**Code K-Means Đối Với Machine Learning**

**import** pandas **as** pd

**import** matplotlib.pyplot **as** plt

raw\_data **=** pd**.**read\_csv("final-dataset.arff",sep**=**" ")

raw\_data**.**to\_csv("SRC\_ADD")

data**=**pd**.**read\_csv('log.csv', skiprows**=**2, names**=**['IP','Dash','Dash1','Date','TimeZone','Request\_Header',

'Status\_code','Byte\_tansfer','Blank','Browser\_Name'])

df **=** pd**.**DataFrame(data**=**data)

data1 **=** df**.**drop(df**.**columns[[1,2,3,4,5,7,8,9]], axis **=** "columns")

data1

IP**=**data1['IP']

data1['IP']**.**value\_counts()

Status\_code**=**data1['Status\_code']

data1['Status\_code']**.**value\_counts()

data12**=**data1**.**groupby(["IP", "Status\_code"])**.**size()**.**reset\_index(name**=**"frequency")

data12

train\_data **=** data12**.**drop(['IP'], axis**=**1)

**from** sklearn.preprocessing **import** StandardScaler

sc **=** StandardScaler()

data\_scaled **=** sc**.**fit\_transform(train\_data)

print(data\_scaled)

**import** numpy **as** np

**from** sklearn.cluster **import** KMeans

model **=** KMeans(n\_clusters**=**3)

pred **=** model**.**fit\_predict(data\_scaled)

pred

data\_pred **=** pd**.**DataFrame(data\_scaled, columns**=**['Status\_Scaled','Frequency\_scaled'])

data\_pred['Clusters'] **=** pred

final\_data **=** pd**.**concat([data12, data\_pred], axis**=**1)

final\_data

plt**.**scatter(final\_data['IP'], final\_data['Status\_code'], c**=**final\_data['Clusters'])

final\_data**.**plot**.**bar(x**=**'IP', y**=**'frequency', rot**=**0)

**Code Các Thuật Toán Khác Của Machine Learning**