

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsRegressor
from sklearn.metrics import mean_squared_error
```

```
data = pd.read_csv('D:/ML/c3/3_Phan Nguyễn Anh Khoa_Ch3_LR.csv')
data.describe()
data = data.select_dtypes(['int64', 'float64'])
data
```

✓ 0.0s

	Hours_Studied	Attendance	Sleep_Hours	Previous_Scores	Tutoring_Sessions	Physical_Activity	Exam_Score
0	23	84	7	73	0	3	67
1	19	64	8	59	2	4	61
2	24	98	7	91	2	4	74
3	29	89	8	98	1	4	71
4	19	92	6	65	3	4	70
...	...	...	...	...	...	...	...
6602	25	69	7	76	1	2	68
6603	23	76	8	81	3	2	69
6604	20	90	6	65	3	2	68
6605	10	86	6	91	2	3	68
6606	15	67	9	94	0	4	64

6607 rows × 7 columns

```
data = data[['Hours_Studied', 'Attendance', 'Exam_Score']]
data.corr()
```

✓ 0.0s

	Hours_Studied	Attendance	Exam_Score
Hours_Studied	1.000000	-0.009908	0.445455
Attendance	-0.009908	1.000000	0.581072
Exam_Score	0.445455	0.581072	1.000000

```
x = data[['Hours_Studied', 'Attendance']]
y = data['Exam_Score']
```

✓ 0.0s

```
# Chia dữ liệu thành tập huấn luyện và tập kiểm tra (80% huấn luyện, 20% kiểm tra)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Khởi tạo mô hình KNN với K=13
knn = KNeighborsRegressor(n_neighbors=3)

# Huấn luyện mô hình trên tập huấn luyện
knn.fit(X_train, y_train)

# Dự đoán với tập kiểm tra
y_pred = knn.predict(X_test)

# Đánh giá mô hình
mse = mean_squared_error(y_test, y_pred)
print(f'Mean Squared Error: {mse}')

# Hiển thị kết quả dự đoán
print('Dự đoán:', y_pred)
```

✓ 0.0s

Mean Squared Error: 7.681122877794587

Dự đoán: [67. 64.66666667 69.66666667 ... 68. 66.33333333  
70.66666667]

```
x = data[['Hours_Studied', 'Attendance']]
y = data['Exam_Score']
✓ 0.0s

# Chia dữ liệu thành tập huấn luyện và tập kiểm tra (80% huấn luyện, 20% kiểm tra)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Khởi tạo mô hình KNN với K=13
knn = KNeighborsRegressor(n_neighbors=7)

# Huấn luyện mô hình trên tập huấn luyện
knn.fit(X_train, y_train)

# Dự đoán với tập kiểm tra
y_pred = knn.predict(X_test)

# Đánh giá mô hình
mse = mean_squared_error(y_test, y_pred)
print(f'Mean Squared Error: {mse}')

# Hiển thị kết quả dự đoán
print('Dự đoán:', y_pred)
✓ 0.0s

Mean Squared Error: 6.7369786038469845
Dự đoán: [66.          67.42857143 70.28571429 ... 67.28571429 66.
69.85714286]
```

- Giá trị MSE không quá lớn, cho thấy mô hình đã học khá tốt từ dữ liệu, nhưng vẫn còn một số chênh lệch giữa dự đoán và thực tế.
- Mô hình KNN có thể hoạt động tốt hơn nếu bạn thử các giá trị K khác( k =7 cho thấy mô hình tốt hơn k = 3)