

[CSC 5825 Fall 2017]

Due. Before Class Sept. 19, 2018 Homework 1

Total Points: 100

Non-Programming Questions:

Question 1. (10 points) Imagine you have two possibilities: You can fax a document, that is, send the image, or you can use an OCR and send the text file. Discuss the advantage and disadvantages of the two approaches in a comparative manner. When would one be preferable over the other?

Question 2. (10 points) Let us say you are given the task of building an automated taxi. Define the constraints. What are the inputs? What is the output? How can you communicate with the passenger? Do you need to communicate with the other automated taxis, that is, do you need a language?

Question 3. (10 points) Let us say our hypothesis class is a circle instead of a rectangle. What are the parameters? How can the parameters of a circle hypothesis be calculated in such a case? What if it is an ellipse? Why does it make more sense to use an ellipse instead of a circle?

Question 4. (10 points) Derive equation 2.17 in textbook.

Programming Questions:

Question 1. (40 points) Logistic Regression with Stochastic Gradient Descent

In this question, you are asked to implement stochastic gradient descent (perceptron learning in slides) to learn the weights for logistic regression. The input variables (X) are the positions of 400 different points, the response variable (y) is the class that each x in X should belong to. Note that you only have two classes “0 and “1. Here are the instructions:

- Use the code provided to generate dataset X and y , visualize the data in a scatter plot with different colors for each class.
- Insert a column of ones as the bias in the input variable matrix X
- The activation function we use in this question is sigmoid. Write a function to calculate the output of sigmoid activation function for a given input t

$$S(t) = \frac{1}{1 + e^{-t}} \tag{1}$$

- Randomly Initialize the weight matrix and initialize a learning rate.
- Pick first row from the matrix X , use the current weight matrix and sigmoid function to calculate the current output.
- Calculate the error and update the weight matrix. (Note: The error is defined as the squares error)
- Repeat several passes over the entire training dataset until converge.
- Average the loss for each epoch and visualizes it. (Note: One epoch means a single pass through the entire training set.)
- Use final weight matrix to plot the line to separate points on the same plot in first step.

Question 2. (20 points) Cross-validation. First divide the data set into 5 partitions. Then use the logistic classifier you built in programming problem 1 and 4 partitions to train the model and the remaining 1 partition to test the model. Please plot the test errors for each of the five folds.

Submission Instructions

Homework must be submitted electronically through Blackboard website on/before the due date/time. Homework assignments are usually due in class at the beginning of lecture on the due date given. Homework must be typed with LaTeX or Word. Late homeworks will be accepted with penalty, i.e., 20% penalty if late for 24 hours or less, 40% penalty if late for 48 hours or less and so on. You will receive no credit if late for 120 hours or more.