Andika William
16/398497/PA/17458

SLP Implementation in Python

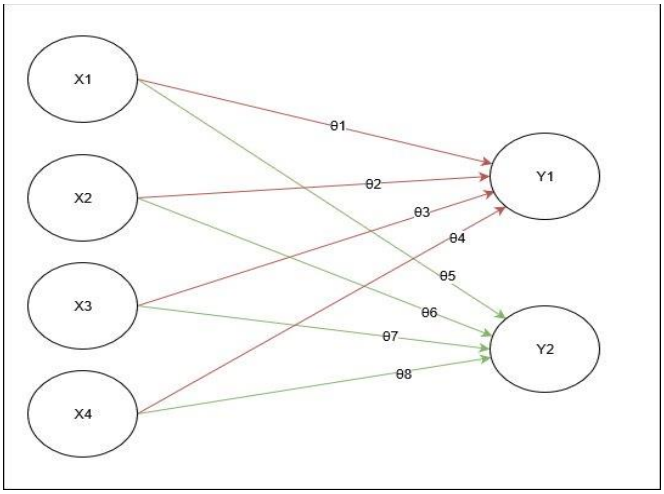
Case Explanation

In this assignment, we try to apply machine learning method of linear classifier by using the Iris data set. The Iris data set that is commonly used for practice. After training, the computer would be able to predict the species/type of the Iris from the length and width of both the petal and sepal.

```
5.1,3.5,1.4,0.2,Iris-setosa
4.9,3.0,1.4,0.2,Iris-setosa
4.7,3.2,1.3,0.2,Iris-setosa
```

Architecture of SLP

Each of the *x* values represent the sepal and petal values, while the *y* values represent the prediction value. Through several iterations we can improve the accuracy of the prediction.



Implementation

In writing the code, the language that was used in this assignment was Python. It was written in a text editor and was tested in the cmd.

https://github.com/andkwv/iris_linear_classifier

Experiment & Result

Running the code:

We start by determining the learning rate that we would like to use (0.1 or 0.8). In our first result the value 0.1 is used.

Learnning rate = 0.1

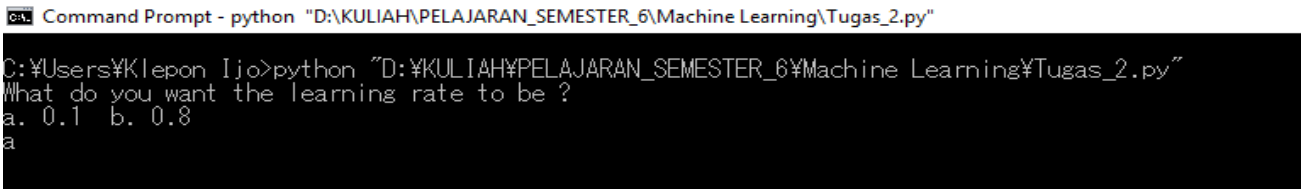


Figure 1.

Each iteration of epoch there are 150 rows that are processed. Every loop will calculate the error average and accuracy average of that epoch. Both the error and the accuracy are calculated by summing all the values and dividing them by 150. Accuracy in this case are only counted if both prediction values are correct.

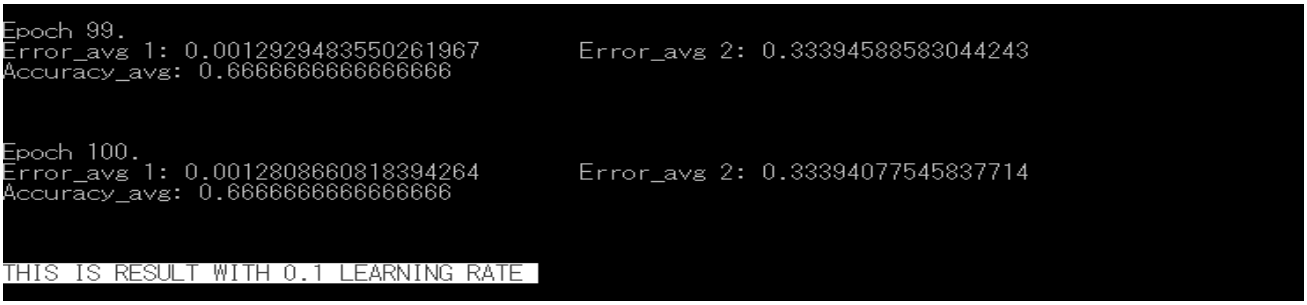


Figure 2.

After getting the result, we can proceed to graph the plot of the error values throughout the epochs iteration to help visualize the data. Supposedly after training the data, it is expected that the error should decrease while accuracy should increase.

In the first figure below, we can see the average values of the first prediction error. It is clear that the error value decreases after each iteration. In the end values converge into a small enough value, below 0.01, that it can be considered negligible. On the other hand the second prediction error does show similar trend to the first error, though with a slight increase decrease changes in convergence. However, the second error seems to always converge on a much larger value, around 0.10.

One of the things to note is that the initial values being used for the thetas here are random. As such, every run of process will result in a different outcome. In this case however, the pattern always shows that the first category prediction error always converges around 0.01, while the *second* category prediction converges around 0.10

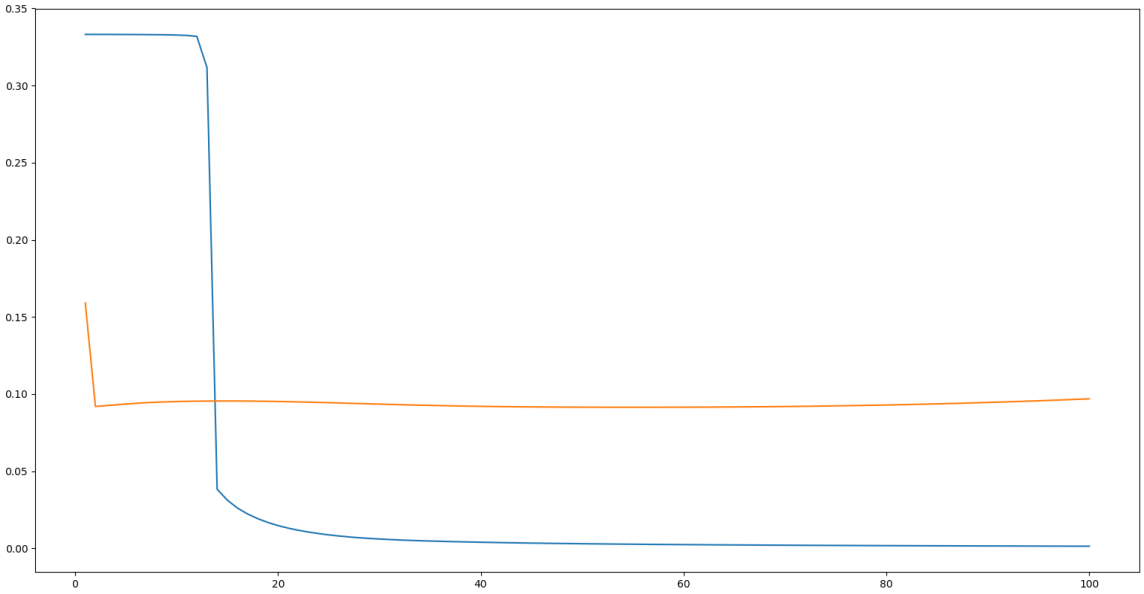


Figure 3.

Then we have the accuracy rate. Here we expect the result to be an increase of accuracy value over each iteration. The figure shows that it does experience an improvement of value over time. The accuracy of the prediction is in fact becoming increasingly accurate. However, what is notable is that while the values here seems to converge on a middle value of around 0.9, we can see that the values are slightly fluctuating or have even decreased. It is possible that optimum has already been reached, and a case of overfitting is happening to the system.

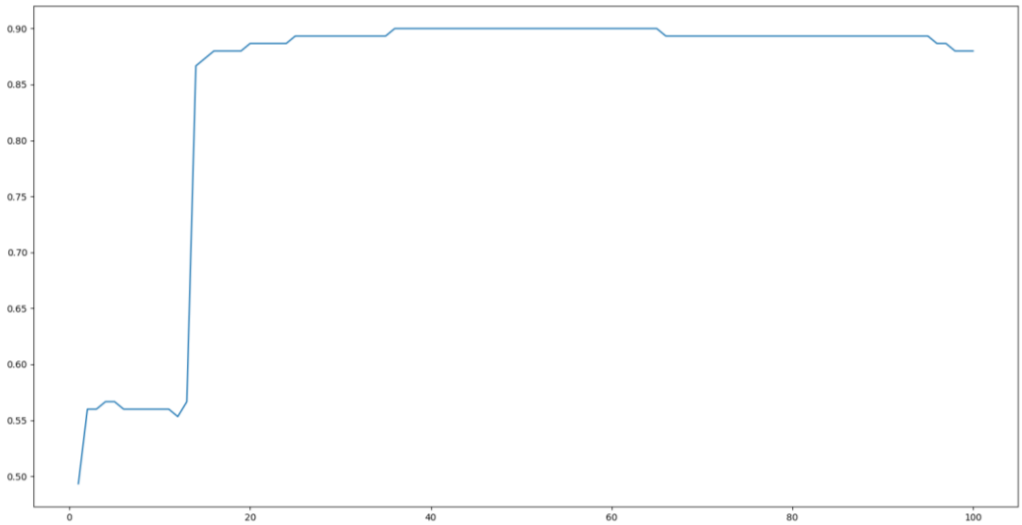


Figure 4.

Learning rate = 0.8

This time we differentiate the learning rate of the dtheta into 0.8. The changing of learning rate will generally help cases when our initial predicted value are actually very far from the target value. In general the bigger the learning rate, the bigger the jump between each iteration.

For the first error we expect to have an error that will decrease overtime the same as the previous experiment. We could see from the Figure 4 there are bigger jumps of values at the beginning, though it shows the decreasing trend. We see that a convergence successfully occurred in the first prediction while the second error experienced big fluctuations.

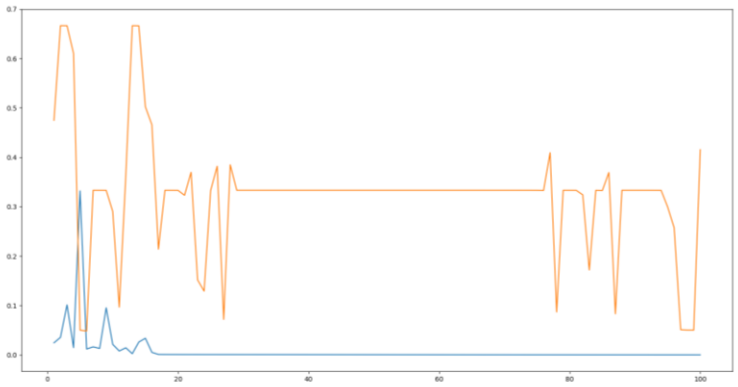


Figure 5.

For the accuracy, the values seems to fluctuate and arrived at a slightly worse result compared to the previous learning rate. Fluctuations can be partly attributed to overshooting of the values, which causes the calculation to not reach an optimum.

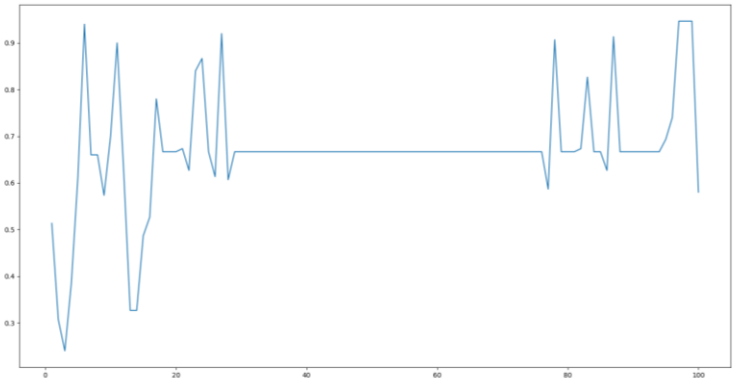


Figure 6.

In conclusion, the accuracy of the prediction will become increasingly accurate overtime. This also means that the rate of error will also decrease. In comparing the two learning rates we see the difference in how the system reaches accuracy. A miscalculation on one of the predictions will have an impact on the accuracy overall.