Machine Learning

Assignment 1 Report

Team members:

Nipun Wahi	Hrithik Kulkarni	Mir Ameen Mohideen
2018A7PS0966H	2018A7PS0278H	2018A7PS0487H

1C - Linear Perceptron :-

First the dataset is read and stored in a pandas dataframe and is then split into training and testing dataset in a 70:30 split. A column of 1 is also added to deal with w0(the bias term). Then the training dataset is used to run the algorithm and in each iteration the coefficients are updated until there are no misclassifications or till the number of iterations is less than 10^6. Finally the resultant coefficients and the testing data are used to find the accuracy of the model.

Accuracy of your model on both the datasets :-

dataset _LP_1.txt -

Training Accuracy is : 99.58333333333333%
Testing Accuracy is : 97.81553398058253%

dataset_LP_2.csv -

Training Accuracy is : 100.0%
Testing Accuracy is : 100.0%

Dataset which was more linearly separable :-

dataset_LP_2.csv is more separable as we get 100% accuracy, which means that there were 0 misclassifications and the data was linearly separable.

Major limitations of the Perceptron classifier :-

1. The perceptron classifier only works for linearly separable data, if the data is not linearly separable then it never reaches a point where all points are classified properly.
2. The algorithm can be only used for a binary classification problem, i.e the linear perceptron tells whether or not the sample belongs to that class(0 or 1). However we can use multilayer perceptron to overcome this problem