Assignment 1C : Linear Perceptron

In the modern sense, the perceptron algorithm is used for learning a binary classifier called an threshold function, a function that maps its input x vector (features) to an output function f(x)

$$f(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{w} \cdot \mathbf{x} + b > 0, \\ 0 & \text{otherwise} \end{cases}$$

Where w is a vector calculator using the perceptron algorithm and b is the bias

In the algorithm, we update the weights based on the below function

$$w_i(t+1) = w_i(t) + r \cdot (d_j - y_j(t))x_{j,i}$$

Where r defines the learning factor in between 0 and 1 and generally it will be 0.1 as higher value makes the w value volatile
y(t) is the output value given by the function f(x) on the input vector x

We keep doing this update until the predetermined number iterations are completed or the iteration error given by the function below is less than or equal to the user defined threshold value

$$\frac{1}{s}\sum_{i=1}^{s} |d_j - y_j(t)|$$

Implementation of the code:

- Importing the python standard libraries like numpy, pandas, random, math
- Importing the data set from the csv and splitting into 3:7 ratio for test and training dataset
- Running the perceptron classifier
  Randomly initialize the w vector
  Running the outer loop on the classifier to check the whether predetermined iterations are completed or not
  Finding the y value and updating it based on the function and running the loop until required threshold is reached
- Checking the accuracy of the model

Analysis on Dataset 1:

The dataset is not linearly separable as there still data points which are misclassified (around 10 - 15)
W = [-48.54461400000811, -27.468915999997733, -32.53475439999668, -4.514892800000065]
b = 49.40000000000016,
The accuracy of the dataset is 99.02%

Analysis on Dataset 2:

The dataset is linearly separable as there are no misclassified points
W = [-0.17087702580000003, 0.2690911075999999, 3.1435274764]
b = 0.8
The accuracy of the dataset is 100%

The Dataset 2 is more separable than Dataset 1 as Dataset 1 has misclassified data points

Major Limitations of Perceptron classifier:

- The classifier is only useful for the binary classification
- The model can not work on the non linearly separable data