In [3]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

In [4]:
```
stud_data = pd.read_csv(r"C:\Users\hp\Music\student_info.csv")
print(stud_data)
```

```
     study_hours  student_marks
0           6.83          78.50
1           6.56          76.74
2            NaN          78.68
3           5.67          71.82
4           8.67          84.19
..           ...            ...
195         7.53          81.67
196         8.56          84.68
197         8.94          86.75
198         6.60          78.05
199         8.35          83.50

[200 rows x 2 columns]
```

In [5]:
```
stud_data.head()
```

Out[5]:

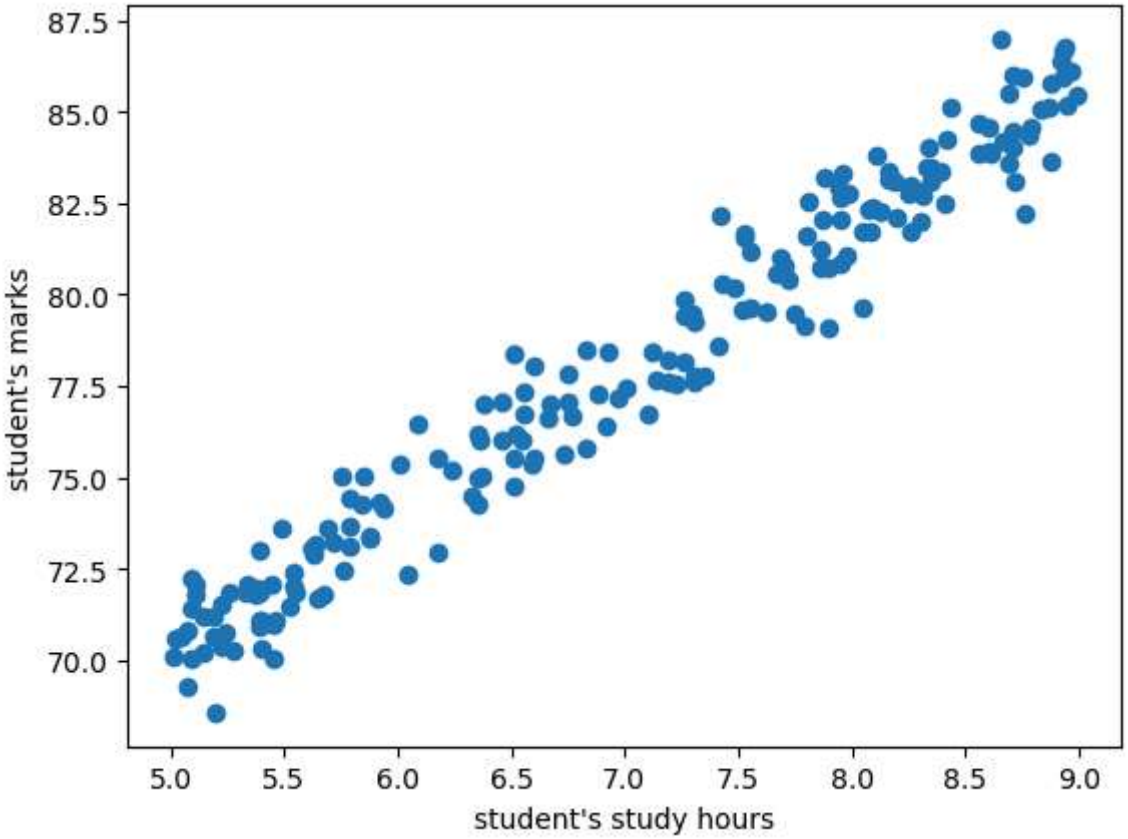
	study_hours	student_marks
0	6.83	78.50
1	6.56	76.74
2	NaN	78.68
3	5.67	71.82
4	8.67	84.19

In [ ]:
```
stud_data.tail()
```

In [ ]:
```
stud_data.describe()
```

In [6]:
```
plt.scatter(x=stud_data.study_hours,y=stud_data.student_marks)
plt.xlabel("student's study hours")
plt.ylabel("student's marks")
```

Out[6]:
```
Text(0, 0.5, "student's marks")
```



In [ ]:
```
stud_data.isnull()
```

In [7]:
```
stud_data.isnull().sum()
```

Out[7]:
```
study_hours      5
student_marks    0
dtype: int64
```

In [8]:
```
stud_data.mean()
```

Out[8]:
```
study_hours       6.995949
student_marks    77.933750
dtype: float64
```

In [9]:
```
student_data=stud_data.fillna(stud_data.mean())
```

In [10]:
```
student_data.isnull().sum()
```

Out[10]:
```
study_hours      0
student_marks    0
dtype: int64
```

In [11]:
```
x=student_data[["study_hours"]]
y=student_data[["student_marks"]]
```

In [12]:
```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.25)
```

In [13]:
```
print("data shape",student_data.shape)
print("x_train shape",x_train.shape)
print("x_test shape",x_test.shape)
print("y_train shape",y_train.shape)
print("y_test shape",x_test.shape)
```

```
data shape (200, 2)
x_train shape (150, 1)
x_test shape (50, 1)
y_train shape (150, 1)
y_test shape (50, 1)
```

In [14]:
```
from sklearn.linear_model import LinearRegression
model=LinearRegression()
model
```

Out[14]:

▾ LinearRegression

```
LinearRegression()
```

In [15]:
```
model.fit(x_train,y_train)
```

Out[15]:

```
LinearRegression
LinearRegression()
```

In [16]:

```
predict=model.predict(x_test)
predict
```

Out[16]:
```
array([[79.86376148],
       [71.65451766],
       [85.31065312],
       [79.16344684],
       [70.64295207],
       [76.12875007],
       [74.80593353],
       [71.07092212],
       [70.48732659],
       [80.40845065],
       [85.5829977 ],
       [74.26124437],
       [80.91423344],
       [71.49889218],
       [76.24546918],
       [78.73547679],
       [72.12139408],
       [70.56513933],
       [84.64924485],
       [70.95420302],
       [84.57143211],
       [84.06564932],
       [79.35797869],
       [72.00467498],
       [82.19814361],
       [70.48732659],
       [77.02359656],
       [81.65345445],
       [81.84798629],
       [83.40424105],
       [76.0898437 ],
       [80.7197016 ],
       [80.79751433],
       [81.10876529],
       [84.26018116],
       [81.49782897],
       [77.98049371],
       [70.56513933],
       [83.24861557],
       [85.73862318],
       [74.14452526],
       [79.20235321],
```

```
       [75.4673418 ],
       [77.87953667],
       [82.50939456],
       [81.80907992],
       [72.66608325],
       [84.45471301],
       [71.30436034],
       [75.54515454]])
```

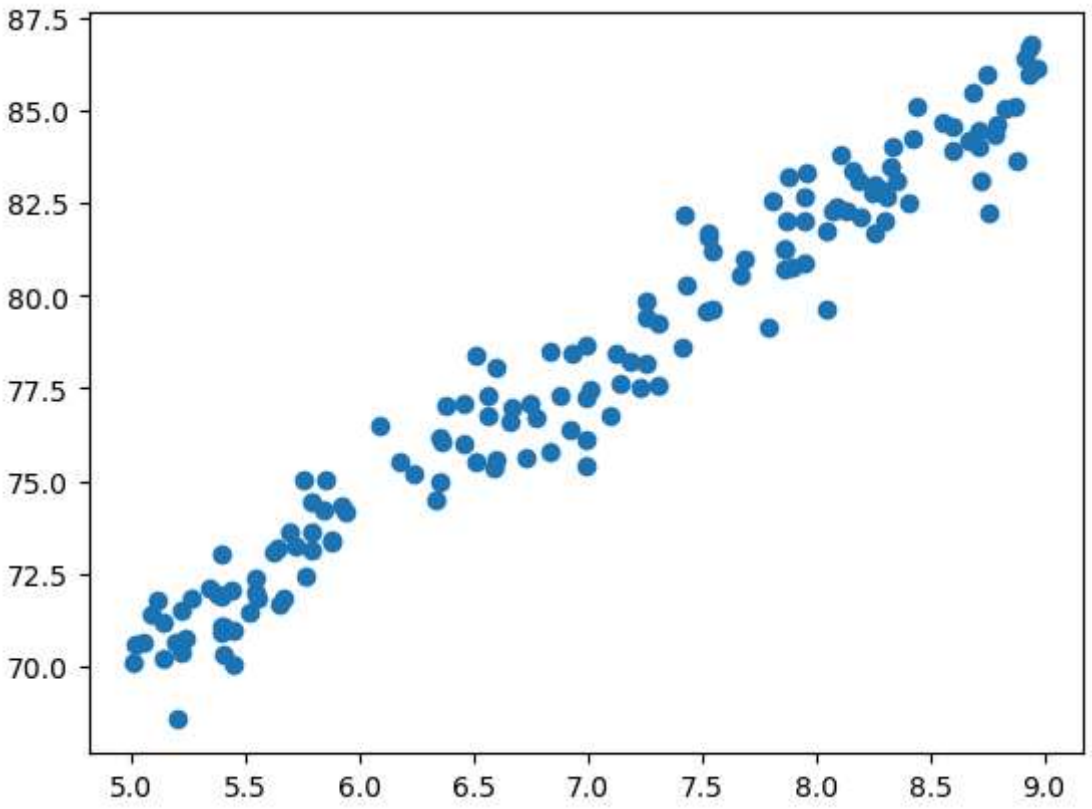
In [ ]:
```
student_data.head()
```

In [17]:
```
model.score(x_test,y_test)
```

Out[17]:
```
0.9582790696752028
```

In [18]:
```
plt.scatter(x_train,y_train)
```

Out[18]:
```
<matplotlib.collections.PathCollection at 0x1c2d806ca10>
```



In [19]:
```
model.predict([[4]])
```

```
C:\Users\hp\OneDrive\Documents\Lib\site-packages\sklearn\base.py:493: UserWarning: X does not have valid feature name
s, but LinearRegression was fitted with feature names
  warnings.warn(
```

Out[19]:
```
array([[66.32434512]])
```

In [20]:
```
model.predict([[8]])
```

```
C:\Users\hp\OneDrive\Documents\Lib\site-packages\sklearn\base.py:493: UserWarning: X does not have valid feature name
s, but LinearRegression was fitted with feature names
  warnings.warn(
```

Out[20]:
```
array([[81.88689266]])
```

In [21]:
```
import joblib
joblib.dump(model,"student_marks_prediction_model.pkl")
```

Out[21]:
```
['student_marks_prediction_model.pkl']
```

In [ ]:
```

```