

## **\*\*VAC ASSIGNMENT\*\***

**\*\*VANKAYALA JAHNAVI\*\***

**\*\*RA1911042020004\*\***

**\*\*CSBS\*\***

AIM:- To implement the student scores using linear regression **with** machine learning

*#import the libraries*

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

*#reading the dataset*

```
dataset = pd.read_csv('student_scores.csv')
```

```
dataset.shape
```

```
(25, 2)
```

```
dataset.head()
```

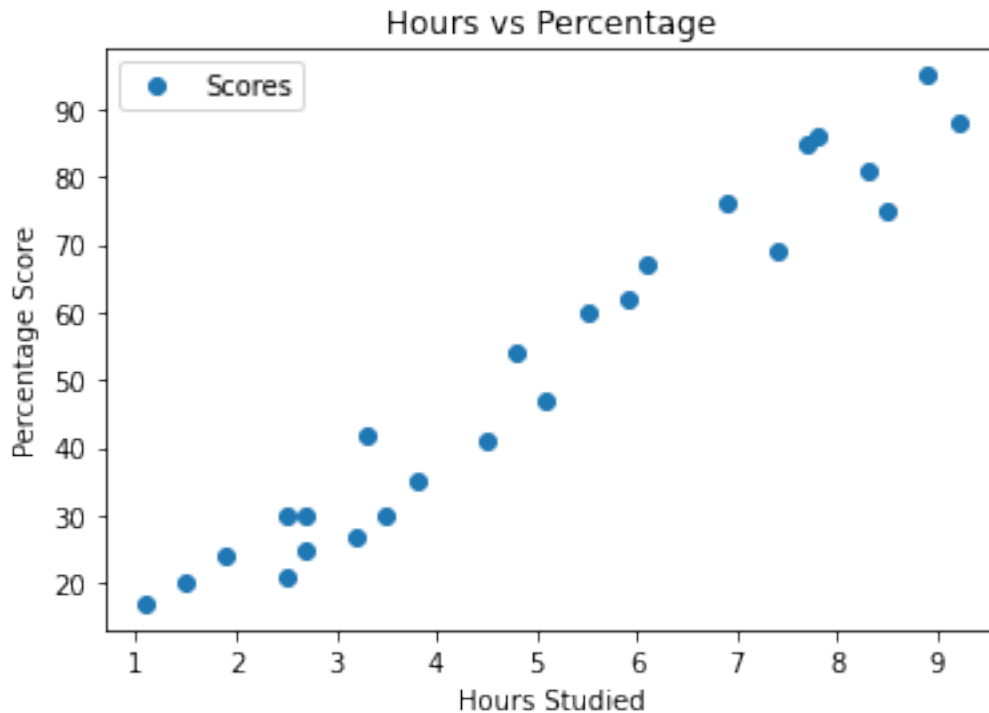
	Hours	Scores
0	2.5	21
1	5.1	47
2	3.2	27
3	8.5	75
4	3.5	30

```
dataset.describe()
```

	Hours	Scores
count	25.000000	25.000000
mean	5.012000	51.480000
std	2.525094	25.286887
min	1.100000	17.000000
25%	2.700000	30.000000
50%	4.800000	47.000000
75%	7.400000	75.000000
max	9.200000	95.000000

*#exploring the data scatter*

```
dataset.plot(x='Hours', y='Scores', style='o')
plt.title('Hours vs Percentage')
plt.xlabel('Hours Studied')
plt.ylabel('Percentage Score')
plt.show()
```



```
#data cleaning/preparation
X = dataset.iloc[:, :-1].values
y = dataset.iloc[:, 1].values
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=0)

#training the machine learning model
from sklearn.linear_model import LinearRegression
regressor = LinearRegression()
regressor.fit(X_train, y_train)

LinearRegression()

#exploring the result
print(regressor.intercept_)

2.018160041434662

print(regressor.coef_)

[9.91065648]

y_pred = regressor.predict(X_test)
df = pd.DataFrame({'Actual': y_test, 'Predicted': y_pred})
df
```

	Actual	Predicted
0	20	16.884145

1	27	33.732261
2	69	75.357018
3	30	26.794801
4	62	60.491033

*#evaluting the algorithm*

```
from sklearn import metrics
print('Mean Absolute Error:', metrics.mean_absolute_error(y_test,
y_pred))
print('Mean Squared Error:', metrics.mean_squared_error(y_test,
y_pred))
print('Root Mean Squared Error:',
np.sqrt(metrics.mean_squared_error(y_test, y_pred)))
```

Mean Absolute Error: 4.183859899002982

Mean Squared Error: 21.598769307217456

Root Mean Squared Error: 4.647447612100373

RESULT:- Implementation of student scores using linear regression was executed successfully.