Spring 2019

ECE471/571 – Homework #4 (Due 04/16)

Note: UG: 100, G: 100

Problem 1: (25/15) On Gradient Descent (GD). Find the global minimum of f(t) = 50 * sin(t) + t^2 over -10<=t<=10. This problem intends to give you a hands-on experience on how gradient descent works and how it can get trapped at the local minima.

a) (5/3) Use MATLAB (or whatever toolbox) to plot this function. Visualize the multiple local minima and the global minimum.

b) Use gradient descent to find the global minimum. (you can use the code provided in the lecture note)

   i) (8/4) Pick a starting point at t=7. What's the minimum? Show the convergence path. Experimenting with different learning rate.

   ii) (7/3) Pick a starting point at t=1. What's the minimum? Show the convergence path. Experimenting with different learning rate.

   iii) (5/5) Comment on the results from the above experiments

Problem 2: (10/10) On Perceptron. Use Perceptron to implement the OR logic and the XOR logic. Show output from each iteration with the maximum number of iterations being 10. You can use either MATLAB or C/C++.

Problem 3: (25/25) Comparison between FLD, PCA, and Perceptron.

Note that FLD and PCA are dimensionality reduction methods that only output a projection direction. Additional classification methods need to be applied to find the decision boundary. Suppose the minimum (Euclidean) distance (MD) classifier is used. On the other hand, Perceptron is a linear classifier that outputs the decision boundary directly. On the same figure, plot the four samples of the AND gate, and compare the decision boundary from FLD+MD, PCA+MD, and Perceptron. You can use whichever language that you feel comfortable (pencil & paper, MATLAB, C/C++).

Problem 4: (20/20) On decision tree. Assume 100 samples are classified to node N, among which 90 samples actually belong to class 1 and 10 samples actually belong to class 2. For the following two split candidates, which one is better according to the Gini impurity?

- Option 1: 70 class 1 samples go to “left”, 20 class 1 samples and 10 class 2 samples go to “right”

- Option 2: 80 class 1 samples go to “left”, 10 class 1 samples and 10 class 2 samples go to “right”

Problem 5: (20/30) For a data set of four samples, divided into two categories as follows: Note that for 571 students, you need to review some materials related to SVM before the lecture.

x1	x2	label
0	0	0
0	1	0
1	0	0
1	1	1

(a) (0/5) Three classification algorithms have been applied, minimum distance, perceptron, and SVM. The figure showed the decision boundary for minimum distance and perceptron. Show the decision boundary for SVM.

(b) (0/5) In a heuristic manner, describe how SVM achieves the best generalization capability.

(c) (20/20) If moving the second sample from (0,1) to (0,0.9), use clustering approaches to classify the samples into two classes using dmin and dmax, respectively.