Assignment 4

Machine Learning COMS 4771

Spring 2014, Itsik Pe'er

Assigned: Feb 19th Due: Class time, Feb 26th

Submission: Your submission folder on Courseworks. Submit folders for `Assignment04.Problem01,` and `Assignment04.Problem02`

1)

a) Implement logistic-regression gradient descent, writing your own function
`ThetaStar = GradDescentLogistic(x,y,eta,epsilon,StartingTheta,StopTime)`
First 5 parameters should be clear from class. `StopTime` is time in seconds after which you decide the program is not converging.
[20 points]

b) Interactively work with Matlab Write to test `GradDescentLogistic` . Save the commands you run into a script `TestGDL.m` that should:
   i) Set `StopTime` to 60
   ii) Choose N=100 `x` values in D=8 dimensions randomly & uniformly within the range $(-5,5)^D$
   iii) Use `SimLogistic` with σ=0 to generate N training classification values `y`, and evaluate your classifier using M=20 test datapoints generated with the same parameters.
   iv) Perform multiple runs on the same input , with the same and different `eta,epsilon,StartingTheta` with `eta,epsilon` that you set as you wish and `StartingTheta` chosen randomly in $(-10,10)^D$.

Perform as many runs as you see fit, till you decide on parameters `eta,epsilon` you think are good (call them η_0, ϵ_0). Measure the time it took `GradDescentLogistic` to converge for each run. The last run is with η_0, ϵ_0.
Report:
`TestGDL.m` the script that does all this
`Inputs.txt` D+1 columns, N+M rows detailing `x,y` (last column is `y`, last M rows are test)
`RealThetas.txt` One column vector of D+1 entries (the first one for the free coefficient)
`Runs.txt` 2D+7 columns with a row for each runs. First two columns report: `eta,epsilon` ; next D+1: `StartingTheta`; next D+1: optimized θ values; next 2 columns: final value of the loss on training and test data; last column: runtime in seconds
[45 points]

c) Rerun with `eta` ← $2^k\eta_0$ for consecutive (positive and negative) integers k to demonstrate the problems with too large/small `eta`. Report `EtaRuns.txt`, formatted exactly as `Runs.txt` above, and `TestEta.m` the script that produces it.
[15 points]

2) Consider the neural network in the attached NN.A4.2.pdf, with two inputs, denoted by subscripts 1 and 2; two 1st -level logistic neurons, denoted by subscripts 3 and 4; two 2nd -level logistic neurons, denoted by subscripts 5 and 6; and a single 3rd –level logistic neuron, denoted by the subscript 7. The intercept for all neurons is modeled as respective coefficients for $x_0=1$. Write down explicit update equations for a gradient descent process with logistic loss.
[20 points]

Good luck!