EX-6 PCR

code:

import numpy as np

import pandas as pd

from sklearn.decomposition import PCA

from sklearn.linear\_model import LinearRegression

from sklearn.metrics import mean\_absolute\_error, mean\_squared\_error

from sklearn.pipeline import Pipeline

data=pd.read\_csv("PCR.csv")

data.shape

#selected the important features

feature=['RID','age','income','student','credit\_rating']

x=data[feature]

y=data.class\_buy\_computer

reg=LinearRegression()

#create a pipeline with PCA and linear regression

pca=PCA(n\_components=2)

pipeline = Pipeline(steps=[('pca',pca),('reg',reg)])

pipeline.fit(x,y)

pred=pipeline.predict(x)

print("Number of features before PCR:",x.shape[1])

print("Number of features after PCR:",pca.n\_components)

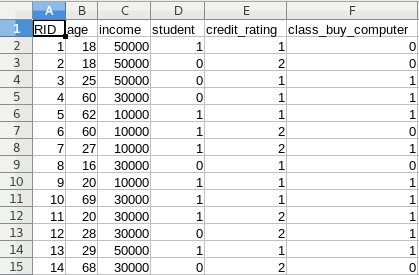
print("MAE",mean\_absolute\_error(y,pred))

print("MSE",mean\_squared\_error(y,pred))

print("RMSE",np.sqrt(mean\_squared\_error(y,pred)))

print("R^2",pipeline.score(x,y))

csv file:



Output:

