**Achinta Gorain**

**Logistic Regression on cancer dataset (q.1)**

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

%matplotlib inline

cancer=pd.read\_csv('cancer\_data.csv')

cancer.head(2)

sns.countplot(x=’diagnosis’,data=cancer)

print(‘total no. of patients = ’,len(cancer.index))

sns.countplot(x=’diagnosis’,hue=’ area\_se’,data=cancer)

sns.countplot(x=’diagnosis’,hue=’ perimeter\_se’,data=cancer)

sns.countplot(x=’diagnosis’,hue=’ concavity\_mean’,data=cancer)

sns.countplot(x=’diagnosis’,hue=’ fractal\_dimension\_mean’,data=cancer)

sns.countplot(x=’diagnosis’,hue=’ radius\_mean’,data=cancer)

cancer.head(2)

cancer.isnull.sum()

sns.heatmap(cancer.isnull())

X=cancer.drop([‘diagnosis’],axis=1)

Y=cancer[‘diagnosis’]

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

X\_train,X\_test,y\_train,y\_test=train\_test\_split(X,y,test\_size=0.5,random\_state=0)

From sklearn.linear\_model import LogisticRegression

Lmodl=LogisticRegression()

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

X\_test

Y\_test

Y\_pred=lmodl.predict(X\_test)

Y\_pred

Y\_test

From sklearn.metrices.import accuracy\_score

Acc=lmodl.score(X\_test,y\_test)

Print(‘The logistic regression model score =’,acc\*100)