Model Optimization and Tuning Phase Report

Date	24 June 2025
Team ID	SWUID20250176341
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	```python
	# 1. Linear Regression Model
	model_lr = LinearRegression()
	model_lr.fit(X_train, y_train)
	pred_test = model_lr.predict(X_test)
	
	mae_lr = mean_absolute_error(y_test, pred_test)
	mse_lr = mean_squared_error(y_test, pred_test)
	r2_lr = r2_score(y_test, pred_test)
	
	print("\nLinear Regression:")
	print("MAE:", mae_lr)
	print("MSE:", mse_lr)
	print("R² Score:", r2_lr)
	``
		``
		Linear Regression:
		MAE: 0.04347336915994117
		MSE: 0.004133703766872829
		R² Score: 0.8012302992785618
		``

Random Forest	```python
	# 2. Random Forest Model
	model_rf = RandomForestRegressor(random_state=42)
	model_rf.fit(X_train, y_train)
	pred = model_rf.predict(X_test)
	
	mae_rf = mean_absolute_error(y_test, pred)
	mse_rf = mean_squared_error(y_test, pred)
	r2_rf = r2_score(y_test, pred)
	
	print("\nRandom Forest:")
	print("MAE:", mae_rf)
	print("MSE:", mse_rf)
	print("R² Score:", r2_rf)
	``
		``
		Random Forest:
		MAE: 0.03253990252410063
		MSE: 0.0035580692923302647
		R² Score: 0.8289097602855915
		``

| Xgboost Model | ```
# 3. XGBoost Model
model_xgb = XGBRegressor(random_state=42, verbosity=
model_xgb.fit(X_train, y_train)
pred3 = model_xgb.predict(X_test)

mae_xgb = mean_absolute_error(y_test, pred3)
mse_xgb = mean_squared_error(y_test, pred3)
r2_xgb = r2_score(y_test, pred3)

print("\nXGBoost:")
print("MAE:", mae_xgb)
print("MSE:", mse_xgb)
print("R² Score:", r2_xgb)
``` | ```
XGBoost:
MAE: 0.033858053386211395
MSE: 0.0037014407571405172
R² Score: 0.8220157623291016
``` |

Performance Metrics Comparison Report (2 Marks):

Model	Optimized Metric
| Random Forest | ```
----- Model Comparison (Based on Test Set) -----
Model              MAE        MSE         R² Score
Linear Regression  0.0435     0.0041      0.8012
Random Forest      0.0325     0.0036      0.8289
XGBoost            0.0339     0.0037      0.8220
✅ Random Forest model and feature order saved successfully.
F2 Score: 0.9046454767726161
``` |

Final Model Selection Justification (2 Marks):

Final Model	Reasoning
Random Forest	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.