```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
```

```
_dataset=pd.read_csv('/content/student_scores - student_scores.csv')
```

```
_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.tail()
```

	Hours	Scores
20	2.7	30
21	4.8	54
22	3.8	35


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

```
[[2.5]
 [5.1]
 [3.2]
 [8.5]
 [3.5]
 [1.5]
 [9.2]
 [5.5]
 [8.3]
 [2.7]
```

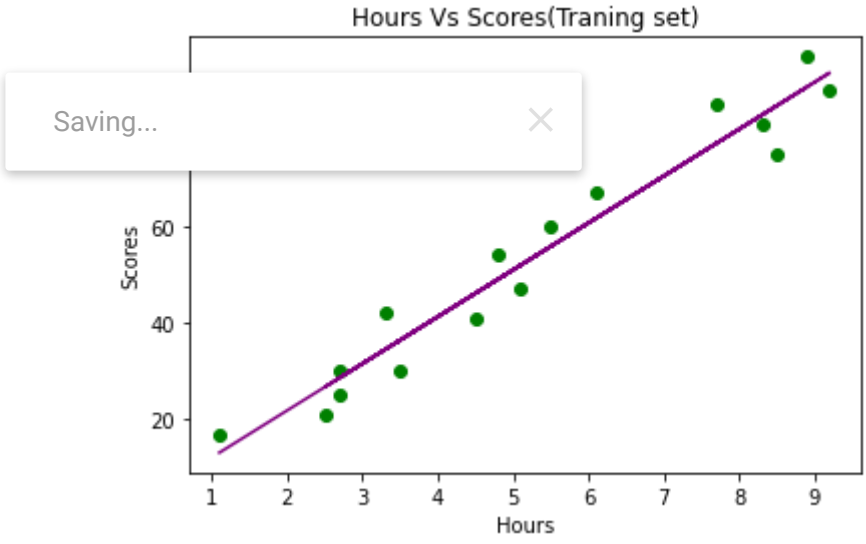
```
       [7.7]
       [5.9]
       [4.5]
       [3.3]
       [1.1]
       [8.9]
       [2.5]
       [1.9]
       [6.1]
       [7.4]
       [2.7]
       [4.8]
       [3.8]
       [6.9]
       [7.8]]
      [21 47 27 75 30 20 88 60 81 25 85 62 41 42 17 95 30 24 67 69 30 54 35 76
       86]
```

```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=1/3,random_state=0)
```

```
regressor=LinearRegression()
regressor.fit(x_train,y_train)
```

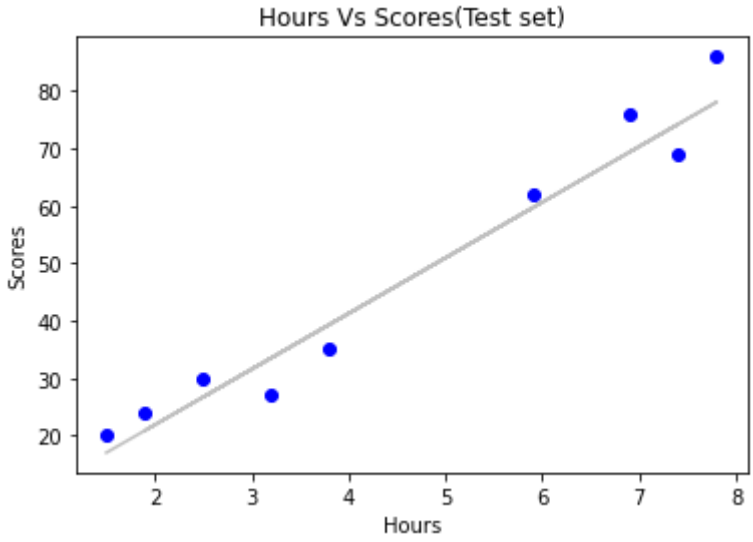
```
     LinearRegression()
```

```
y_pred=regressor.predict(x_test)
plt.scatter(x_train,y_train,color='green')
plt.plot(x_train,regressor.predict(x_train),color='purple')
plt.title("Hours Vs Scores(Traning set)")
plt.xlabel("Hours")
plt.ylabel("Scores")
plt.show()
```




```
plt.scatter(x_test,y_test,color='blue')
plt.plot(x_test,regressor.predict(x_test),color='silver')
plt.title("Hours Vs Scores(Test set)")
plt.xlabel("Hours")
```

```
    plt.ylabel("Scores")
    plt.show()
```



Saving...