

Task 1 : Prediction Using Supervised ML

Predict the percentage of student based on the number of study hours this is simple linear regression task that involves just two variables. Also predict the percentage of student if he/she studies 9.25hours/day.

Author : Mansi Rahul Shah

Importing required libraries

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

Importing Dataset

```
In [8]: data = "http://bit.ly/w-data"
dataset = pd.read_csv(data)
```

```
In [9]: print(dataset)
```

	Hours	Scores
0	2.5	21
1	5.1	47
2	3.2	27
3	8.5	75
4	3.5	30
5	1.5	20
6	9.2	88
7	5.5	60
8	8.3	81
9	2.7	25
10	7.7	85
11	5.9	62
12	4.5	41
13	3.3	42
14	1.1	17
15	8.9	95
16	2.5	30
17	1.9	24
18	6.1	67
19	7.4	69
20	2.7	30
21	4.8	54
22	3.8	35
23	6.9	76
24	7.8	86

```
In [10]: 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

Spilitting the data

```
In [14]: x = dataset.iloc[:, :-1].values
y = dataset.iloc[:, 1].values
print(x)
```

```
[[2.5]
[5.1]
[3.2]
[8.5]
[3.5]
[1.5]
[9.2]
[5.5]
[8.3]
[2.7]
[7.7]
[5.9]
[4.5]
[3.3]
[1.1]
[8.9]
[2.5]
[1.9]
[6.1]
[7.4]
[2.7]
[4.8]
[3.8]
[6.9]
[7.8]]
```

```
In [15]: print(y)
```

```
[21 47 27 75 30 20 88 60 81 25 85 62 41 42 17 95 30 24 67 69 30 54 35 76
86]
```

Applying Train test split

```
In [18]: 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)
```

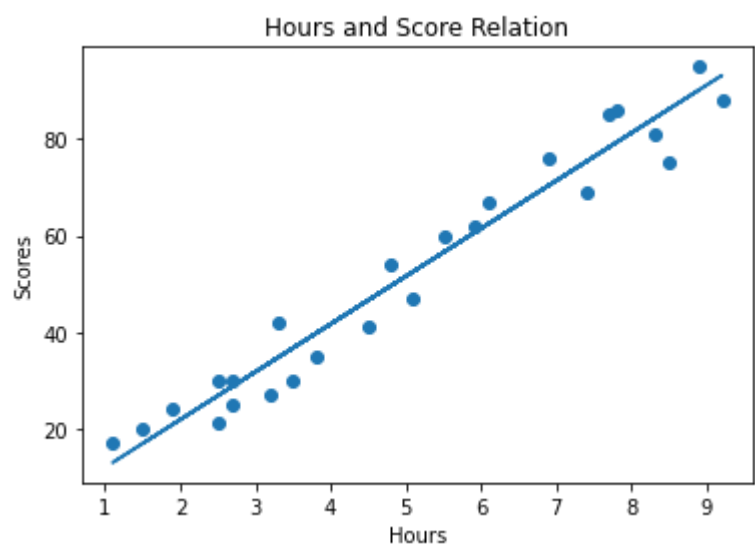
```
In [19]: from sklearn.linear_model import LinearRegression
regressor = LinearRegression()
regressor.fit(x_train, y_train)

print("Training complete.")
```

Training complete.

Visualization

```
In [21]: line = regressor.coef_*x+regressor.intercept_
plt.scatter(x, y)
plt.title("Hours and Score Relation")
plt.xlabel("Hours")
plt.ylabel("Scores")
plt.plot(x, line);
plt.show()
```



Prediction

If student studies for 9.25 hours/day

```
In [24]: t = 9.25
prediction = regressor.predict([[t]])
print("Hours studied per day :", t)
print("Predicted Score :", prediction)
```

Hours studied per day : 9.25  
Predicted Score : [93.69173249]

Thank-You