By Bhagya M

Task no.1-Predict the percentage of marks of an student based on the number of study hours

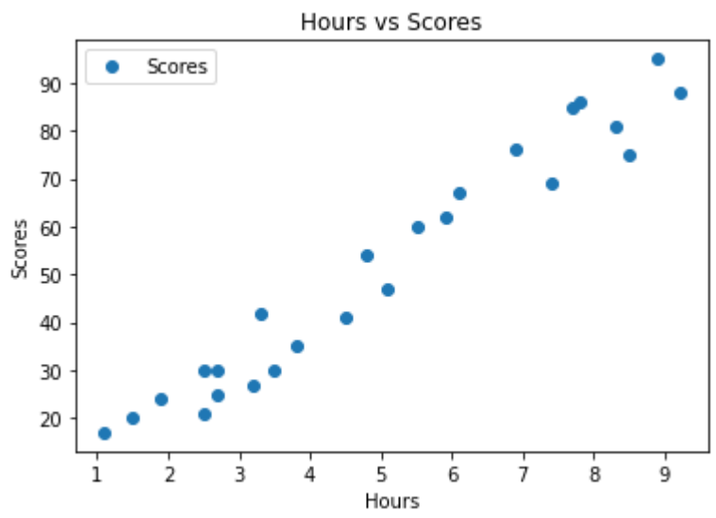
In [ ]:
```
#importing libraries

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

In [23]:
```
#importing dataset from the provided URL

url="http://bit.ly/w-data"
data=pd.read_csv(url)
data.head()
```

Out[23]:

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

In [3]:
```
#plot data on graph

data.plot(x="Hours",y="Scores",style="o")
plt.title("Hours vs Scores")
plt.xlabel("Hours")
plt.ylabel("Scores")
plt.show()
```



In [5]:
```
#preparing the data

X = data.iloc[:, :-1].values
y = data.iloc[:, :1].values
```

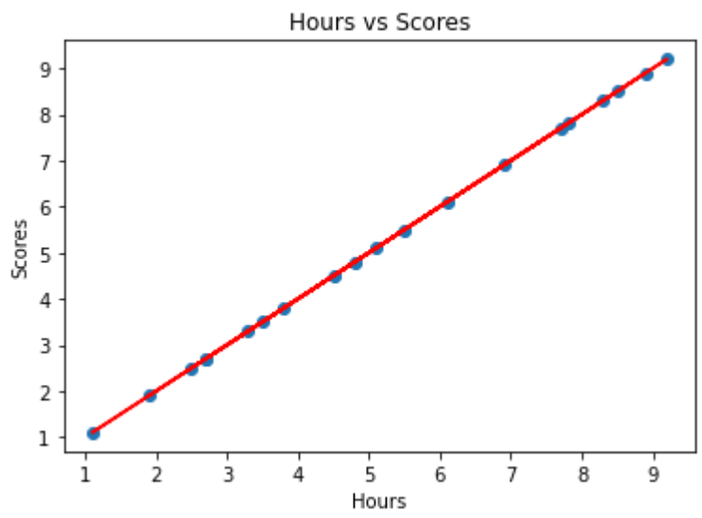
In [6]:
```
#split the data into train set and test set

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 [7]:
```
#train our model

from sklearn.linear_model import LinearRegression
regressor=LinearRegression()
regressor.fit(X_train,y_train)
```

Out[7]:
```
LinearRegression()
```

In [8]:
```
#predict the result

y_pred = regressor.predict(X_test)
```

In [13]:
```
#plot the Regression line on graph

plt.scatter(X_train,y_train)
plt.plot(X_train,regressor.predict(X_train),color='r')
plt.title("Hours vs Scores")
plt.xlabel('Hours')
plt.ylabel('Scores')
plt.show()
```



In [20]:
```
#testing with data

Hours=[[9.25]]
Score_pred=regressor.predict(Hours)
print("No.of Hours={}".format(Hours))
print("Predicted score={}*100".format(Score_pred[0]))
```

```
No.of Hours=[[9.25]]
Predicted score=[9.25]*100
```

In [22]:
```
#Evaluating our model

from sklearn import metrics
print('Mean Absolute Error:',metrics.mean_absolute_error(y_test,y_pred))
```

```
Mean Absolute Error: 3.552713678800501e-16
```

Thank you