567]] Classification Re pro 0 1 accuracy	eport: ecision 0.93 0.93	0.94 0.92	f1-score 0.93 0.93	support 9280 9262 18542	
Gradient Boost	[[8678 602] [695 8567]] Classification Re pro 0 1 accuracy macro avg	eport: ecision 0.93 0.93	0.94 0.92 0.93	f1-score 0.93 0.93 0.93 0.93	support 9280 9262 18542 18542	
Gradient Boost	[[8678 602] [695 8567]] Classification Represented for the property of the p	eport: ecision 0.93 0.93	0.94 0.92 0.93 0.93	f1-score 0.93 0.93 0.93 0.93 0.93	support 9280 9262 18542 18542 18542	
Gradient Boost	[[8678 602] [695 8567]] Classification Reproduction of the production of the produ	eport: ecision 0.93 0.93 0.93	0.94 0.92 0.93 0.93	f1-score 0.93 0.93 0.93 0.93	support 9280 9262 18542 18542 18542	• •
Gradient Boost	[[8678 602] [695 8567]] Classification Reproduction of the production of the produ	eport: ecision 0.93 0.93 0.93	0.94 0.92 0.93 0.93 Grad	f1-score 0.93 0.93 0.93 0.93 iientBoost: rning_rate	support 9280 9262 18542 18542 18542 ingClassifier e=0.2, max_depth=	5, min_samples_split=5, ==42, subsample=0.8)







Final Model Selection Justification (2 Marks):

Final Model	Reasoning
Random Forest	I chose the Random Forest model as the final model for predicting employee promotions due to its superior accuracy (96%) compared to other models like Decision Tree, KNN, and Gradient Boosting. Random Forest is robust, handles overfitting well, and provides insights into feature importance. It captures complex, non-linear relationships within the data and is scalable for large datasets. Additionally, hyperparameter tuning further optimized its performance, making it a reliable and efficient choice for this task