

Model Optimization and Tuning Phase Report

Date	24 June 2025
Team ID	SWUID20250176345
Project Title	Machine Learning Approach for Employee Performance Prediction
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	<pre> 61 62 # Linear Regression 63 lr = LinearRegression() 64 lr.fit(X_train, y_train) 65 lr_preds = lr.predict(X_test) 66 lr_mae = mean_absolute_error(y_test, lr_preds) 67 lr_mse = mean_squared_error(y_test, lr_preds) 68 lr_r2 = r2_score(y_test, lr_preds) 69 print("\nLinear Regression:") 70 print(f"MAE: {lr_mae:.4f}") 71 print(f"MSE: {lr_mse:.4f}") 72 print(f"R² Score: {lr_r2:.4f}") </pre>	<p>Linear Regression:</p> <p>MAE: 0.04347336915994117</p> <p>MSE: 0.004133703766872829</p> <p>R² Score: 0.8012302992785618</p>
Random Forest	<pre> 74 # Random Forest 75 rf = RandomForestRegressor(random_state=42) 76 rf.fit(X_train, y_train) 77 rf_preds = rf.predict(X_test) 78 rf_mae = mean_absolute_error(y_test, rf_preds) 79 rf_mse = mean_squared_error(y_test, rf_preds) 80 rf_r2 = r2_score(y_test, rf_preds) 81 print("\nRandom Forest:") 82 print(f"MAE: {rf_mae:.4f}") 83 print(f"MSE: {rf_mse:.4f}") 84 print(f"R² Score: {rf_r2:.4f}") </pre>	<p>Random Forest:</p> <p>MAE: 0.03253990252410063</p> <p>MSE: 0.0035580692923302647</p> <p>R² Score: 0.8289097602855915</p>

Xgboost Model	<pre> 86 # XGBoost 87 xgb_model = xgb.XGBRegressor() 88 xgb_model.fit(X_train, y_train) 89 xgb_preds = xgb_model.predict(X_test) 90 xgb_mae = mean_absolute_error(y_test, xgb_preds) 91 xgb_mse = mean_squared_error(y_test, xgb_preds) 92 xgb_r2 = r2_score(y_test, xgb_preds) 93 print("\nXGBoost:") 94 print(f"MAE: {xgb_mae:.4f}") 95 print(f"MSE: {xgb_mse:.4f}") 96 print(f"R² Score: {xgb_r2:.4f}") 97 </pre>	<pre> XGBoost: MAE: 0.033858053386211395 MSE: 0.0037014407571405172 R² Score: 0.8220157623291016 </pre>
---------------	--	---

Performance Metrics Comparison Report (2 Marks):

Model	Optimized Metric
Random Forest	<pre> Best Model: Random Forest Model and feature order saved successfully. Process finished with exit code 0 </pre>

Final Model Selection Justification (2 Marks):

Final Model	Reasoning
Random Forest	<p>The Random Forest Regressor was selected as the final model due to its strong predictive performance, low mean absolute error, and robustness against overfitting. It effectively handled both numerical and categorical features without the need for extensive scaling or transformation. Its ability to manage complex, non-linear relationships in the employee productivity data made it a reliable and efficient choice aligned with the project's goals.</p>