

Model	Tuned Hyperparameters	Optimal Values
Linear regression	-----	<pre>lr = LinearRegression() lr.fit(train_X,train_y) print('Attempting to fit Linear Regressor')</pre> <p>Attempting to fit Linear Regressor</p>

Random Forest	-----	<pre>from sklearn.ensemble import RandomForestRegressor rf = RandomForestRegressor() rf.fit(train_X,train_y) print('Attempting to fit Random Forest Regressor')</pre> <p>Attempting to fit Random Forest Regressor</p>
---------------	-------	--

Model Optimization and Tuning Phase Report

Date	20 June 2024
Team ID	740114
Project Title	Customer Acquisition Cost Estimation Using ML
Maximum Marks	10 Marks

Model Optimization and Tuning Phase

In the model optimization and tuning phase for customer acquisition cost estimation using machine learning, split the data, select key hyperparameters (e.g., `estimators`, `max_depth`), and use `Randomized SearchCV` or `GridSearchCV` to identify optimal values. Evaluate performance using metrics like Mean Absolute Error (MAE) or Mean Squared Error (MSE).

Hyperparameter Tuning Documentation (6 Marks):

--	--	--

Performance Metrics Comparison Report (2 Marks):

Model	Optimized Metric
-------	------------------

Random Forest	<pre> y_pred_val_rf = rf.predict(val_X) print('MAE on Validation set :',metrics.mean_absolute_error(val_y, y_pred_val_rf)) print("\n") print('MSE on Validation set :',metrics.mean_squared_error(val_y, y_pred_val_rf)) print("\n") print('RMSE on Validation set :',np.sqrt(metrics.mean_squared_error(val_y, y_pred_val_rf))) print("\n") print('R2 Score on Validation set :',metrics.r2_score(val_y, y_pred_val_rf)) print("\n") </pre> <p>MAE on Validation set : 0.0925440344201496</p> <p>MSE on Validation set : 1.7262364914711157</p> <p>RMSE on Validation set : 0.3042105100422232</p> <p>R2 Score on Validation set : 0.9980566391348797</p>
Linear Regressor	<pre> y_pred_val_lr = lr.predict(val_X) print('MAE on Validation set :',metrics.mean_absolute_error(val_y, y_pred_val_lr)) print("\n") print('MSE on Validation set :',metrics.mean_squared_error(val_y, y_pred_val_lr)) print("\n") print('RMSE on Validation set :',np.sqrt(metrics.mean_squared_error(val_y, y_pred_val_lr))) print("\n") print('R2 Score on Validation set :',metrics.r2_score(val_y, y_pred_val_lr)) print("\n") </pre> <p>MAE on Validation set : 25.212882223695512</p> <p>MSE on Validation set : 862.7559482129169</p> <p>RMSE on Validation set : 5.0212430954590825</p> <p>R2 Score on Validation set : 0.028727434398146512</p>

Final Model Selection Justification (2 Marks):

Final Model	Reasoning
Random Forest Model	The Random Forest model was selected for its superior performance, exhibiting high accuracy. Its ability to handle complex relationships, minimize overfitting, and optimize predictive accuracy aligns with project objectives, justifying its selection as the final model.

