

```
Out[7]:
```

	ID	age	income	student	credit_rating
0	0	2	0	0	1
1	1	2	0	0	0

2	2	0	0	0	1
3	3	1	2	0	1
4	4	1	1	1	1
5	5	1	1	1	0
6	6	0	1	1	0
7	7	2	2	0	1
8	8	2	1	1	1
9	9	1	2	1	1
10	10	2	2	1	0
11	11	0	2	0	0
12	12	0	0	1	1
13	13	1	2	0	0

In [8]: X.columns

Out[8]: Index(['ID', 'age', 'income ', 'student', 'credit\_rating'], dtype='object')

In [9]: #X.classes\_ # Cột các lớp

...

In [10]: X.shape # Cột bao nhiêu hàng và cột (Hàng, cột)

Out[10]: (14, 5)

In [11]: X.info() # Cột thông tin

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 14 entries, 0 to 13
Data columns (total 5 columns):
#   Column      Non-Null Count  Dtype
---  -
0   ID           14 non-null    int64
1   age          14 non-null    int32
2   income       14 non-null    int32
3   student      14 non-null    int32
4   credit_rating 14 non-null    int32
dtypes: int32(4), int64(1)
memory usage: 464.0 bytes
```

In [14]: X.describe()

Out[14]:

	ID	age	income	student	credit_rating
count	14.00000	14.000000	14.000000	14.000000	14.000000
mean	6.50000	1.071429	1.142857	0.500000	0.571429
std	4.18333	0.828742	0.864438	0.518875	0.513553
min	0.00000	0.000000	0.000000	0.000000	0.000000
25%	3.25000	0.250000	0.250000	0.000000	0.000000
50%	6.50000	1.000000	1.000000	0.500000	1.000000
75%	9.75000	2.000000	2.000000	1.000000	1.000000
max	13.00000	2.000000	2.000000	1.000000	1.000000

In [15]: #genre\_mappings = {index: Label for index, Label in enumerate(X.classes\_)}  
#genre\_mappings

In [16]: from sklearn.tree import DecisionTreeClassifier  
regressor = DecisionTreeClassifier()  
regressor.fit(X.iloc[:,1:5],y) # hàm .fit Là Huấn Luyện mô hình (train module)  
# X\_in = np.array([2,0,0,0])  
X\_in = np.array([1,0,0,0])  
y\_pred = regressor.predict([X\_in]) # .predict : Hàm dự đoán . Đưa ra kết quả kiểu thứ  
y\_pred # Trả về nhánh bất kỳ

# Đánh giá độ chính xác

# Nhận diện chữ số viết tay: Bước kiểm thử -

# BTVN: Sử dụng dữ liệu dòng 5 (bất kỳ), trong bảng để trả về không mua máy tính

C:\ProgramData\Anaconda3\lib\site-packages\sklearn\base.py:450: UserWarning:

X does not have valid feature names, but DecisionTreeClassifier was fitted with feature names

Out[16]: array(['no'], dtype=object)

In [7]: #pip install pydotplus sklearn

In [17]: from sklearn.externals.six import StringIO  
from IPython.display import Image  
from sklearn.tree import export\_graphviz  
import pydotplus  
dot\_data = StringIO()  
export\_graphviz(regressor, out\_file=dot\_data, filled=True, rounded=True, special\_characters=True)  
graph = pydotplus.graph\_from\_dot\_data(dot\_data.getvalue())  
graph.write\_png('Tree\_Buy\_Computer.png')  
  
# Conda: conda install graphviz  
# Homework : Lưu tấm hình cây với tên, Show ảnh

Out[17]: True

## Bài tập

- Chọn dữ liệu để xuất hiện qđ: Yes
- Lưu hình cây quyết (qđ) với tên: Cay\_Quyết\_Dinh\_.csv
- Hiếm thị hình ảnh cây (qđ) trên của số Colab
- Thay đổi dữ liệu của file csv

In [18]: # Lưu tấm hình cây với tên, Show ảnh - Xong  
from IPython.display import Image  
Image('Tree\_Buy\_Computer.png')

Out[18]:



