

Trần Xuân Triển - B20DCCN691

câu 1

```
In [1]: import pandas as pd
import numpy as np
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split

for i in range(1,3):
    fname = 'data_pttk_'+str(i)
    ex_file = fname + '.xlsx'
    df = pd.read_excel(ex_file)
    df = df.dropna(axis = 1, how = 'all')
    df = df.dropna(axis = 0, how = 'any')

    df.to_csv(fname + '.csv')
```

câu 2 3

```
In [2]: df1 = pd.read_csv('data_pttk_1.csv', index_col=None)
df1 = df1[['0.1','0.1.1','0.2','điểm thi']]
df2 = pd.read_csv('data_pttk_2.csv', index_col=None)
df2 = df2[['0.1','0.1.1','0.2','điểm thi']]
df3 = pd.read_csv('data_pttk_3'+'.csv', index_col=None)
df3 = df3[['10%','20%','20%','Thi']]
df3.columns = ['0.1','0.1.1','0.2','điểm thi']

merged_df = pd.concat([df1, df2, df3], ignore_index=True)
merged_df.columns = ['10%','10%','20%',"Điểm thi"]
merged_df.to_csv('data_pttk.csv', index=False)
```

```
In [3]: df = pd.read_csv('data_pttk.csv')
print(df)
```

	10%	10%.1	20%	Điểm thi
0	9.00	7.00	6.00	6.00
1	9.00	7.00	6.00	6.00
2	10.00	9.00	7.00	7.25
3	9.00	7.00	6.00	6.00
4	10.00	8.00	7.00	7.50
..
146	8.25	7.75	8.25	8.00
147	3.50	3.25	6.00	6.50
148	7.25	6.50	6.00	6.50
149	7.75	3.50	7.50	7.50
150	0.00	0.00	0.00	0.00

[151 rows x 4 columns]

```
In [4]: X = df[['10%', '10%.1', '20%']]
y = df['Điểm thi']
print(X)
print(y)
```

```

      10% 10%.1  20%
0      9.00  7.00  6.00
1      9.00  7.00  6.00
2     10.00  9.00  7.00
3      9.00  7.00  6.00
4     10.00  8.00  7.00
...
146    8.25  7.75  8.25
147    3.50  3.25  6.00
148    7.25  6.50  6.00
149    7.75  3.50  7.50
150    0.00  0.00  0.00
```

```
[151 rows x 3 columns]
```

```

0      6.00
1      6.00
2      7.25
3      6.00
4      7.50
```

```

...
146    8.00
147    6.50
148    6.50
149    7.50
150    0.00
```

```
Name: Điểm thi, Length: 151, dtype: float64
```

Câu 4

```
In [5]: import pandas as pd
import numpy as np
from sklearn.linear_model import LinearRegression
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 = 40)

model = LinearRegression(fit_intercept=False)
model.fit(X=X_train, y=y_train)
```

```
Out[5]:
*      LinearRegression
LinearRegression(fit_intercept=False)
```

```
In [6]: d1 = float(input('Điểm 10%: '))
d2 = float(input('Điểm 10%: '))
d3 = float(input('Điểm 20%: '))

predicted = model.predict([[d1, d2, d3]])
print(predicted)
```

```

Điểm 10%: 6.5
Điểm 10%: 7
Điểm 20%: 8
[8.1725737]
```

Câu 5 6

```
In [7]: import tensorflow as tf
        from sklearn.model_selection import train_test_split

        # Chia dữ liệu thành tập huấn luyện và tập kiểm tra
        X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=40)

        # Xây dựng mô hình Sequential
        model2 = tf.keras.Sequential([
            tf.keras.layers.Dense(16, activation='relu', input_shape=(3,)), # Lớp ẩn với 16 nơ-ron và ReLU activation
            tf.keras.layers.Dense(8, activation='relu'), # Lớp ẩn với 8 nơ-ron và ReLU activation
            tf.keras.layers.Dense(1) # Lớp đầu ra
        ])

        # Biên dịch mô hình
        model2.compile(optimizer='adam', loss='mean_squared_error') # Sử dụng mean squared error Loss

        # Huấn luyện mô hình trên tập huấn luyện
        model2.fit(X_train, y_train, epochs=100, verbose=1) # Chạy 100 epoch

        #câu 6

        # Đánh giá mô hình trên tập kiểm tra
        loss = model2.evaluate(X_test, y_test, verbose=1)
        print("Loss trên tập kiểm tra:", loss)

        # Dự đoán điểm thi cho một ví dụ mới
        d1 = float(input('Điểm 10%: '))
        d2 = float(input('Điểm 10%: '))
        d3 = float(input('Điểm 20%: '))

        new_data = np.array([[d1, d2, d3]])
        predicted = model2.predict(new_data)

        print("Dự đoán điểm thi:", predicted)
```

```
Epoch 95/100
4/4 [=====] - 0s 4ms/step - loss: 0.4706
Epoch 96/100
4/4 [=====] - 0s 3ms/step - loss: 0.4691
Epoch 97/100
4/4 [=====] - 0s 3ms/step - loss: 0.4692
Epoch 98/100
4/4 [=====] - 0s 3ms/step - loss: 0.4682
Epoch 99/100
4/4 [=====] - 0s 2ms/step - loss: 0.4664
Epoch 100/100
4/4 [=====] - 0s 3ms/step - loss: 0.4688
1/1 [=====] - 0s 164ms/step - loss: 0.5195
Loss trên tập kiểm tra: 0.5194907784461975
Điểm 10%: 6
Điểm 10%: 7
Điểm 20%: 8
1/1 [=====] - 0s 121ms/step
Dự đoán điểm thi: [[8.195924]]
```