

NUTAN COLLEGE OF ENGINEERING & RESEARCH (NCER)



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING - ARTIFICIAL INTELLIGENCE

Experiment No: 05

Code 01:import pandas as pd import numpy as np import matplotlib.pyplot as plt import seaborn as sns # Import warnings import warnings warnings.filterwarnings("ignore") # We will use some methods from the sklearn module from sklearn import linear model from sklearn.linear model import LinearRegression from sklearn import metrics from sklearn.metrics import mean squared error, mean absolute error from sklearn.model selection import train test split, cross val score # Load the dataset df = pd.read csv("Student Performance.csv") print(df.corr(numeric only=True)) # Use numeric only=True to only get numeric correlations print(df.describe()) # Setting the value for X and y

X = df[['Hours Studied', 'Previous Scores']] # Adjusted features

NUTAN COLLEGE OF PASSED BASED BASED

NUTAN MAHARASHTRA VIDYA PRASARAK MANDAL'S

NUTAN COLLEGE OF ENGINEERING & RESEARCH (NCER)



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING - ARTIFICIAL INTELLIGENCE

y = df['Performance Index'] # Adjusted target variable # Visualizing the features fig, axs = plt.subplots(2, figsize=(5, 10)) # Adjusted height for better spacing plt1 = sns.boxplot(x=df['Hours Studied'], ax=axs[0]) # Adjusted for the new feature plt2 = sns.boxplot(x=df['Previous Scores'], ax=axs[1]) # Adjusted for the new feature plt.tight layout() sns.histplot(df]'Performance Index'], kde=True) # Changed to histplot with kde for better visualization sns.pairplot(df, x vars=['Hours Studied', 'Previous Scores'], y vars='Performance Index', height=4, aspect=1, kind='scatter') plt.show() # Create the correlation matrix and represent it as a heatmap. sns.heatmap(df.corr(numeric only=True), annot=True, cmap='coolwarm') # Use numeric only=True for heatmap plt.show()

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=100)

Fitting the Multiple Linear Regression model

reg_model = LinearRegression().fit(X_train, y_train)

Printing the model coefficients

print('Intercept: ', reg_model.intercept_)



NUTAN COLLEGE OF ENGINEERING & RESEARCH (NCER)



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING - ARTIFICIAL INTELLIGENCE

Pair the feature names with the coefficients
print(list(zip(X.columns, reg_model.coef_))) # Corrected to use .columns for feature names
Predicting the Test and Train set result
<pre>y_pred = reg_model.predict(X_test)</pre>
x_pred = reg_model.predict(X_train) # This is actually predicting on the train set
<pre>print("Prediction for test set: {}".format(y_pred))</pre>
Actual value and the predicted value
reg_model_diff = pd.DataFrame({'Actual value': y_test, 'Predicted value': y_pred})
print(reg_model_diff) # Added print to display the DataFrame
mae = metrics.mean_absolute_error(y_test, y_pred)
mse = metrics.mean_squared_error(y_test, y_pred)
r2 = np.sqrt(mse) # Changed to use mse instead of mean_squared_error directly
print('Mean Absolute Error:', mae)
print('Mean Square Error:', mse)
print('Root Mean Square Error:', r2)

Code 01 Output:-



NUTAN COLLEGE OF ENGINEERING & RESEARCH (NCER)

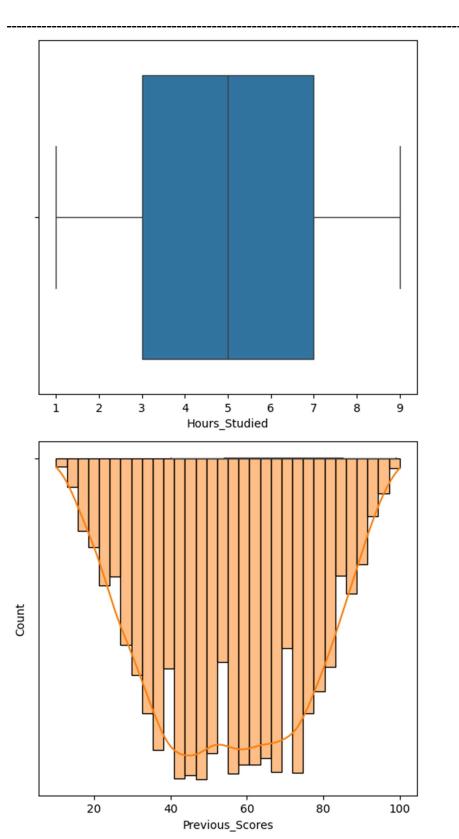


Hours Studied Previous Scores New Hours Ne	IX.	7/:				
Hours Studied T. 000000			Hours Studied	Previous Scores	Sleep Hours	1
Sleep Hours	Hours_Stu	udied	1.000000		0.001245	
Sample Question Papers Practiced Performance Index Sample Question Papers Practiced No. 17463						
Sample Question Papers Practiced Performance Index Sample Question Papers Practiced O.017463 O.003990 O.0043268						
Sample Question Papers Practiced No.017463 Previous Scores 0.007888 Sleep Hours Sample Question Papers Practiced 1.0000000 No.043268 Performance Index No.043268 No.04268 No						
Hours Studied Previous Scores Performance Index	Periorman	ice_index	0.3/3/30	0.915189	0.048100	
Previous Scores 0.007888 Sleep Hours 0.003999 Sample Question Papers Practiced Performance Index Performance Index O.043268			Sample_Quest:	on_Papers_Practice	/ b	
Sample_Question_Papers_Practiced Performance_Index						
None						
Performance_Index			esticad			
Performance Index			icticed			
Hours_Studied 0.373730 Previous_Scores 0.915189 Sleep_Hours 0.048106 Sample_Question_Papers_Practiced 0.943268 Performance_Index 1.000000 Hours_Studied Previous_Scores Sleep_Hours count 10000.000000 10000.000000 10000.000000 mean 4.992900 69.445700 6.530600 std 2.589309 17.343152 1.695863 min 1.000000 40.000000 4.000000 25% 3.000000 54.000000 5.000000 50% 5.000000 69.000000 7.000000 75% 7.000000 85.000000 8.000000 max 9.000000 99.000000 9.000000 std 2.867348 19.212558 min 0.000000 10.000000 std 2.867348 19.212558 min 0.000000 40.000000 25% 2.000000 40.000000 25% 2.000000 55.000000 50% 5.000000 55.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000	Terror	icc_Index		0.04320		
Hours_Studied 0.373730 Previous_Scores 0.915189 Sleep_Hours 0.048106 Sample_Question_Papers_Practiced 0.943268 Performance_Index 1.000000 Hours_Studied Previous_Scores Sleep_Hours count 10000.000000 10000.000000 10000.000000 mean 4.992900 69.445700 6.530600 std 2.589309 17.343152 1.695863 min 1.000000 40.000000 4.000000 25% 3.000000 54.000000 5.000000 50% 5.000000 69.000000 7.000000 75% 7.000000 85.000000 8.000000 max 9.000000 99.000000 9.000000 std 2.867348 19.212558 min 0.000000 10.000000 std 2.867348 19.212558 min 0.000000 40.000000 25% 2.000000 40.000000 25% 2.000000 55.000000 50% 5.000000 55.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000 75% 7.000000 71.000000			Perf	ormance Index		
Previous_Scores 0.915189 Sleep_Hours 0.048106 Sample_Question_Papers_Practiced 0.043268 Performance_Index 1.000000 Hours_Studied Previous_Scores Sleep_Hours count 10000.000000 10000.000000 10000.000000 mean 4.992900 69.445700 6.530600 std 2.589309 17.343152 1.695863 min 1.000000 40.000000 4.000000 25% 3.000000 54.000000 5.000000 50% 5.000000 69.000000 7.000000 75% 7.000000 85.000000 8.000000 max 9.000000 99.000000 10000.00000 mean 4.583300 55.224800 std 2.867348 19.212558 min 0.000000 10.000000 10.000000 25% 2.000000 40.000000 55.000000 50% 5.000000 50.000000 55.000000 75% 7.000000 71.000000 71.000000	Hours S	Studied				
Sleep_Hours	and the second s					
Sample_Question_Papers_Practiced 0.043268 Performance_Index 1.000000 Hours_Studied Previous_Scores Sleep_Hours count 10000.000000 10000.000000 mean 4.992900 69.445700 6.530600 std 2.589309 17.343152 1.695863 min 1.000000 40.000000 4.000000 25% 3.000000 54.000000 5.000000 50% 5.000000 69.000000 7.000000 75% 7.000000 85.000000 9.000000 max 9.000000 99.000000 10000.00000 std 2.867348 19.212558 min 0.000000 40.00000 25% 2.000000 40.00000 50% 5.000000 55.00000 50% 5.000000 71.000000						
Performance Index			s Practiced			
Hours Studied Previous Scores Sleep Hours \ count 10000.000000 10000.000000 10000.0000000 mean 4.992900 69.445700 6.530600 std 2.589309 17.343152 1.695863 min 1.000000 40.000000 5.000000 50% 5.000000 54.000000 7.000000 50% 5.000000 85.000000 7.000000 max 9.000000 99.000000 99.000000 99.000000 Score Std 2.867348 19.212558 min 0.000000 10.000000 10.000000 50% 2.000000 25% 2.000000 10.000000 10.000000 50% 5.000000 55.000000 55.000000 55% 5.000000 55.000000 55.000000 55% 5.000000 55.0000000 55.0000000 55.						
count 10000.000000 10000.000000 10000.000000 mean 4.992900 69.445700 6.530600 std 2.589309 17.343152 1.695863 min 1.000000 40.000000 4.000000 25% 3.000000 54.000000 5.000000 50% 5.000000 69.000000 7.000000 75% 7.000000 85.000000 8.000000 max 9.000000 99.000000 Performance_Index count 10000.000000 10000.000000 mean 4.583300 55.224800 std 2.867348 19.212558 min 0.000000 10.00000 25% 2.000000 40.00000 50% 5.000000 55.00000 75% 7.000000 71.000000	1 01 1011		Previous Scores			
mean 4.992900 69.445700 6.530600 std 2.589309 17.343152 1.695863 min 1.000000 40.000000 4.000000 25% 3.000000 54.000000 5.000000 50% 5.000000 69.000000 7.000000 75% 7.000000 85.000000 8.000000 max 9.000000 99.000000 Performance Index count 10000.000000 10000.000000 mean 4.583300 55.224800 std 2.867348 19.212558 min 0.000000 10.000000 25% 2.000000 40.000000 50% 5.000000 55.000000 75% 7.000000 71.000000	count					
std 2.589309 17.343152 1.695863 min 1.000000 40.000000 4.000000 25% 3.000000 54.000000 5.000000 50% 5.000000 69.000000 7.000000 75% 7.000000 85.000000 8.000000 max 9.000000 99.000000 9.000000 std 10000.000000 10000.000000 std 2.867348 19.212558 min 0.000000 10.000000 25% 2.000000 40.000000 50% 5.000000 55.000000 75% 7.000000 71.000000						
min 1.000000 40.000000 4.000000 25% 3.000000 54.000000 5.000000 50% 5.000000 69.000000 7.000000 75% 7.000000 85.000000 8.000000 max 9.000000 99.000000 9.000000 Sample Question Papers Practiced Performance Index 10000.000000 count mean 4.583300 55.224800 std 2.867348 19.212558 min 0.000000 10.000000 25% 2.000000 40.000000 50% 5.000000 55.000000 75% 7.000000 71.000000	100000000000000000000000000000000000000					
25% 3.000000 54.000000 5.000000 5.000000 50% 5.000000 69.000000 7.000000 7.000000 7.000000 85.000000 8.000000 8.000000 9.000000 9.000000 9.000000 9.000000 9.000000 9.000000 10000.000000 10000.000000 std 2.867348 19.212558 min 0.000000 10.000000 10.000000 25% 2.000000 40.000000 50% 5.000000 55.000000 71.000000 71.000000 71.000000						
50% 5.000000 69.000000 7.000000 75% 7.000000 85.000000 8.000000 max 9.000000 99.000000 9.000000 Sample_Question_Papers_Practiced Performance_Index 10000.000000 10000.000000 mean 4.583300 55.224800 std 2.867348 19.212558 min 0.000000 10.000000 25% 2.000000 40.000000 50% 5.000000 55.000000 75% 7.000000 71.000000						
75% 7.000000 85.000000 8.000000 9.000000						
max 9.000000 99.000000 9.000000 Sample_Question_Papers_Practiced count Performance_Index 10000.000000 10000.000000 mean 4.583300 55.224800 std 2.867348 19.212558 min 0.000000 10.000000 25% 2.000000 40.000000 50% 5.000000 55.000000 75% 7.000000 71.000000						
Sample_Question_Papers_Practiced count	. 515355					
count 10000.000000 10000.000000 mean 4.583300 55.224800 std 2.867348 19.212558 min 0.000000 10.000000 25% 2.000000 40.000000 50% 5.000000 55.000000 75% 7.000000 71.000000	max	3.00000	33.00000	3.00000		
mean 4.583300 55.224800 std 2.867348 19.212558 min 0.000000 10.000000 25% 2.000000 40.000000 50% 5.000000 55.000000 75% 7.000000 71.000000	S	Sample Question	Papers Practiced	Performance_Inc	dex	
std 2.867348 19.212558 min 0.000000 10.000000 25% 2.000000 40.000000 50% 5.000000 55.000000 75% 7.000000 71.000000	count	W 70 III	10000.000000	10000.0000	900	
min 0.000000 10.000000 25% 2.000000 40.000000 50% 5.000000 55.000000 71.000000	mean		4.583300	55.2248	300	
25% 2.000000 40.000000 50% 5.000000 55.000000 75% 7.000000 71.000000	std		2.867348	19.212	558	
50% 5.000000 55.000000 75% 7.000000 71.000000	min		0.00000	10.000	900	
75% 7.000000 71.000000	25%		2.000000	40.000	900	
	50%		5.000000	55.000	900	
max 9.000000 100.000000	75%		7.000000	71.000	900	
	max		9.000000	100.000	900	



NUTAN COLLEGE OF ENGINEERING & RESEARCH (NCER)

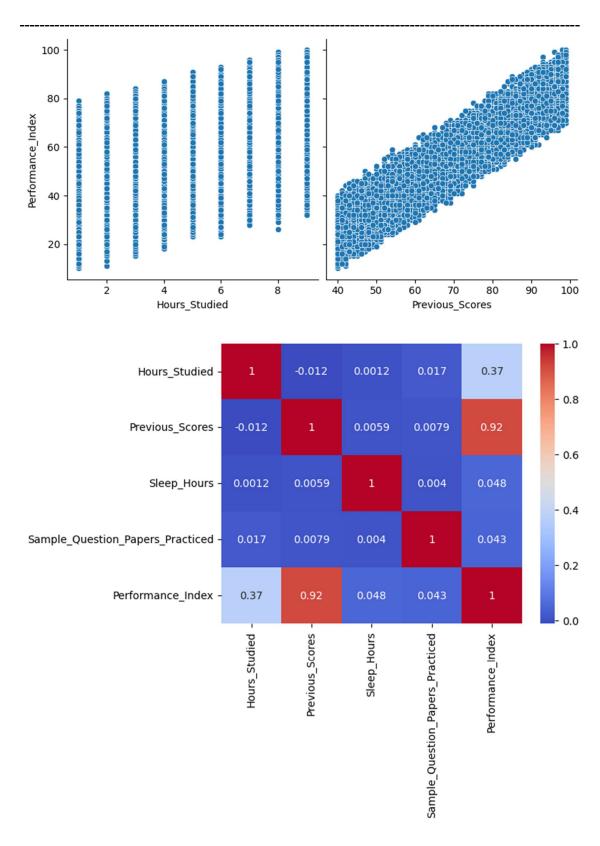






NUTAN COLLEGE OF ENGINEERING & RESEARCH (NCER)







NUTAN COLLEGE OF ENGINEERING & RESEARCH (NCER)



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING - ARTIFICIAL INTELLIGENCE

Intercept: -29.759696031285607

[('Hours_Studied', np.float64(2.8458785973597283)), ('Previous_Scores', np.float64(1.0192750395624741))]

Prediction for test set: [65.62826464 49.70329949 34.41417389 ... 71.9558614 35.22150241

14.87645919]

	Actual va	alue	Predicted value
8018		65	65.628265
9225		51	49.703299
3854		35	34.414174
2029		97	92.896287
3539		87	84.530140
4208		69	67.242922
5335		60	58.705286
2554		73	71.955861
13		33	35.221502
4468		15	14.876459

[3000 rows x 2 columns]

Mean Absolute Error: 1.8372246647574142

Mean Square Error: 5.3409802420571

Root Mean Square Error: 2.3110560880379127

Code 02:-

Import necessary libraries

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn import linear model

from sklearn.model selection import train test split

from sklearn import metrics

Load the dataset

df = pd.read csv("Student Performance.csv")

Prepare the features and target variable

X = df[['Hours Studied', 'Previous Scores']] # Adjusted features



NUTAN COLLEGE OF ENGINEERING & RESEARCH (NCER)



```
y = df['Performance Index'] # Adjusted target variable
# Perform exploratory data analysis
print(df.corr(numeric only=True)) # Only get numeric correlations
sns.pairplot(df, x vars=['Hours Studied', 'Previous Scores'], y vars='Performance Index',
height=4, aspect=1, kind='scatter')
plt.show()
# Split the data into training and testing sets
X train, X test, y train, y test = train test split(X, y, test size=0.3, random state=42)
# Create and fit the multivariable regression model
reg model = linear model.LinearRegression()
reg_model.fit(X_train, y_train)
# Print model coefficients
print('Intercept:', reg model.intercept )
print('Coefficients:', reg model.coef )
# Make predictions
y pred = reg model.predict(X test)
# Evaluate the model
mae = metrics.mean_absolute_error(y_test, y_pred)
mse = metrics.mean squared error(y test, y pred)
r2 = metrics.r2_score(y_test, y_pred)
```

NUTAN COLLEGE OF IMMEDIATES AND RESEARCH SHOWN AND

NUTAN MAHARASHTRA VIDYA PRASARAK MANDAL'S

NUTAN COLLEGE OF ENGINEERING & RESEARCH (NCER)



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING - ARTIFICIAL INTELLIGENCE

```
print('Mean Absolute Error:', mae)

print('Mean Squared Error:', mse)

print('R-squared:', r2)

# Visualize actual vs predicted values

plt.scatter(y_test, y_pred)

plt.xlabel('Actual Performance Index')

plt.ylabel('Predicted Performance Index')

plt.title('Actual vs Predicted Performance Index')

plt.plot([y.min(), y.max()], [y.min(), y.max()], 'r--') # 45-degree line

plt.show()
```

Code 02 Output: -

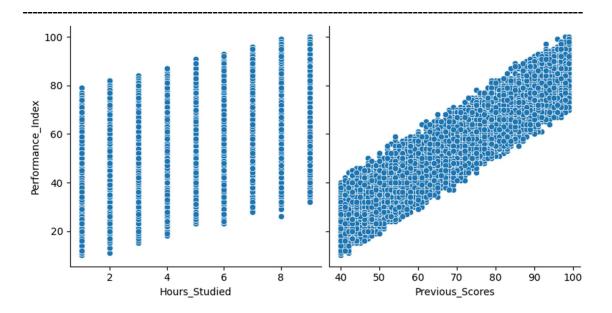
	Hours Studied	Previous Scores	Sleep Hours	١
Hours Studied	1.000000	-0.012390	0.001245	
Previous Scores	-0.012390	1.000000	0.005944	
Sleep Hours	0.001245	0.005944	1.000000	
Sample Question Papers Practiced	0.017463	0.007888	0.003990	
Performance_Index	0.373730	0.915189	0.048106	
	Sample Questio	n Papers Practice	d \	
Hours Studied	1.5	0.01746	3	
Previous Scores		0.00788	8	
Sleep Hours		0.00399	Θ	
Sample Question Papers Practiced		1.00000	Θ	
Performance_Index		0.04326	8	
	Performance In	dex		
Hours Studied	0.373	730		
Previous Scores	0.915	189		
Sleep Hours	0.048	106		
Sample Question Papers Practiced	0.043	268		
Performance Index	1.000	000		



NUTAN COLLEGE OF ENGINEERING & RESEARCH (NCER)



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING - ARTIFICIAL INTELLIGENCE



Intercept: -29.60536090045995

Coefficients: [2.86449217 1.01615225] Mean Absolute Error: 1.8101437492369952 Mean Squared Error: 5.201346140631689

R-squared: 0.9859445764337783

