PERCEPTRON – A BINARY CLASSIFCATION ALGORITHM

ASHNA SWAIKA	***MUDIT CHATURVEDI***	***ROHAN SACHAN***
(2018A7PS0027H)	(2018A7PS0248H)	(2018B3A70992H)

INTRODUCTION

Perceptron is a binary classification algorithm, introduced by Frank Rosenblatt in 1957. This algorithm works only on linearly separable data. This is one of it's limitations since if a linearly inseparable data is given the algorithm goes into an infinite loop. The data may be linearly separable, but if it takes a large amount of iterations and if we stop the algorithm at some finite timestep before that (due to computational overhead) then we may never find out, that it is indeed separable.

MODEL DESCRIPTION

INPUT :

1. *Input matrix, '* $X_{N\,x\,M}$ *',* where - 'N' training instances each with 'M' features,
2. *Output vector, '* $T_{N\,x\,1}$ *'* which gives the correct labels of N training features.

OUTPUT :

1. *Weight vector '* $W_{1\,x\,M}$ *'* , which gives the boundary which linearly separates all training instances

MODEL :

During training, each training instance is classified by the formula $W^T X$ and if it is misclassified (1) then loss (2) is calculated for that training instance and subtracted from the model (3). This procedure is continued until all points are correctly classified.

$$t^n . W^T \phi(x^n) < 0 \qquad (1)$$

$$-t^n . \phi(x^n) \qquad (2)$$

$$W^{\tau+1} = (W^\tau)^T - (-t^n . \phi(x^n)) \qquad (3)$$

TESTING :

During testing accuracy is calculated based on fraction of correctly classified examples in total dataset. If (4) is positive then point is classified positive else negative.

$$W^T \phi(x^n) \qquad (4)$$

IMPLEMENTATION

We have been given two datasets namely – *dataset_LP_1.txt* and *dataset_LP_2.csv*. The dataset is divide in 70/30 ratio for training and testing . The 70% training examples are put inside the perceptron model to get the weight vector. After the weight vector has been calculated, testing is done on the remaining 30% of the dataset. Finally, *test accuracy* is calculated – alongside *number of iteration it took to train*, the *final weight vector* and *training accuracy*.

RESULTS

1. *DATASET :: dataset_LP_1.txt*

   WEIGHTS: [0.1772573 -0.16648173 -0.09337221 -0.10022434 -0.01652668]

   ITERATIONS: 1000000

   TRAIN ACCURACY: 0.98

   TEST ACCURACY: 0.99

2. *DATASET :: dataset_LP_2.csv*

   WEIGHTS: [0.18727955 -0.05347587 0.2191062 1.10302439]

   ITERATIONS: 42286

   TRAIN ACCURACY: 1.0

   TEST ACCURACY: 1.0

CONCLUSION

dataset_LP_2.csv is more likely linearly separable because both its training and testing accuracy is 100% and it took less than maximum number of iterations (10^6) to reach a weight which separates all points into positive and negative.

LIMITATIONS

Perceptron works only on linearly separable data. It goes inside an infinite loop for linearly inseparable datasets.