

Model Optimization and Tuning Phase Template

Date	24 April 2024
Team ID	team-739704
Project Title	Identifying Airline Passenger Satisfaction 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>from sklearn.model_selection import GridSearchCV param_grid = { 'criterion': ['gini', 'entropy'], 'max_depth': [None, 5, 10, 15], 'min_samples_split': [2, 5, 10], 'min_samples_leaf': [1, 2, 4] } tree = DecisionTreeClassifier() grid_search = GridSearchCV(estimator=tree, param_grid=param_grid,</pre>	<pre>grid_search= GridSearchCV(estimator= tree,param_grid=param_grid) grid_search=grid_search.fit(X_train,Y_train) print("Best accuracy=",grid_search.best_score_) print("Best parameters=",grid_search.best_params_) warnings.warn(Best accuracy= 0.9244474806826352 Best parameters= {'criterion': 'entropy',</pre>

Random Forest	<pre>parameter={ 'n_estimators':[25,50,200,300], 'criterion':['gini','entropy'], 'max_depth':[14,20,25,30]}</pre>	<pre>grid_search= GridSearchCV(estimator= forest,param_grid=parameter,s grid_search=grid_search.fit(X_train,Y_train) print("Best accuracy=",grid_search.best_score_) print("Best parameters=",grid_search.best_params_) /usr/local/lib/python3.10/dist-packages/sklearn/model_selection/_s warnings.warn(Best accuracy= 0.9432758844558482 Best parameters= {'criterion': 'gini', 'max_depth': 30, 'n_estimat</pre>
---------------	---	---

Performance Metrics Comparison Report (2 Marks):

Model	Optimized Metric
Decision Tree	<pre>-----Decision Tree----- Model accuracy {0.9116179615110478} Accuracy in Percentage 91.2% precision recall f1-score support 1 0.92 0.93 0.92 2357 2 0.91 0.89 0.90 1852 accuracy 0.91 4209 macro avg 0.91 0.91 4209 weighted avg 0.91 0.91 4209</pre> <pre>cm=confusion_matrix(Y_test,Y_pred) cm array([[2201, 156], [204, 1648]])</pre>

Random Forest	<pre> -----RandomForest classifier----- Model accuracy {0.9453551912568307} Accuracy in Percentage 94.5% precision recall f1-score support 1 0.93 0.97 0.95 2357 2 0.96 0.91 0.94 1852 accuracy 0.95 0.95 4209 macro avg 0.95 0.94 0.94 4209 weighted avg 0.95 0.95 0.95 4209 </pre> <pre> cm=confusion_matrix(Y_test,Y_pred) cm array([[1119, 48], [88, 861]]) </pre>
---------------	--

Final Model Selection Justification (2 Marks):

Final Model	Reasoning
Random Forest	<p>The Random Forest Boosting 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.</p>