Model Development Phase

Date	23 July 2025
ProjectName	Machine Learning Approach for Employee Performance Prediction
Maximum 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 Model

```
from sklearn.linear_model import LinearRegression
liR = LinearRegression()
liR_model= liR.fit(X_train, Y_train)
```

```
X_test.head()
```

Show hidden output

```
Y_pred = liR_model.predict(X_test)
Y_pred
```

```
r2 = r2_score(Y_test, Y_pred)
r2 # (0.290)
```

```
0.29075725756610216
```

2. Random Forest Model

```
from sklearn.ensemble import RandomForestRegressor
Rf = RandomForestRegressor(n_estimators=100, random_state=42)
Rf_model = Rf.fit(X_train, Y_train)
```

```
Y_pred = Rf_model.predict(X_test)
Y_pred
```

```
r2 = r2_score(Y_test, Y_pred)
r2 # (0.470)
```

```
0.4707626410260015
```

3. XgBoost Model

```
xgb = XGBRegressor(objective='reg:squarederror', n_estimators=250, max_depth=3,learning_rate=0.1, random_state=40)
xgb_model = xgb.fit(X_train, Y_train)
```

```
Y_pred = xgb_model.predict(X_test)
Y_pred
```

```
r2 = r2_score(Y_test, Y_pred)
r2 # (0.521)
```

```
0.5217457607580249
```

Model Validation and Evaluation Report:

Model	Mean Absolute Error	Mean Squared Error	R2 Score
Linear Regression	0.105	0.020	0.290
Random Forest	0.079	0.015	0.470
XGBoost	0.078	0.014	0.521