INF552: Programming Assignment 4

Part 1: Implementation [7 points]

[2 points] Implement the Perceptron Learning algorithm. Run it on the data file "classification.txt" ignoring the 5th column. That is, consider only the first 4 columns in each row. The first 3 columns are the coordinates of a point; and the 4th column is its classification label +1 or -1. Report your results.

[1 point] Implement the Pocket algorithm and run it on the data file "classification.txt" ignoring the 4th column. That is, consider only the first 3 columns and the 5th column in each row. The first 3 columns are the coordinates of a point; and the 5th column is its classification label +1 or -1. Plot the number of misclassified points against the number of iterations of the algorithm. Run up to 7000 iterations.

[3 points] Implement Logistic Regression and run it on the points in the data file "classification.txt" ignoring the 4th column. That is, consider only the first 3 columns and the 5th column in each row. The first 3 columns are the coordinates of a point; and the 5th column is its classification label +1 or -1. Use the sigmoid function $\Theta(s) = e^s/(1+e^s)$. Run up to 7000 iterations.

[1 point] Implement Linear Regression and run it on the points in the data file "linear-regression.txt". The first 2 columns in each row represent the independent X and Y variables; and the 3rd column represents the dependent Z variable. Report your results.

You can write your programs in any programming language. However, you will have to implement the algorithms yourself instead of using library functions, except for solving linear equations. Provide a description of the data structures you use, any code-level optimizations you perform, any challenges you face, and of course, the requested outputs.

Part 2: Software Familiarization [2 points]

Do your own research and find out about library functions that offer good implementations of linear classification, linear regression, and logistic regression. Learn how to use them. Compare them against your implementations and suggest some ideas for how you can improve your code. Describe all this in your report.

Part 3: Applications [1 point]

Do your own research and describe some interesting applications of linear classification, linear regression, and logistic regression.

Submission Guidelines

In your report, please include the names of all group members and mention their individual contributions. The report should be in PDF format. Your submission should include the code as well as the report and is due before **10/28**, **11:59pm** in an archive in zip, tar.gz or tar.xz format. Only one

submission is required for each group by one of the group members. Please submit your homework on **BlackBoard** (do NOT email the homework to the instructor or the TA).