

# CSE 472 : Assignment 1 : Report

Name : Fardin Zaman

ID : 1605008

Logistic Regression :

Dataset 1(telco) :

Performance measure	Training	Test
Accuracy	0.8028044018459354	0.7977288857345636
Sensitivity	0.5336442371752165	0.49728260869565216
Specificity	0.9005564964916526	0.90393852065321810
Precision	0.6608910891089109	0.6466431095406361
False discovery rate	0.3391089108910891	0.3533568904593639
F1 score	0.5904902322152599	0.5622119815668202

Dataset 2(adult) :

Performance measure	Training	Test
Accuracy	0.8254353367525568	0.8288188686198636
Sensitivity	0.482336436678995	0.4823192927717109
Specificity	0.9342637540453075	0.9359871330920788
Precision	0.6994636582208249	0.6997359486986043
False discovery rate	0.3005363417791751	0.30026405130139566
F1 score	0.5709541062801933	0.5710327843620132

Dataset 3(creditcard) :

Performance measure	Training	Test
Accuracy	0.9745013115354114	0.9717004147353013
Sensitivity	0.0	0.0
Specificity	1.0	1.0
Precision	nan	nan
False discovery rate	nan	nan
F1 score	nan	nan

AdaBoost :

Dataset 1(telco) :

No. of boosting rounds	Training	Test
5	0.80386936457224	0.7977288857345636
10	0.7997870074547391	0.7892122072391767
15	0.797124600638977	0.794889992902768
20	0.7925097621583245	0.7785663591199432

Dataset 2(adult) :

No. of boosting rounds	Training	Test
5	0.8255274715150026	0.8275290215588723
10	0.820490771167961	0.8245193784165592
15	0.815085531771137	0.817333087648178
20	0.8093424649120113	0.811743750383883

Dataset 3(creditcard) :

No. of boosting rounds	Training	Test
5	0.9745013115354114	0.9717004147353013
10	0.9745013115354114	0.9717004147353013
15	0.9745013115354114	0.9717004147353013
20	0.9745013115354114	0.9717004147353013

For 3 datasets, there are 3 preprocess steps :

1. preprocess1
2. preprocess2  
preprocess2\_2
3. preprocess3

There are 3 portions in the code. Two of them should be commented when using a dataset.  
They start with

```
dataset_dummies = preprocess1()  
dataset_train = preprocess2()  
dataset = preprocess3()
```

from the end of one portion, another begins.  
The code was implemented in google colab(python notebook).  
So, the python file may have some issues