CSD361: Introduction to Machine Learning
Assignment #2: Linear discriminants

Due on: 23-2-2019, 23.59 13-2-2022
MM: 170

- *Pl. do not copy. If copying is established you will get zero marks in the assignment. For repeat offences you can be given a failing grade in the course.*

1. In this assignment you will experiment with linear discriminants on artificial data sets and then use a binary classifier to classify the Iris data.

   Generate two data sets D1 (a separable) data set and D2 (a non-separable) data set. One way to do this is to construct a random hyper-plane in m dimensions and then generate points on either side of the hyper-plane. Remember that the hyper-plane equation is $g(\mathbf{x}) = \mathbf{w'}^T\mathbf{x'} = 0$ and the perpendicular distance of an arbitray vector $\mathbf{x}$ from the hyper-plane is $\frac{|g(\mathbf{x})|}{\|\mathbf{w}\|}$. So, if $g(\mathbf{x}) > 0$ then $\mathbf{x} \in C_1$ and it is in C_2 if $g(\mathbf{x}) < 0$. Accordingly, generate n vectors (choose a suitable $n > 100$) equally divided in classes C_1 and C_2 respectively. To generate a non-separable set, first generate a separable set then randomly choose some vectors from each set and flip their labels to obtain D2. While this does not guarantee the non-separability of D2 it will be non-separable with very high probablity.

   Your programs should have m (dimension of the feature vector) and n (number of feature vectors or size of the learning set) as parameters. Run your algorithms with at least two different values of m and n.

   (a) Implement the batch and incremental perceptron algorithms with data set D1. Try with two types of $\rho(t)$ i) a suitably chosen constant value and ii) a value that depends on the iteration number t. Report the hyper-plane found and the number of iterations when perfect classification is achieved in each case.

   (b) Implement the pocket algorithm. You should use D2 in this case. Report the hyper-plane, number of iterations and the number of vectors classified correctly.

   (c) Implement the algorithm that minimizes squared loss. Do this on both data sets D1 and D2. In each case report the hyper-plane and the number of vectors classifed correctly.

   (d) Classify the Iris data set (of assignment 1) using both one versus rest and one versus one classification. The data set has 3 labels. You can use any of the algorithms in parts (b) or (c) above to build the binary classifiers. Report the accuracy results you get by both methods.

[40,40,40,50=170]