ML.ex2.linear 2

May 11, 2023

[12]:
```
import pandas as pd

from word2number import w2n
```

[31]:
```
df=pd.read_csv(r"C:\Users\Rakesh\Downloads\hiring.csv")
df
```

[31]:

	experience	test_score(out of 10)	interview_score(out of 10)	salary($)
0	NaN	8.0	9	50000
1	NaN	8.0	6	45000
2	five	6.0	7	60000
3	two	10.0	10	65000
4	seven	9.0	6	70000
5	three	7.0	10	62000
6	ten	NaN	7	72000
7	eleven	7.0	8	80000

[32]:
```
df.experience=df.experience.fillna("zero")
df
```

[32]:

	experience	test_score(out of 10)	interview_score(out of 10)	salary($)
0	zero	8.0	9	50000
1	zero	8.0	6	45000
2	five	6.0	7	60000
3	two	10.0	10	65000
4	seven	9.0	6	70000
5	three	7.0	10	62000
6	ten	NaN	7	72000
7	eleven	7.0	8	80000

[33]:
```
df.experience=df.experience.apply(w2n.word_to_num)
df
```

[33]:

	experience	test_score(out of 10)	interview_score(out of 10)	salary($)
0	0	8.0	9	50000
1	0	8.0	6	45000
2	5	6.0	7	60000
3	2	10.0	10	65000

4	7	9.0	6	70000
5	3	7.0	10	62000
6	10	NaN	7	72000
7	11	7.0	8	80000

[45]:
```
import math

x=math.floor(df['test_score(out of 10)'].mean())
x
```

[45]: 7

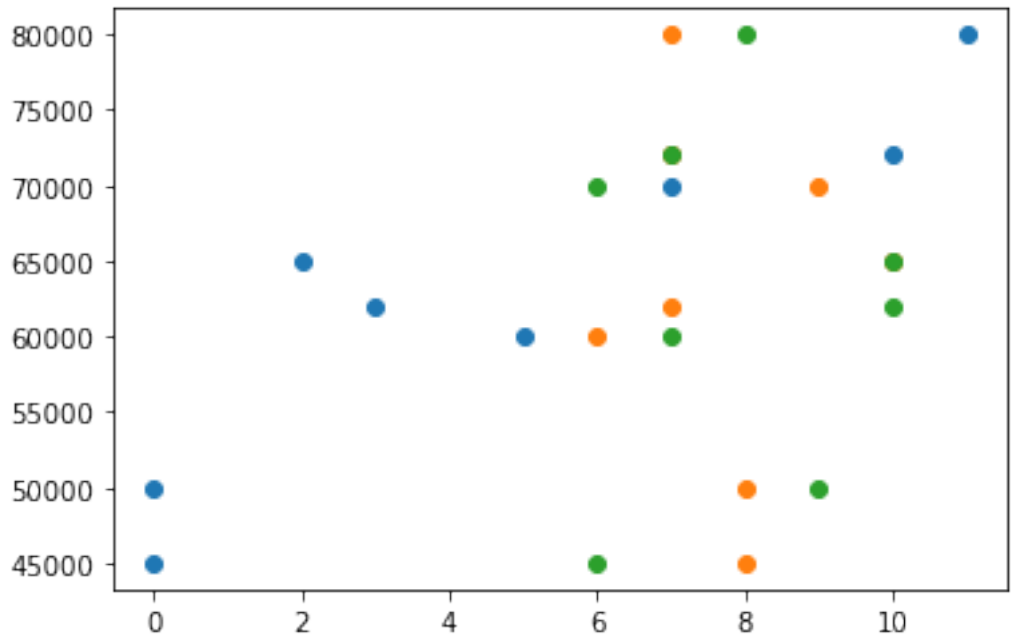
[50]:
```
df['test_score(out of 10)']=df['test_score(out of 10)'].fillna(x)
df
```

[50]:

	experience	test_score(out of 10)	interview_score(out of 10)	salary($)
0	0	8.0	9	50000
1	0	8.0	6	45000
2	5	6.0	7	60000
3	2	10.0	10	65000
4	7	9.0	6	70000
5	3	7.0	10	62000
6	10	7.0	7	72000
7	11	7.0	8	80000

[69]:
```
import matplotlib.pyplot as plt
%matplotlib inline
plt.scatter(df.experience,df['salary($)'])
plt.scatter(df['test_score(out of 10)'],df['salary($)'])
plt.scatter(df['interview_score(out of 10)'],df['salary($)'])

plt.show()
```



[51]:
```
from sklearn import linear_model
```

[58]:
```
reg=linear_model.LinearRegression()
reg.fit(df[['experience','test_score(out of 10)','interview_score(out of␣
↪10)']],df['salary($)'])
```

[58]:
```
LinearRegression()
```

[59]:
```
reg.predict([[2,9,6]])
```

```
C:\Users\Rakesh\anaconda3\lib\site-packages\sklearn\base.py:450: UserWarning: X
does not have valid feature names, but LinearRegression was fitted with feature
names
  warnings.warn(
```

[59]:
```
array([53713.86677124])
```

[60]:
```
reg.predict([[12,10,10]])
```

```
C:\Users\Rakesh\anaconda3\lib\site-packages\sklearn\base.py:450: UserWarning: X
does not have valid feature names, but LinearRegression was fitted with feature
names
  warnings.warn(
```

[60]:
```
array([93747.79628651])
```

[ ]: