

## Model Development Phase Template

Date	07 July 2024
Team ID	team-739715
Project Title	House Rent Price Prediction Using Machine Learning
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:

##### Linear Regression Model

```

Linear Regression model

linReg = LinearRegression()
linReg.fit(x_train,y_train)

[ ] y_pred = linReg.predict(x_test)

[ ] accuracy = linReg.score(x_test,y_test)
print(accuracy)

0.8139527448447011
  
```

##### Random Forest Model

```

Random Forest Model

[ ] rf = RandomForestRegressor(n_estimators = 100 , random_state = 0)
rf.fit(x,y)

[ ] y_pred = rf.predict(x_test)

[ ] accuracy = rf.score(x_test,y_test)
print(accuracy)

0.9863832466567757
  
```

##### XGBoost Regression Model

```

XGBoost Regression

import xgboost
from xgboost import XGBRegressor
xgb_model = XGBRegressor()
xgb_model.fit(x_train, y_train)
pred_xgb = xgb_model.predict(x_test)
mae_xgb = mean_absolute_error(y_test, pred_xgb)
mse_xgb = mean_squared_error(y_test, pred_xgb)
rmse_xgb = np.sqrt(mse_xgb)
rsq_xgb = r2_score(y_test, pred_xgb)
print('MAE: %.3f' % mae_xgb)
print('MSE: %.3f' % mse_xgb)
print('RMSE: %.3f' % rmse_xgb)
print('R-Square: %.3f' % rsq_xgb)
print(accuracy)

MAE: 3266.968
MSE: 25973983.707
RMSE: 5096.468
R-Square: 0.917
0.9863832466567757

```

## Decision Tree Model

```

Decision Tree Model:

from sklearn.tree import DecisionTreeRegressor
dt = DecisionTreeRegressor(random_state = 0)
dt.fit(x,y)

DecisionTreeRegressor
DecisionTreeRegressor(random_state=0)

[ ] y_pred = dt.predict(x_test)

[ ] accuracy = dt.score(x_test,y_test)
print(accuracy)

0.9968193356037073

```

## Model Validation and Evaluation Report:

Model	Regression Report	Accuracy	Regression Matrix
Linear Regression	<pre> # Importing the libraries import numpy as np import pandas as pd from sklearn.linear_model import LinearRegression  # Splitting the dataset into the training set and the test set X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)  # Fitting the Linear Regression model to the training set regressor = LinearRegression() regressor.fit(X_train, y_train)  # Predicting the test set results y_pred = regressor.predict(X_test)  # Evaluating the model performance print('Mean Absolute Error: ', mean_absolute_error(y_test, y_pred)) print('Mean Squared Error: ', mean_squared_error(y_test, y_pred)) print('Root Mean Squared Error: ', np.sqrt(mean_squared_error(y_test, y_pred))) print('R-squared: ', r2_score(y_test, y_pred)) </pre>	81.3%	<pre> def model_compare(x_train, x_test, y_train, y_test):     # Importing the libraries     from sklearn.linear_model import LinearRegression     y_pred = LinearRegression().fit(x_train, y_train).predict(x_test)     print('Prediction Evaluation using Linear Regression')     print('Mean Absolute Error: ', mean_absolute_error(y_test, y_pred))     print('Mean Squared Error: ', mean_squared_error(y_test, y_pred))     print('Root Mean Squared Error: ', np.sqrt(mean_squared_error(y_test, y_pred)))     print('R-squared: ', r2_score(y_test, y_pred))  # Similarly for other models (rf, xgb_model), use their predict methods # ...  model_compare(x_train, x_test, y_train, y_test)  Prediction Evaluation using Linear Regression Mean Absolute Error: 341.87684615 Mean Squared Error: 5808104.81221 Root Mean Squared Error: 2409.920379 R-squared: 0.81302744847011 0.813 </pre>
Random Forest Regressor	<pre> # Importing the libraries import numpy as np import pandas as pd from sklearn.ensemble import RandomForestRegressor  # Splitting the dataset into the training set and the test set X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)  # Fitting the Random Forest Regressor model to the training set regressor = RandomForestRegressor() regressor.fit(X_train, y_train)  # Predicting the test set results y_pred = regressor.predict(X_test)  # Evaluating the model performance print('Mean Absolute Error: ', mean_absolute_error(y_test, y_pred)) print('Mean Squared Error: ', mean_squared_error(y_test, y_pred)) print('Root Mean Squared Error: ', np.sqrt(mean_squared_error(y_test, y_pred))) print('R-squared: ', r2_score(y_test, y_pred)) </pre>	98.6%	<pre> def model_compare(x_train, x_test, y_train, y_test):     # Importing the libraries     from sklearn.ensemble import RandomForestRegressor     y_pred = RandomForestRegressor().fit(x_train, y_train).predict(x_test)     print('Prediction Evaluation using Random Forest Regressor')     print('Mean Absolute Error: ', mean_absolute_error(y_test, y_pred))     print('Mean Squared Error: ', mean_squared_error(y_test, y_pred))     print('Root Mean Squared Error: ', np.sqrt(mean_squared_error(y_test, y_pred)))     print('R-squared: ', r2_score(y_test, y_pred))  # Similarly for other models (rf, xgb_model), use their predict methods # ...  model_compare(x_train, x_test, y_train, y_test)  Prediction Evaluation using Random Forest Regressor Mean Absolute Error: 124.3613820803 Mean Squared Error: 4247661.86975192 Root Mean Squared Error: 2060.81847211805 R-squared: 0.9863832466567757 0.986 </pre>
XGBoost Regression	<pre> # Importing the libraries import numpy as np import pandas as pd from sklearn.xgboost import XGBRegressor  # Splitting the dataset into the training set and the test set X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)  # Fitting the XGBoost Regressor model to the training set regressor = XGBRegressor() regressor.fit(X_train, y_train)  # Predicting the test set results y_pred = regressor.predict(X_test)  # Evaluating the model performance print('Mean Absolute Error: ', mean_absolute_error(y_test, y_pred)) print('Mean Squared Error: ', mean_squared_error(y_test, y_pred)) print('Root Mean Squared Error: ', np.sqrt(mean_squared_error(y_test, y_pred))) print('R-squared: ', r2_score(y_test, y_pred)) </pre>	91.6%	<pre> def model_compare(x_train, x_test, y_train, y_test):     # Importing the libraries     from sklearn.xgboost import XGBRegressor     y_pred = XGBRegressor().fit(x_train, y_train).predict(x_test)     print('Prediction Evaluation using Gradient Boosting Regressor')     print('Mean Absolute Error: ', mean_absolute_error(y_test, y_pred))     print('Mean Squared Error: ', mean_squared_error(y_test, y_pred))     print('Root Mean Squared Error: ', np.sqrt(mean_squared_error(y_test, y_pred)))     print('R-squared: ', r2_score(y_test, y_pred))  # Similarly for other models (rf, xgb_model), use their predict methods # ...  model_compare(x_train, x_test, y_train, y_test)  Prediction Evaluation using Gradient Boosting Regressor Mean Absolute Error: 200.26042001076 Mean Squared Error: 1587103.70711517 Root Mean Squared Error: 1260.16707091144 R-squared: 0.916222890949421 0.916 </pre>

Decision Tree	<pre> 1 # Importing the libraries to be used in the programme and to a series of arrays. 2 # Importing the datasets 3 # Importing the train and test datasets 4 # Splitting the dataset into the training set and the testing set 5 # Training the Decision Tree Classifier 6 # Predicting the results 7 # Evaluating the model 8 # Saving the model </pre>	99.6%	<pre> 1 # Importing the libraries to be used in the programme and to a series of arrays. 2 # Importing the datasets 3 # Importing the train and test datasets 4 # Splitting the dataset into the training set and the testing set 5 # Training the Decision Tree Classifier 6 # Predicting the results 7 # Evaluating the model 8 # Saving the model </pre>
---------------	---	-------	---