Assignment No:3

Name: Ankitkumar Sigh

Div: B

Seat no: B150694205

import pandas as pd

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

from sklearn.linear\_model import LogisticRegression

from sklearn.metrics import accuracy\_score

df = pd.read\_csv("2010-capitalbikeshare-tripdata.csv")

#print df["Member type"].unique()

df.loc[df["Member type"] == 'Member', "Member type"] = 1

df.loc[df["Member type"] == 'Casual', "Member type"] = 2

df.loc[df["Member type"] == 'Unknown', "Member type"] = 3

print df["Member type"]

train,test = train\_test\_split(df,test\_size=0.3)

X = train[["Duration","Start station number","End station number"]]

y = train["Member type"].astype(int)

X\_test = test[["Duration","Start station number","End station number"]]

y\_test = test["Member type"].astype(int)

classifier = LogisticRegression(multi\_class="ovr")

classifier.fit(X,y)

y\_predicted = classifier.predict(X\_test)

score = accuracy\_score(y\_test,y\_predicted)

print "Accuracy ",score

/\*OUTPUT

C:\Users\HP\Desktop>python bikeanalysis.py

0 1

1 1

2 1

3 1

4 1

5 1

6 1

7 1

8 1

9 1

10 1

11 1

12 1

13 1

14 1

15 1

16 1

17 1

18 1

19 1

20 1

21 1

22 1

23 1

24 1

25 1

26 1

27 1

28 1

29 1

..

115567 2

115568 1

115569 1

115570 1

115571 1

115572 1

115573 2

115574 1

115575 1

115576 1

115577 1

115578 1

115579 1

115580 2

115581 1

115582 1

115583 1

115584 1

115585 1

115586 1

115587 1

115588 2

115589 2

115590 1

115591 1

115592 2

115593 1

115594 1

115595 1

115596 2

Name: Member type, Length: 115597, dtype: int64

Accuracy 0.8519031141868512

\*/