

Homework Assignment 4

(100 points)

Assigned Date: Tuesday March 8, 2016

Due Date: 4:00 PM Thursday March 24, 2016

Educational Goal

Become familiar with logistic regression and regularization.

Requirements

Two datasets:

http://www.cs.umb.edu/~ding/classes/438_638/homework/hwk4/

1. Logistic Regression:

hwk4_data1.txt provides two examination scores, *Exam1*, *Exam 2*, *Status*, for each applicant and their correspondent admission decision. 0 is for not admitted and 1 is for admitted.

- Visualize the data.** X-axis for Exam 1 score, Y-axis for Exam 2 score, + (plus) for admitted, and circle for not admitted.
- Implement cost function and gradient descent.** Plot the cost with the number of iterations.
- Report optimal final θ .** Plot the decision boundary with the training data
- If a student has an Exam 1 score 100 and an Exam 2 score 50, what is the admission probability of this student?**

2. Regularized Logistic Regression:

hwk4_data2.txt provides two test scores, *test1*, *test 2*, *Status*, for each device and their correspondent acceptance decision. 0 is for rejection and 1 is for acceptance.

- Visualize the data.** X-axis for test 1 score, Y-axis for test 2 score, + (plus) for accepted, and circle for rejected.

- b. Implement cost function and gradient descent.** Plot the cost with the number of iterations. Note you should construct features from two features test 1 and test 2 scores to the 7th power. For example, $x_1, x_2, x_1^2, x_1x_2, x_2^2, x_1^3, \dots, x_1x_2^6, x_2^7$.
- c. Report optimal final θ .** Plot the decision boundary with the training data
- d. Demonstrate the impact of the regularization parameter λ with 1, 0, 100, 1000.** Plot the decision boundary of for the 4 different λ values.

Submission Requirements

1. Provide a single runnable script for your homework code. You will demonstrate your program to your TA Yahui Di in the Unix Lab from 3:00 to 4:00 PM on Monday March 28 or Tuesday March 29. You may bring your own laptop or use a machine in the Unix lab.
2. Report: Prepare experimental results in PPT slides (can be more than 1 slide; saved as a PDF file) with your hwk4_firstname_lastname.pdf and submit the PDF file to the Blackboard.
3. Zip your source code into one ZIP file and save it as hwk4_firstname_lastname.zip and submit the ZIP file to the Blackboard.
4. Submit paper copies of your PPT slides report and source code in class.
5. In order to receive credits for this homework, you must submit hard copies (in class) and soft copies (to the Blackboard) and demonstrate your program to your TA in the Unix lab during the given time period.