

Model Development Phase Template

Date	25 June 2025
Team ID	SWUID20250176209
Project Title	Machine Learning Approach for Employee Performance Prediction
Maximum Marks	4 Marks

Initial Model Training Code, Model Validation and Evaluation Report

The initial model training code will be showcased in the future through a screenshot. The model validation and evaluation report will include classification reports, accuracy, and confusion matrices for multiple models, presented through respective screenshots.

Initial Model Training Code:

1. Linear Regression:

```
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score

# Initialize Linear Regression model:
model_lr = LinearRegression()

# Train the model:
model_lr.fit(X_train, y_train)

# Predict on the test set:
pred_test = model_lr.predict(X_test)

# Model Evaluation:
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)

# Results:
print("Linear Regression Evaluation:")
print(f"Mean Absolute Error (MAE): {mae_lr:.4f}")
```

2. Random Forest:

```
# Random Forest:
from sklearn.ensemble import RandomForestRegressor

# Initialize Random Forest model:
model_rf = RandomForestRegressor()

# Train the model:
model_rf.fit(X_train, y_train)

# Predict on the test set:
pred = model_rf.predict(X_test)

# Model Evaluation:
mae_rf = mean_absolute_error(y_test, pred)
mse_rf = mean_squared_error(y_test, pred)
r2_rf = r2_score(y_test, pred)

# Results:
print("Random Forest Evaluation:")
print(f"Mean Absolute Error (MAE): {mae_rf:.4f}")
print(f"Mean Squared Error (MSE): {mse_rf:.4f}")
print(f"R2 Score: {r2_rf:.4f}")
```

3. XGBoost:

```
# XGBoost:
from xgboost import XGBRegressor

# Initialize XGBoost Regressor:
model_xgb = XGBRegressor()

# Train the model:
model_xgb.fit(X_train, y_train)

# Predict on the test set:
pred3 = model_xgb.predict(X_test)

# Model Evaluation:
mae_xgb = mean_absolute_error(y_test, pred3)
mse_xgb = mean_squared_error(y_test, pred3)
r2_xgb = r2_score(y_test, pred3)

# Results:
print("XGBoost Evaluation:")
print(f"Mean Absolute Error (MAE): {mae_xgb:.4f}")
print(f"Mean Squared Error (MSE): {mse_xgb:.4f}")
print(f"R2 Score: {r2_xgb:.4f}")
```

Model Validation and Evaluation Report:

Model	Performance Summary	R ²	Prediction vs Actual Plot / Residual Plot
Linear Regression	<pre> Linear Regression Evaluation: Mean Absolute Error (MAE): 0.1040 Mean Squared Error (MSE): 0.0205 R² Score: 0.3025 </pre>	30%	<pre> ===== Model Performance Comparison: Model MAE MSE R² Score 0 Linear Regression 0.103975 0.020518 0.302503 1 Random Forest 0.068143 0.011757 0.600309 2 XGBoost 0.067346 0.010990 0.626399 ===== </pre>
Random Forest	<pre> ===== Random Forest Evaluation: Mean Absolute Error (MAE): 0.0681 Mean Squared Error (MSE): 0.0118 R² Score: 0.6003 ===== </pre>	60%	<pre> ===== Model Performance Comparison: Model MAE MSE R² Score 0 Linear Regression 0.103975 0.020518 0.302503 1 Random Forest 0.068143 0.011757 0.600309 2 XGBoost 0.067346 0.010990 0.626399 ===== </pre>
XGBoost	<pre> ===== XGBoost Evaluation: Mean Absolute Error (MAE): 0.0673 Mean Squared Error (MSE): 0.0110 R² Score: 0.6264 ===== </pre>	62%	<pre> ===== Model Performance Comparison: Model MAE MSE R² Score 0 Linear Regression 0.103975 0.020518 0.302503 1 Random Forest 0.068143 0.011757 0.600309 2 XGBoost 0.067346 0.010990 0.626399 ===== </pre>