

Model Optimization and Tuning Phase Template

Date	9 July 2024
Team ID	SWTID1720104754
Project Title	Cereal Analysis Based On Rating By Using Machine Learning Techniques
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># Calculate evaluation metrics for each model models = ["Linear Regression", "Ridge Regression", "Lasso Regression", "Decision Tree Regressor", "Random Forest Regressor"] y_preds = [y_pred_lr, y_pred_r, y_pred_l, y_pred_dt, y_pred_rf] for model, y_pred in zip(models, y_preds): r2 = r2_score(y_test, y_pred) rmse = mean_squared_error(y_test, y_pred, squared=False) # Square root for interpretability mape = mean_absolute_percentage_error(y_test, y_pred) * 100 # Percentage error print(f"\nModel: {model}") print(f"R-squared: {r2:.4f}") print(f"Root Mean Squared Error (RMSE): {rmse:.4f}") print(f"Mean Absolute Percentage Error (MAPE): {mape:.4f}%")</pre>	Model: Linear Regression R-squared: 1.0000 Root Mean Squared Error (RMSE): 0.0000 Mean Absolute Percentage Error (MAPE): 0.0000%
RandomForest		Model: Ridge Regression R-squared: 0.9941 Root Mean Squared Error (RMSE): 1.1395 Mean Absolute Percentage Error (MAPE): 2.0762%
Linear Regression		Model: Lasso Regression R-squared: 1.0000 Root Mean Squared Error (RMSE): 0.0859 Mean Absolute Percentage Error (MAPE): 0.1717%
Ridge Regression		Model: Decision Tree Regressor R-squared: 0.7235 Root Mean Squared Error (RMSE): 7.7926 Mean Absolute Percentage Error (MAPE): 16.3354%
Lasso Regression		Model: Random Forest Regressor R-squared: 0.7841 Root Mean Squared Error (RMSE): 6.8870 Mean Absolute Percentage Error (MAPE): 16.9964%

Performance Metrics Comparison Report (2 Marks):

Model	Baseline Metric	Optimized Metric
Decision Tree	Model: Linear Regression R-squared: 1.0000 Root Mean Squared Error (RMSE): 0.0000 Mean Absolute Percentage Error (MAPE): 0.0000%	<div>lr.score(x_test,y_test)</div> <div>0.9999999999999994</div>
Random Forest	Model: Ridge Regression R-squared: 0.9941 Root Mean Squared Error (RMSE): 1.1395 Mean Absolute Percentage Error (MAPE): 2.0762%	<div>r.score(x_test,y_test)</div> <div>0.9940890789552553</div>
Linear Regression	Model: Lasso Regression R-squared: 1.0000 Root Mean Squared Error (RMSE): 0.0859 Mean Absolute Percentage Error (MAPE): 0.1717%	<div>l.score(x_test,y_test)</div> <div>0.9999663757215922</div>
Ridge Regression	Model: Decision Tree Regressor R-squared: 0.7235 Root Mean Squared Error (RMSE): 7.7926 Mean Absolute Percentage Error (MAPE): 16.3354%	<div>dt.score(x_test,y_test)</div> <div>0.7235434757748809</div>
Lasso Regression	Model: Random Forest Regressor R-squared: 0.7841 Root Mean Squared Error (RMSE): 6.8870 Mean Absolute Percentage Error (MAPE): 16.9964%	<div>rf.score(x_test,y_test)</div> <div>0.7840688820583703</div>

Final Model Selection Justification (2 Marks):

Final Model	Reasoning
Ridge Regression	While Linear and Lasso Regression show perfect fit scores on the given data, Ridge Regression might be the better model due to its potential for better generalization and robustness against overfitting. The slightly lower performance metrics in Ridge Regression indicate it maintains a good fit while reducing the risk of overfitting, which is crucial for making reliable predictions on new data.