**Read Me:**

We submitted this project as team of 2 students and divided the work in following manner:

|  |  |
| --- | --- |
| Dataset 1 | Breast Cancer Dataset |
| Dataset 2 | Customer Churn Dataset |

|  |  |  |
| --- | --- | --- |
|  | Implemented By | |
| Algorithm | Mrunmayi | Anup |
| KNN | Dataset 1 | Dataset 2 |
| ANN | Dataset 2 | Dataset 1 |

Each one of us contributed to report in such a way that the one who worked on the dataset using algorithm, wrote the conclusion in the report.

**Steps to run the code:**

* In total 3 datasets are present.
* First the KNN is implemented on 2 datasets, later the ANN is implemented on the 2 datasets.
* For ANN implementation of customer churn dataset,

We have included an interim generated file Churn\_h2o.csv which is the preprocessed version of Churn dataset

**For some of the graphs we have used h2o Flow, which can get activated in following simple steps:**

To download & use h2o Flow please follow instructions on the below URL -

http://h2o-release.s3.amazonaws.com/h2o/rel-weierstrass/7/index.html

Get started with H2O in 3 easy steps

1. Download H2O. This is a zip file that contains everything you need to get started.

2. From your terminal, run:

cd ~/Downloads

unzip h2o-3.14.0.7.zip

cd h2o-3.14.0.7

java -jar h2o.jar

3. Point your browser to http://localhost:54321

4. Once you run the R code, you should be able to see the data-sets as well as models in the h2o Flow web interface under -

Admin -> Jobs