README

The python program is saved as lr.py and at runtime train and test file (**NOTE: either provide absolute path or just invoke the pyhton program from the submission itself**) should be provided.

$ python3 lr.py train.csv test.csv

Note: Program is an interactive one just input necessary input at each stage for the needed output

```
atul@linux:~/college/cs725/Assignment$ python3 lr.py /home/atul/college/cs725/Assignment/train.csv /home/atul/college/cs725/Assignme
nt/test.csv
Reading Files...
Done
::::::::::::::LINEAR REGRESSION:::::::::::::::::
1. L2 regularisation
2. P-Norm regularisation
3. L2 regularisation with basis
4. Optimised one(P-6 with sigmoidal basis)
Your choice: 1
lambda value : 10000
Learn rate :0.001
No. of iterations: 100000
```

Figure 1: Interactive Console

The following model settings is common across all norms of regularisation and basis functions applied:

1. **Min-Max Normalization is used.** $\quad \forall x \dfrac{x - min(\forall x)}{max(\forall x) - min(\forall x)}$

2. **Held Out Validation is used, in proportion of 20 – 80, 20% for validation set and 80% validation for train set.**

3. **Error rate = abs(cost of train-set at i^{th} iteration - cost of train-set at $(i-1)^{th}$ iteration) is taken 10^{-18}**

4. **For the negative prediction, negative cutoff is used.**

Hyperparameters used for given tasks are as follows:

(λ = Regularisation Factor, learn = Learning rate, Iterations = No. Of Iterations)

NOTE: FOR MSE and RMSE refer report.pdf

1. **Task -1:**

   First Select L2 regularisation option then enter the followings(see Figure1):

   λ   = 10000

   learn   = 0.001

Iterations = 100000

For unregularised :

λ   = 0

learn   = 0.03

Iterations = 100000

2. **Task -2:**

To have the graphs, at last of the execution, you will be asked to whether you want to see the top features or not.[Press y]
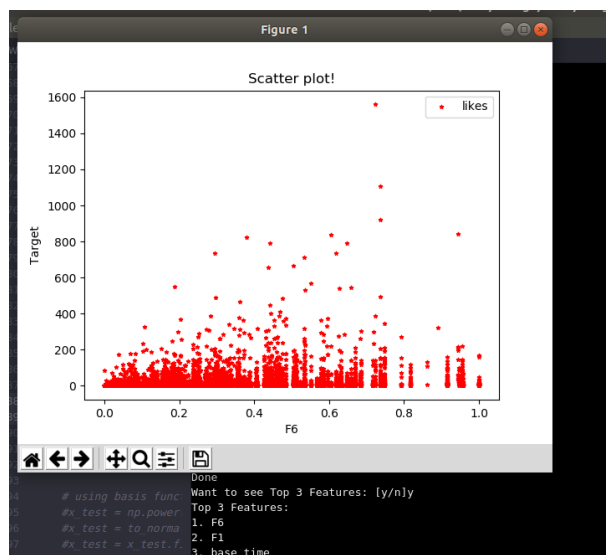

Figure 2: Scatter plot for top feature

3. **Task -3:**

First select option -2 viz., P-Norm from the option.(see Figure1)
p   = 4

λ   = 109000
learn   = 0.001

Iterations = 100000

For P-6 Norm:

p   = 4

λ   = 119000
learn   = 0.001

Iterations = 100000

4. Task -4

First select option-3 viz., L2 regularisation with basis function(see Figure1)

1. **Inverse Sigmoidal**

λ = 740
learn = 0.001

Iterations = 100000

2. **Gaussian Function**

λ = 13000
learn = 0.001

Iterations = 100000

5. **Task -5**

First select option-4 Optimised(P6-Norm with Inverse sigmoid basis)

(see Figure1)

λ = 706
learn = 0.001

Iterations = 100000