Single-layer Linear Neural Networks

Hasan ALQudah

Introduction:

The report implements three binary classifiers Perceptron, Adaline, and Stochastic Gradient Descent (SGD). The aim of this report is to compare the different classifier performance in 2 datasets. And to implement a multiclass technique called one versus all using binary classifier.

Datasets

The first dataset we test our classifier is iris dataset which has been published in UCI machine learning repository for small data. The data was originally produced in 1936 by biologist and statistician Ronald Fisher. The data contains 150 sample divided into 50 samples among three classes. Iris data set measures four features of Iris flowers spetal length and width as well as petals length and widths in centimetres. These measurements classify Iris flowers into 3 classes :Iris setosa, Iris virginica and Iris versicolor.

The other dataset is called seed dataset which has been conducted as an experiment study to classify 3 kind of wheat grains to three different kinds or classes: Kama, Rosa and Canadian, 70 elements each, randomly selected for the experiment. The dataset consists of 210 kernels or instances belongs to three wheat varieties, characterized by 7 geometric features, all of these real-valued continuous.

Experiment:

First we implement three binary classifiers Perceptron, Adaline, and Stochastic Gradient Descent (SGD) using Python. Classes programmed as 3 separate modules in addition to main module. The second step was to prepare data sets for training and classifying. Each datasets have been partitioned to 75% of data as training data and the remaining data have been considered as testing dataset.

Iris data classification:

Iris data has been trained and testing using the three algorithms (classifiers Perceptron, Adaline,and SGD). Table 1 shows the result of testing the data in different algorithms using Iris dataset. These information illustrates the convergence of the algorithm when eta=0.1 and number of iterations are greater than 4. Figure 1 and 2 shows theses attempts visually.

attempts	1	2	3
Success rate	100%	98%	98%
Error rate	0%	0.02%	0.02%
eta	0.1	0.01	0.001
Number of iteration	10	10	10

Table1: three attempts of perceptron algorithm in iriss data

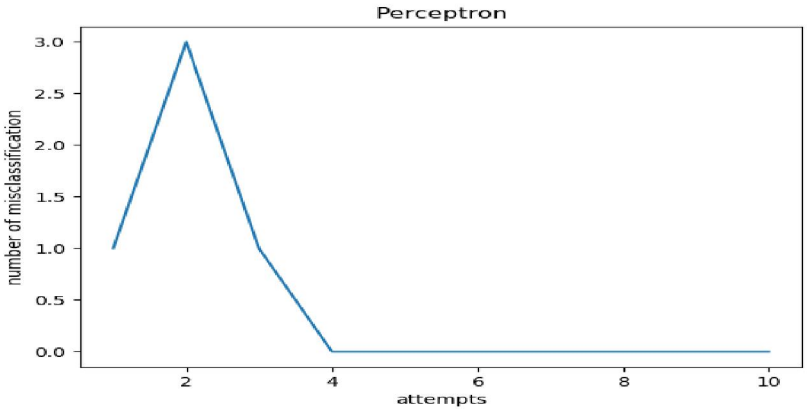

Figure:1 eta=0.1 and iteration 10

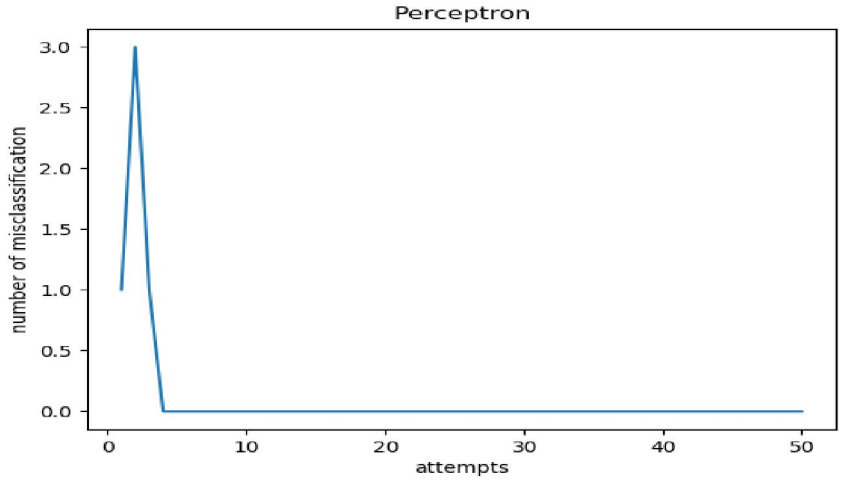

Figure2: eta=0.1 and iteration 10

Addaline: Table 2 shows the different attempt with Addaline algorithm in Iris data. It is clearly seen that when learning rate is small this increase the accuracy but it need more iteration to allow the mode to converge. Figure 3 and 4 shows the relation of eta and iterations number.

attempts	1	2	3
Success rate	0.33	0.56%	100%
Error rate	0.67%	0.44%	0%
eta	0.1	0.0001	0.0001
Number of iteration	10	10	100

Table2: three attempts of Addaline algorithm in iris data

Hasan A

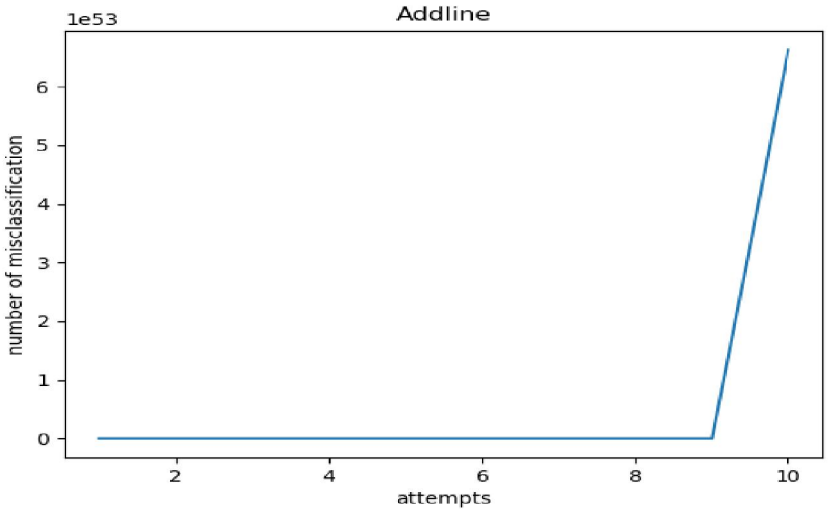

Figure 3: eta=0.1 and the iteration is 10

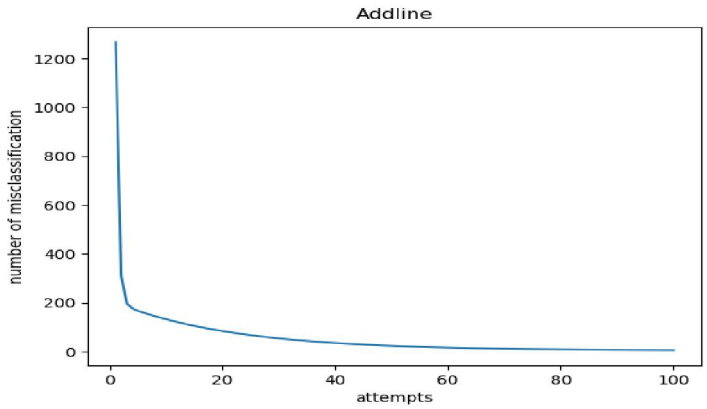

Figure 4: eta=0.0001 and the iteration is 100

SGD: Stochastic Gradient Descent in table 3 shows that the data need to converge take less number of iteration than in adaline figure 5 and 6

attempts	1	2	3
Success rate	0.33%	0.67%	100.0%
Error rate	0.67%	0.33%	0.0%
eta	0.1	0.001	0.0001
Number of iteration	10	10	100

Table3 three attempts of SGD algorithm in iris data

Single-layer Linear Neural Networks

Hasan ALQudah

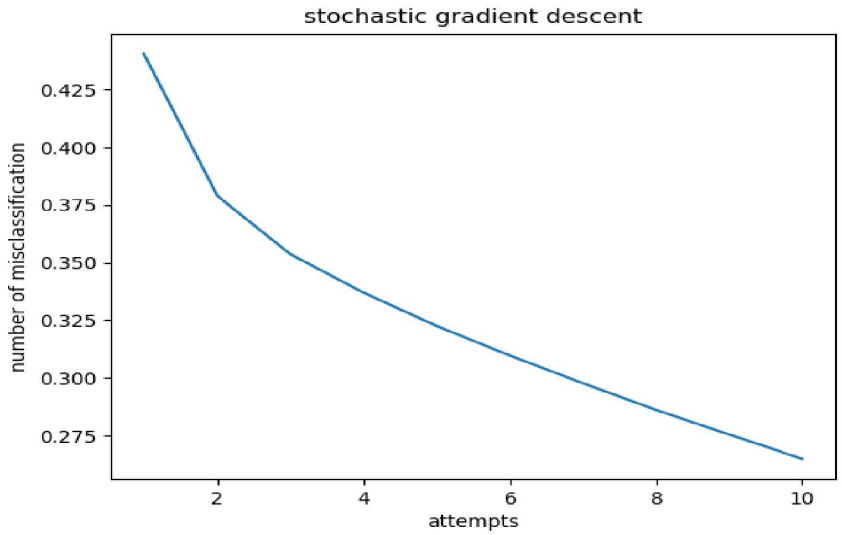

Figure 5: eta=0. 0001 and the iteration is 10

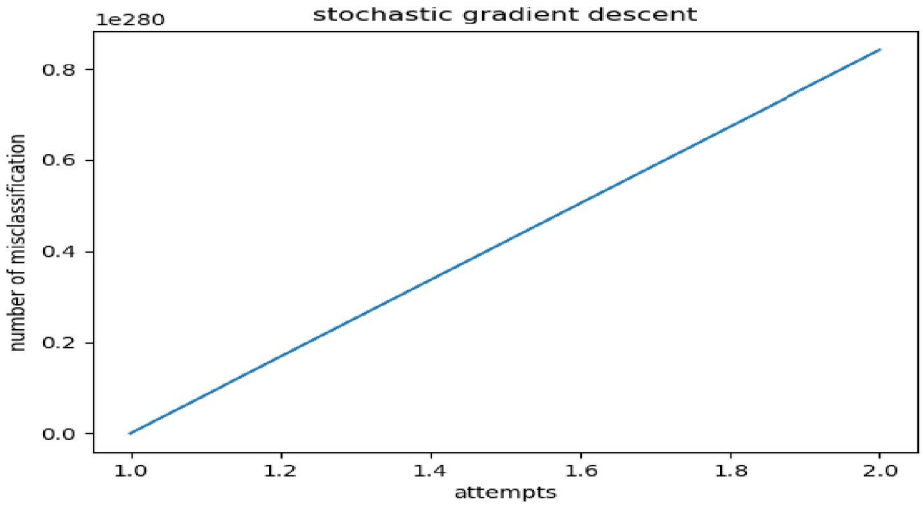

Figure 6: eta=0.1 and the iteration is 10

Single-layer Linear Neural Networks

Hasan ALQudah

SEED DATASET

The second experiments is to test a new three binary classifiers Perceptron, Adaline, and Stochastic Gradient Descent (SGD) with other data set that have more attributes and instances called seed data set. Dataset has been separated by 75% for training data and 25% for test data. Similarly, like iriss dataset the models shows same attitudes when we change the parameters of learning rate and iteration number. Figures

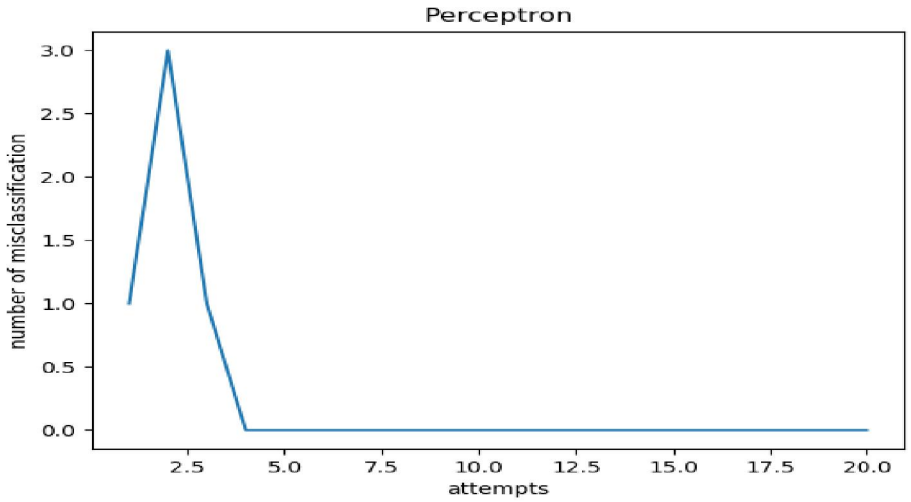

Figure 7 eta=0.1 and the iteration is 20 shows the convergence of data after 4 iteration

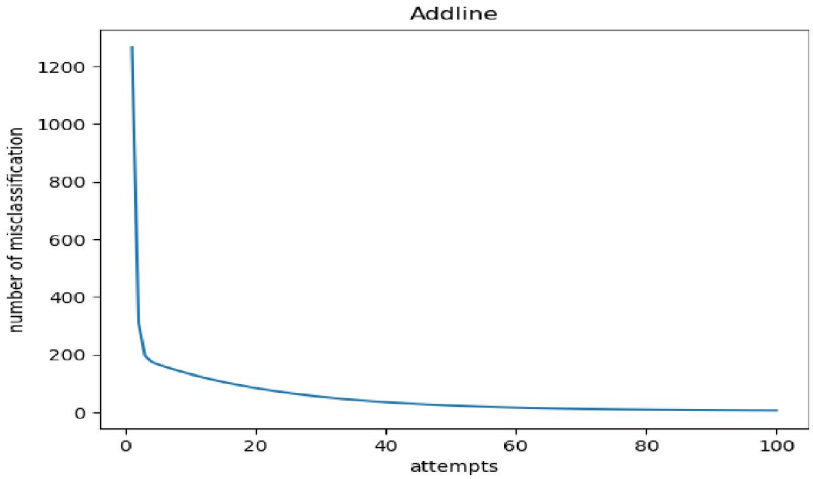

Figure 8 eta=0.0001 and the iteration are 100 iteration in addline

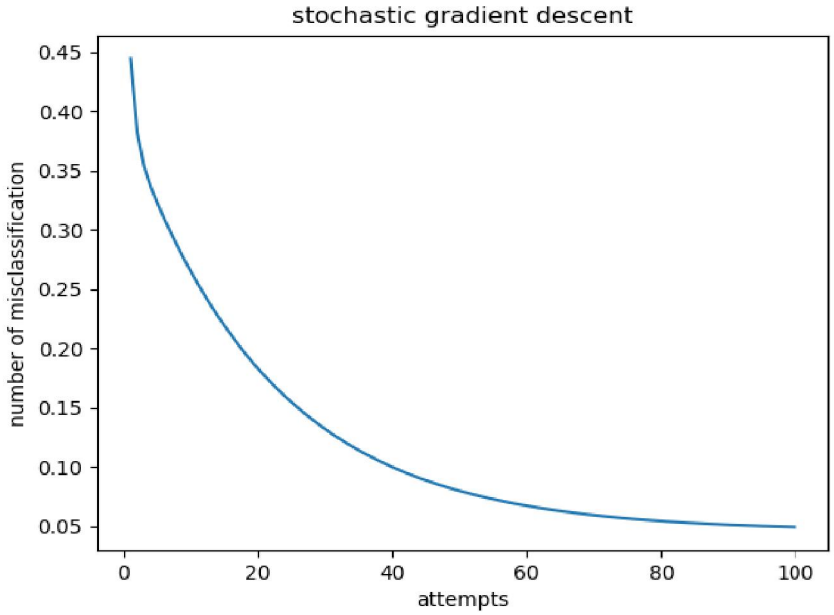

Figure 9 eta=0.0001 and the iteration are 100 iteration in SGD

algorithm	perceptron			addline			**SGD**		
attempts	1	2	3	1	2	3	1	2	3
Success rate	100%	98%	0.98%	0.33%	0.33%	100%	35%	0.33%	100%
Error rate	0%	0.02%	0.02	0.67%	0.67%	0%	0.65%	0.67%	0%
eta	0.1	0.001	0.0001	0.1	0.01	0.0001	0.1	0.01	0.0001
Number of iteration	10	10	100	10	100	100	10	100	100

Table 4:shows the result of running different classifier on the seed dataset

In summary, the numbers obtained in this test for different kind of binary classifier can be change according to change of the input parameter. Also, data pre-processing such as normalization and feature scaling can be enhanced by standardizing the training and testing files.

Multiclass classification using binary classifier (Stochastic Gradient Descent (SGD))

This strategies allows the binary classifiers to work with multi classes dataset. One of these in one versus all is that we assign every class to be positive instance and other class to be negative and make that interchangeably. That mean in our iris data we have 3 classes so we have to train 3 models. In every model we make class positive and other to be predict negative.

Single-layer Linear Neural Networks

Hasan ALQudah

The result after implementing this technique in our datasets we got the following result.

In iris data the output of program is with eta 0.1 and 10 iterations **i**

```
# of instances in class  1 that classified correctly are  15 from  15  with percentage of 1.0%
 # of instances in class  2 that classified correctly are  0 from  15  with percentage of 0.0%
 # of instances in class  3 that classified correctly are  15 from  15  with percentage of 1.0%
class 1 statistics:
success Rate=   1.0%
error Rate=     0.0%
class 2 statistics:
success Rate=   1.0%
error Rate=     0.0%
class 3 statistics:
success Rate=   1.0%
```

running the program in iris dataset using SGD classifier

```
please input eta0.01
please number of iteration20
 # of instances in class  1 that classified correctly are  15 from  15  with percentage of 1.0%
 # of instances in class  2 that classified correctly are  0 from  15  with percentage of 0.0%
 # of instances in class  3 that classified correctly are  7 from  15  with percentage of
0.4666666666666666%
```

Running on seed dataset

```
please input eta0.001
please number of iteration100
 # of instances in class  1 that classified correctly are  16 from  18  with percentage of 0.88%
 # of instances in class  2 that classified correctly are  10 from  18  with percentage of 0.55%
 # of instances in class  3 that classified correctly are  16 from  18  with percentage of 0.88%
class 1 statistics:
success Rate=   0.89%
error Rate=     0.11%
class 2 statistics:
success Rate=   0.89%
error Rate=     0.11%
class 3 statistics:
success Rate=   0.89%
error Rate=     0.11%
```