Task 1: PREDICTION USING SUPERVISED MACHINE LEARNING

Percentage scored by a student based on study hours.

In [19]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

In [20]:
```
#importing data from the url
from urllib.request import urlretrieve
url= 'https://raw.githubusercontent.com/AdiPersonalWorks/Random/master/student_scores%20-%20student_scores.csv'
urlretrieve(url, 'student_data.txt')
```

Out[20]:
```
('student_data.txt', <http.client.HTTPMessage at 0x220a2434c70>)
```

In [21]:
```
#printing the input data
student_data=pd.read_csv('data.csv')
print(student_data)
```

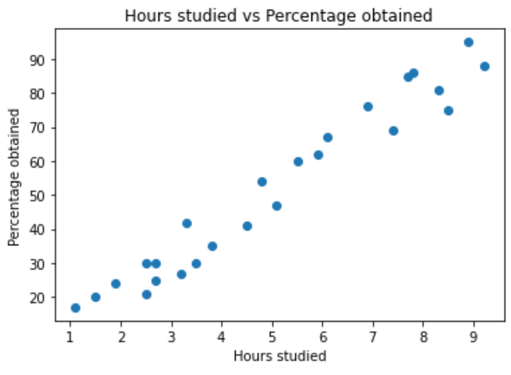
```
    Hours  Scores
0     2.5      21
1     5.1      47
2     3.2      27
3     8.5      75
4     3.5      30
5     1.5      20
6     9.2      88
7     5.5      60
8     8.3      81
9     2.7      25
10    7.7      85
11    5.9      62
12    4.5      41
13    3.3      42
14    1.1      17
15    8.9      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 [22]:
```
student_data.describe()
```

Out[22]:

	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

In [26]:
```
#plotting
x=student_data['Hours']
y=student_data['Scores']
plt.scatter(x, y)
plt.title('Hours studied vs Percentage obtained')
plt.xlabel('Hours studied')
plt.ylabel('Percentage obtained')
student_data.describe()
plt.show()
```



In [24]:
```
#training the data
from sklearn.linear_model import LinearRegression
model=LinearRegression()
x= np.array(x)
x= x.reshape(-1,1)
model.fit(x,y)
```

Out[24]:
```
LinearRegression()
```

In [25]:
```
#finding score for 9.25 hours of study per day
hour=9.25
h= np.array(hour)
h= h.reshape(-1,1)
score=model.predict(h)
print("Hours studied is",hour)
print("Percentage scored is {}".format(score[0]))
```

```
Hours studied is 9.25
Percentage scored is 92.90985477015731
```