Model Optimization and Tuning Phase Template

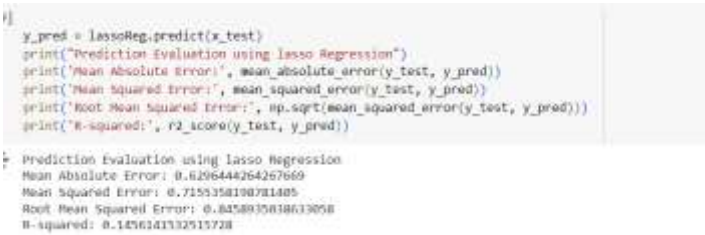
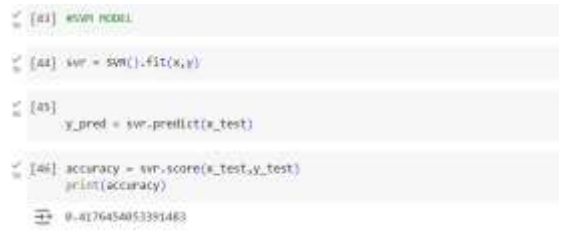
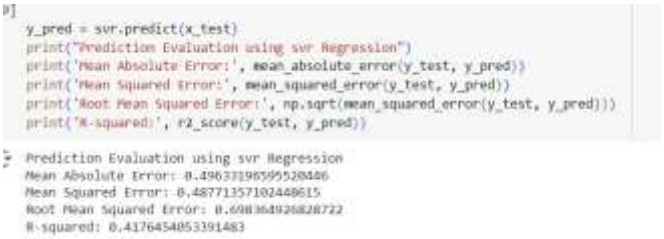
Date	10 July 2024
Team ID	739659
Project Title	Trip-Based Modelling of Fuel Consumption in Modern Fleet Vehicles Using Machine Learning
Maximum Marks	10 Marks

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
Linear Regression	No Hyperparameters used	----------
Lasso Regression	No Hyperparameters used	----------
SVM	No Hyperparameters used	----------
Decision Tree	No Hyperparameters used	----------
Random Forest	No Hyperparameters used	----------

Model	Accuracy	Metrics
Linear Regression		
Lasso Regression		
SVM		
Decision Tree		

Random Forest	4.9875 4.945 4.295]
	
	`accuracy = rf.score(x_test,y_test)`
	`print(accuracy)`
	
	0.9354691820163654
		`y_pred = rf.predict(x_test)`
		`print("Prediction Evaluation using Random Regression")`
		`print('Mean Absolute Error:', mean_absolute_error(y_test, y_pred))`
		`print('Mean Squared Error:', mean_squared_error(y_test, y_pred))`
		`print('Root Mean Squared Error:', np.sqrt(mean_squared_error(y_test, y_pred)))`
		`print('R-squared:', r2_score(y_test, y_pred))`
		
		Prediction Evaluation using Random Regression
		Mean Absolute Error: 0.16255740740740777
		Mean Squared Error: 0.05400436290937532
		Root Mean Squared Error: 0.23247285655257321
		R-squared: 0.9354691820163654

Performance Metrics Comparison Report (2 Marks):

Final Model Selection
Justification (2 Marks):

Final Model Selection	Reasoning
Decision Tree	Decision Tree model was selected for its superior performance, exhibiting high accuracy than any other models . We chose the decision tree because it gives very accurate predictions, can handle complex patterns in data, and avoids overfitting. It works well with different types of data and allows us to see which features are most important. This makes it a reliable and effective model for our task