SAMPLE CODE

**Model\_fit.py:**

# -\*- coding: utf-8 -\*-

"""

Created on Sun Feb 11 23:12:14 2018

@author: hp

"""

import pandas as pd

import matplotlib.pyplot as plt

#importing the dataset

columns=['buying\_price','maintainence\_cost','number\_of\_doors','number\_of\_seats','luggage\_boot\_size','safety\_rating']

train=pd.read\_csv("train.csv")

test=pd.read\_csv("test.csv",names=columns)

#columns=['buying\_price','maintainence\_cost','number\_of\_doors','number\_of\_seats','luggage\_boot\_size','safety\_rating']

#fitting the model

from sklearn import svm

X=train[columns]

y=train.popularity

clf=svm.SVC(kernel="rbf",C=300)

clf=clf.fit(X,y)

pred=clf.predict(test[columns])

#preparing the csv file

submission\_df={"0":pred}

submission=pd.DataFrame(submission\_df)

m = submission.to\_csv("prediction.csv",index=False,header=False)

#print('pred',pred)

**Data\_process.py**

# -\*- coding: utf-8 -\*-

"""

Created on Sun Feb 11 23:03:39 2018

@author: hp

"""

import pandas as pd

import matplotlib.pyplot as plt

train=pd.read\_csv("train.csv")

#barplot of popularity

fig=plt.figure()

ax = fig.add\_subplot(1,1,1)

ax.hist(train.popularity,bins = 7)

plt.title('car popularity')

plt.xlabel('popularity')

plt.ylabel('cars')

plt.show()

#train\_description

train.describe()

#boxplot of popularity

figure1=plt.figure()

bx=figure1.add\_subplot(1,1,1)

bx.boxplot(train.popularity)

plt.ylabel('popularity')

plt.show()

columns=['buying\_price','maintainence\_cost','number\_of\_doors','number\_of\_seats','luggage\_boot\_size','safety\_rating']

#getting all relationship plots

for i in columns:

fig2=plt.figure()

ax=fig2.add\_subplot(111)

ax.scatter(train[i],train.popularity,color='grey',alpha=0.05)

plt.ylabel('popularity')

plt.xlabel(i)

plt.show()

#checking for the null values

for i in columns:

print(i)

print(train[i].isnull().sum())