import numpy as np

import pandas as pd

from sklearn.metrics import accuracy\_score

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

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

df=pd.DataFrame(data)

df

df.replace('?',np.nan,inplace=True)

df.dropna(axis=1,inplace=True)

df

df=pd.get\_dummies(df)

df

Y=df.iloc[:,0]

X=df.iloc[:,2:]

X\_train,X\_test,y\_train,y\_test=train\_test\_split(X,Y,test\_size=0.2,random\_state=2)

num\_training=y\_train.shape[0]

num\_test=y\_test.shape[0]

print('training:'+str(num\_training)+',test:'+str(num\_test))

from sklearn.naive\_bayes import GaussianNB

gnb =GaussianNB()

gnb.fit(X\_train,y\_train)

y\_pred=gnb.predict(X\_test)

print("Misclassified samples:%d out of %d" %((y\_test!=y\_pred).sum(),y\_test.shape[0]))

print('Accuracy:'+str(accuracy\_score(y\_test,y\_pred)))