

machinelearning_final_exam.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

RAM Disk

Comment Share

Name: Nowshad Mahmud

ID: 20228054

[14] import numpy as npimport pandas as pdfrom sklearn.model_selection import train_test_splitfrom sklearn.linear_model import LinearRegression

[15] # Load the California Housing datasetfrom sklearn.datasets import fetch_california_housingcalifornia_housing = fetch_california_housing()X = pd.DataFrame(california_housing.data, columns=california_housing.feature_names)y = pd.Series(california_housing.target, name='MedHouseVal')

[16] X

	MedInc	HouseAge	AveRooms	AveBedrms	Population	AveOccup	Latitude	Longitude
0	8.3252	41.0	6.984127	1.023810	322.0	2.555556	37.88	-122.23
1	8.3014	21.0	6.238137	0.971880	2401.0	2.109842	37.86	-122.22
2	7.2574	52.0	8.288136	1.073446	496.0	2.802260	37.85	-122.24
3	5.6431	52.0	5.817352	1.073059	558.0	2.547945	37.85	-122.25
4	3.8462	52.0	6.281853	1.081081	565.0	2.181467	37.85	-122.25
...
20635	1.5603	25.0	5.045455	1.133333	845.0	2.560606	39.48	-121.09
20636	2.5568	18.0	6.114035	1.315789	356.0	3.122807	39.49	-121.21
20637	1.7000	17.0	5.205543	1.120092	1007.0	2.325635	39.43	-121.22
20638	1.8672	18.0	5.329513	1.171920	741.0	2.123209	39.43	-121.32
20639	2.3886	16.0	5.254717	1.162264	1387.0	2.616981	39.37	-121.24

20640 rows × 8 columns

[17] y

0	4.526
1	3.585
2	3.521
3	3.413
4	3.422
...	...
20635	0.781
20636	0.771
20637	0.923
20638	0.847
20639	0.894

Name: MedHouseVal, Length: 20640, dtype: float64

[18] # Split the dataset into training and testing setsX_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

[19] # Train a linear regression modelregressor = LinearRegression()regressor.fit(X_train, y_train)

LinearRegression

LinearRegression()

[20] # Make predictions on the test sety_pred = regressor.predict(X_test)

[21] from sklearn.metrics import mean_squared_error, mean_absolute_error, r2_score

Evaluate the model

mse = mean_squared_error(y_test, y_pred)rmse = np.sqrt(mse)print("Mean Squared Error (MSE): {:.4f}".format(mse))print("Root Mean Squared Error (RMSE): {:.4f}".format(rmse))

Mean Squared Error (MSE): 0.5559Root Mean Squared Error (RMSE): 0.7456

we first load the California Housing dataset using the fetch_california_housing() function from scikit-learn. We then split the dataset into training and testing sets using train_test_split() function, with 80% of the data used for training and 20% for testing. Next, we train a linear regression model using LinearRegression() class from scikit-learn. Finally, we evaluate the model by calculating the mean squared error (MSE) and root mean squared error (RMSE) between the predicted and actual house prices on the test set

Colab paid products - Cancel contracts here

0s completed at 14:59