PATTERN RECOGNITION LAB
CSE 4214

LAB EXPERIMENT 2

IMPLEMENTING THE PERCEPTRON ALGORITHM FOR FINDING THE WEIGHTS OF A LINEAR DISCRIMINANT FUNCTION. (MANY-AT-A-TIME APPROACH)

SUBMITTED BY

MD. MUKITUL ISLAM 14.02.04.076

SECTION: B1

AHSANULLAH UNIVERSITY OF SCIENCE & TECHNOLOGY

Implementing the Perceptron Algorithm for Finding the Weights of a Linear Discriminant Function. (many-at-a-time approach)

Md. Mukitul Islam (140204076)
Dept. of Computer Science & Engineering
Ahsanullah University of Science & Technology

I. Objective

Objective of this report is to show the step by step calculation process of the Perceptron algorithm for finding the weights of a linear discriminant function. Here, two epochs including step by step numerical example have shown.

II. Problem Description

Considering a two class problem. The two classes are-
$w1$ = (1, 1), (1, -1), (4, 5)
$w2$ = (2, 2.5), (0, 2), (2, 3)
These two classes cannot be separated with a linear boundary. So, we have to convert those sample points to high dimensional sample points. Then we have to normalize any one of two class and apply Perceptron algorithm (many-at-a-time approach) to find the weights of a linear discriminant function.

III. High Dimensional Sample Points & Normalization

Here, we are going to generate high dimensional sample points from the sample points of class $w1$ and $w2$.
For this purpose we are using formula-
$y = [x_1^2 \quad x_2^2 \quad x_1 * x_2 \quad x_1 \quad x_2 \quad 1]$
In each class there are 3 sample points.

High dimensional points for class $w1$-
(i) for sample point $(1, 1)$
Here, $x_1 = 1$ and $x_2 = 1$
So, $y_{11} = [1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1]$
(ii) for sample point $(1, -1)$
Here, $x_1 = 1$ and $x_2 = -1$
So, $y_{12} = [1 \quad 1 \quad -1 \quad 1 \quad -1 \quad 1]$
(iii) for sample point $(4, 5)$
Here, $x_1 = 4$ and $x_2 = 5$
So, $y_{13} = [16 \quad 25 \quad 20 \quad 4 \quad 5 \quad 1]$

High dimensional points for class $w2$-
(i) for sample point $(2, 2.5)$
Here, $x_1 = 2$ and $x_2 = 2.5$
So, $y_{21} = [4.00 \quad 6.25 \quad 5.00 \quad 2.00 \quad 2.50 \quad 1.00]$
(ii) for sample point $(0, 2)$
Here, $x_1 = 0$ and $x_2 = 2$
So, $y_{22} = [0 \quad 4.00 \quad 0 \quad 0 \quad 2.00 \quad 1.00]$
(iii) for sample point $(2, 3)$
Here, $x_1 = 2$ and $x_2 = 3$
So, $y_{23} = [4.00 \quad 9.00 \quad 6.00 \quad 2.00 \quad 3.00 \quad 1.00]$

Now, we have to normalize any one of the two classes. We are normalizing high dimensional sample points of class $w2$. So, after normalizing we get-
$y_{21} = [-4.00 \quad -6.25 \quad -5.00 \quad -2.00 \quad -2.50 \quad -1.00]$
$y_{22} = [0 \quad -4.00 \quad 0 \quad 0 \quad -2.00 \quad -1.00]$
$y_{23} = [-4.00 \quad -9.00 \quad -6.00 \quad -2.00 \quad -3.00 \quad -1.00]$

IV. Applying Perceptron Alogorithm: Many-At-A-Time Approach

Let us consider our learning rate $alpha = 0.7$ and our initial weight vector

$$a = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Now $Epoch : 1$
In this epoch each sample point y is multiplied with the transpose of weight vector a. Two sample matrix multiplication are shown below-

	y	a	$g = a^T y$	g<0
y_{11}	[1 1 1 1 1 1]	[1 1 1 1 1 1]	6.00	
y_{12}	[1 1 -1 1 -1 1]		2.00	
y_{13}	[16 25 20 4 5 1]		71.00	
y_{21}	[-4.00 -6.25 -5.00 -2.00 -2.50 -1.00]		-20.75	misclassified
y_{22}	[0 -4.00 0 0 -2.00 -1.00]		-7.00	misclassified
y_{23}	[-4.00 -9.00 -6.00 -2.00 -3.00 -1.00]		-25.00	misclassified

Weight vector a is a column vector and each sample y is a column vector. Now, for

$$y_{11} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \text{ and } a = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$g = a^T y$

$$g = [1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1] \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$g = (1*1) + (1*1) + (1*1) + (1*1) + (1*1) + (1*1)$
$g = 6$
Now for

$$y_{21} = \begin{bmatrix} -4.00 \\ -6.25 \\ -5.00 \\ -2.00 \\ -2.50 \\ -1.00 \end{bmatrix} \text{ and } a = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$g = a^T y$

$$g = [1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1] \begin{bmatrix} -4.00 \\ -6.25 \\ -5.00 \\ -2.00 \\ -2.50 \\ -1.00 \end{bmatrix}$$

$g = (1*(-4.00)) + (1*(-6.25)) + (1*(-5.00)) + (1*(-2.00)) + (1*(-2.50)) + (1*(-1.00))$
$g = -4.00 - 6.25 - 5.00 - 2.00 - 2.50 - 1.00$
$g = -20.75$
After these calculation for 3 cases we have got negative g. So, those y are misclassified. So, we have to find a new weight for starting new epoch.
As our approach is many-at-a-time, so we have to add those misclassified data with the current weight to get a new weight, just like shown below-
$y_{21} = [-4.00 \quad -6.25 \quad -5.00 \quad -2.00 \quad -2.50 \quad -1.00]$
$y_{22} = [0 \quad -4.00 \quad 0 \quad 0 \quad -2.00 \quad -1.00]$
$y_{23} = [-4.00 \quad -9.00 \quad -6.00 \quad -2.00 \quad -3.00 \quad -1.00]$

After adding this column wise, we have got-
$Sum = [-8.00 \quad -19.00 \quad -11.00 \quad -4.00 \quad -7.50 \quad -3.00]$
Now we need to multiply each of these value with our learning rate $alpha$
Now, we have got
$[-5.60 \quad -13.30 \quad -7.70 \quad -2.80 \quad -5.25 \quad -2.10]$
Now we have to add this vector with our current weight vector a to get new weight for next epoch
New $a = [-4.60 \quad -12.48 \quad -6.70 \quad -1.80 \quad -4.25 \quad -1.10]$
Now we can start new epoch
$Epoch2$:

y		a	$g = a^T y$	g<0
y_{11}	[1 1 1 1 1 1]	[-4.60 -12.48 -6.70 -1.80 -4.25 -1.10]	-30.93	misclassified
y_{12}	[1 1 -1 1 -1 1]		-9.03	misclassified
y_{13}	[16 25 20 4 5 1]		-549.03	misclassified
y_{21}	[-4.00 -6.25 -5.00 -2.00 -2.50 -1.00]		145.19	
y_{22}	[0 -4.00 0 0 -2.00 -1.00]		59.50	
y_{23}	[-4.00 -9.00 -6.00 -2.00 -3.00 -1.00]		188.33	

In this epoch we have also got 3 misclassified y. So, we need to calculate new weight for next epoch. The process of calculation is same as shown for $epoch1$.

If we do all the calculation accurately, then our new weight vector for next epoch will be-
New $a = [8.00 \quad 6.42 \quad 7.30 \quad 2.40 \quad -0.75 \quad 1.00]$
This iteration will end when we got all the values of column $g = a^T y$ positive.