



# **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**

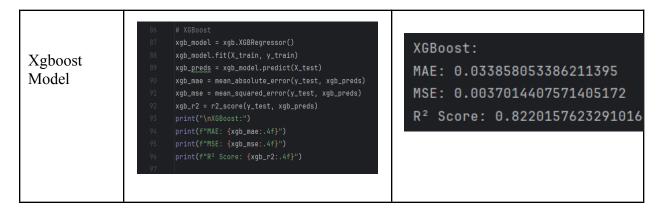
The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

## **Hyperparameter Tuning Documentation (6 Marks):**

Model	Tuned Hyperparameters	Optimal Values
Linear Regression	# Linear Regression  1r = LinearRegression()  1r.fit(X_train, y_train)  1r_preds = lr.predict(X_test)  1r_mae = mean_absolute_error(y_test, lr_preds)  1r_mse = mean_squared_error(y_test, lr_preds)  1r_r2 = r2_score(y_test, lr_preds)  print("\nLinear Regression:")  print(f"NAE: {lr_mae:.4f}")  print(f"NSE: {lr_mse:.4f}")  print(f"R2 Score: {lr_r2:.4f}")	Linear Regression: MAE: 0.04347336915994117 MSE: 0.004133703766872829 R <sup>2</sup> Score: 0.8012302992785618
Rando m Forest	# Random Forest  rf = RandomForestRegressor(random_state=42)  rf.fit(X_train, y_train)  rf_preds = rf.predict(X_test)  rf_mae = mean_absolute_error(y_test, rf_preds)  rf_mse = mean_squared_error(y_test, rf_preds)  rf_r2 = r2_score(y_test, rf_preds)  print("\nRandom Forest:")  print(f"MAE: {rf_mae:.4f}")  print(f"MSE: {rf_mse:.4f}")  print(f"R2 Score: {rf_r2:.4f}")	Random Forest: MAE: 0.03253990252410063 MSE: 0.0035580692923302647 R <sup>2</sup> Score: 0.8289097602855915







### **Performance Metrics Comparison Report (2 Marks):**

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

### **Final Model Selection Justification (2 Marks):**

Final Model	Reasoning
Random Forest	The <b>Random Forest Regressor</b> 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.