import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

df = pd.read\_csv("heart.csv")

df.info()

df.target.value\_counts()

features = df.iloc[:, 0:13]

label = df['target']

features

pd.crosstab(df.sex,df.target).plot(kind="bar",figsize=(25,6))

plt.title('Heart Disease Frequency for chol')

plt.xlabel('chol')

plt.ylabel('Frequency')

plt.show()

pd.crosstab(df.age,df.target).plot(kind="bar",figsize=(25,6))

plt.title('Heart Disease for fbs')

plt.xlabel('fbs')

plt.ylabel('Frequency')

plt.show()

from sklearn import ensemble

clf\_rand = ensemble.RandomForestClassifier()

from sklearn import linear\_model

clf\_log = linear\_model.LogisticRegression()

from sklearn.model\_selection import train\_test\_split

X\_train, X\_test, y\_train, y\_test = train\_test\_split(features, label, test\_size=0.1, random\_state=2)

X\_train

pred\_rand = clf\_rand.predict(X\_test)

pred\_log = clf\_log.predict(X\_test)

print(pred\_rand)

print(pred\_log)

from sklearn.metrics import confusion\_matrix

cm = confusion\_matrix(y\_test, pred\_rand)

print(cm)

from sklearn.metrics import accuracy\_score

accuracy\_score(y\_test, pred\_rand)

from sklearn.metrics import accuracy\_score

accuracy\_score(y\_test, pred\_log)

clf\_rand.predict([[39, 0, 0, 136, 208, 0, 0, 171, 0, 1.5, 2, 0, 2]])

from joblib import dump, load

loaded\_model = load('model\_rand.joblib')

pred = loaded\_model.predict(X\_test)

pred