Machine Learning Algorithms

Assignment 1 Report

Name:	Omar Adel Ahmed Ahmed Amr
ID:	28-10999
Tut. No.:	T-08

1. Description of the perceptron structure:

The structure of the algorithm represents a Single Layer Perceptron. There are two input nodes plus an extra node which represents the bias node with fixed input equal to -1. The first input node represents the Sepal length of a planet and the other input represents its Petal length. Concerning the output, there is only one output. There are three weights, one for each input. Sepal weight is the weight of the sepal input, petal weight is the weight of the petal input and bias weight is the weight of the bias node. The following Figure demonstrates the structure of the inputs and output of the algorithm.

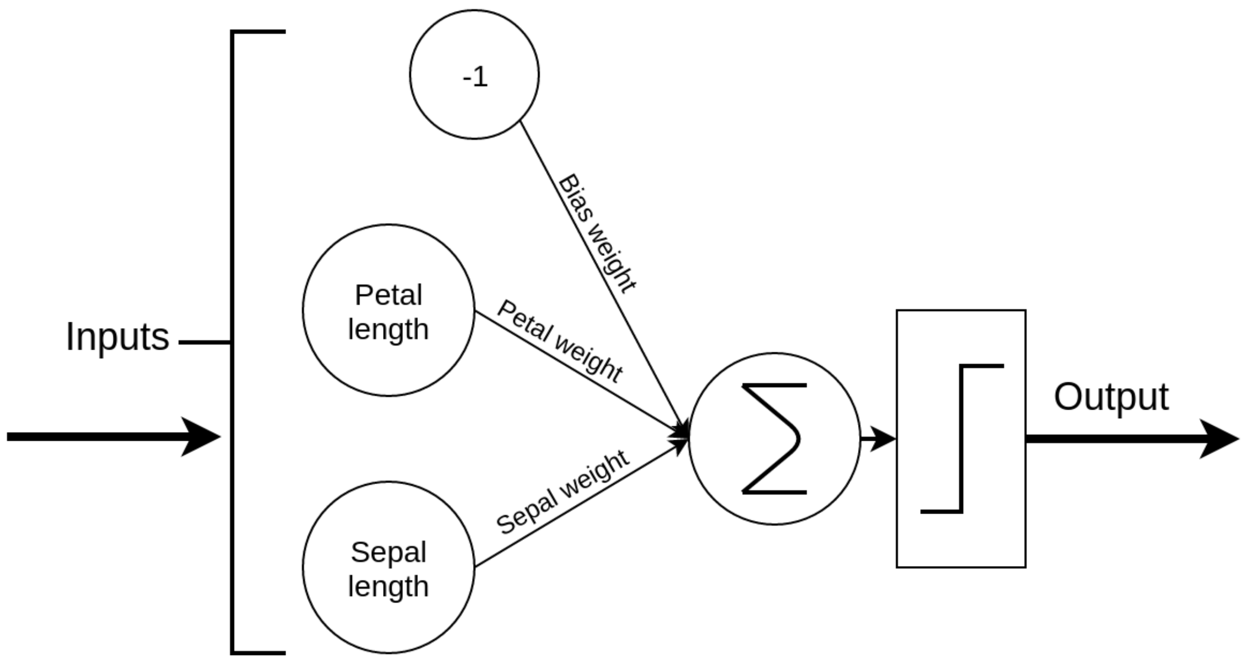

Figure 1

2. Training the perceptron implementation using the provided dataset.

a. Plotting Error values versus the instance indexes:

1- The following figure is a plot for the value of error during the training process on the y-axis versus the index values on x-axis. The values on the x-axis is ranged from 0 to 100 which represents only one iteration of the training process that is why the difference in the results is not very significant. However, a slight change can be spotted concerning the error values. The difference between the values of the error is getting smaller and a bit closer to the zero value on the y-axis. Figure 2 represents the value of error without calculating the threshold. On the other hand, Figure 3 will show the same representation as Figure 2 with only one difference. This difference is represented in applying the threshold then calculating the error.

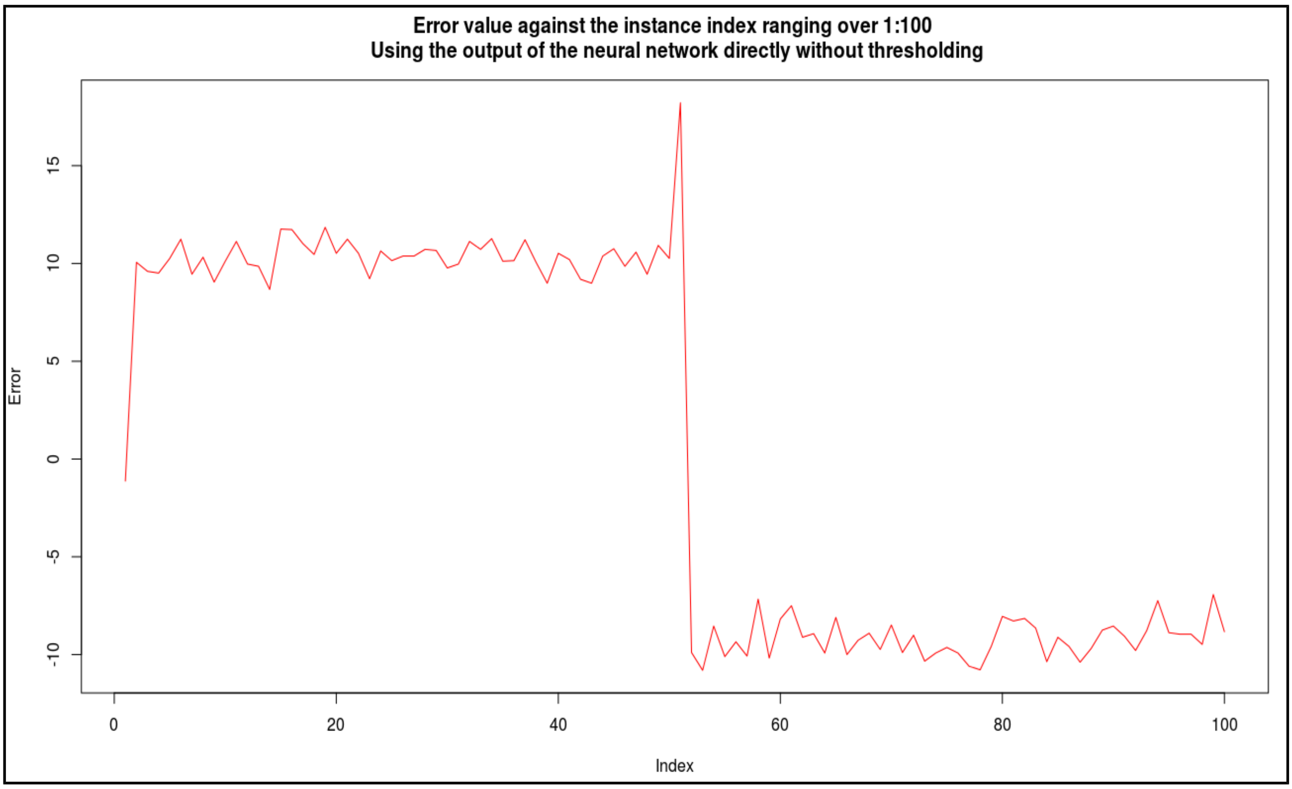

Figure 2

Figure 3 demonstrates the error change after applying the threshold to limit the values to either 1 or -1. It can be observed that at the beginning of the training during the first iteration there exist some errors however the number of these errors decreased at the end of the iteration. Next, in point b, a figure of the whole training process is shown to show the significant change in the error values.

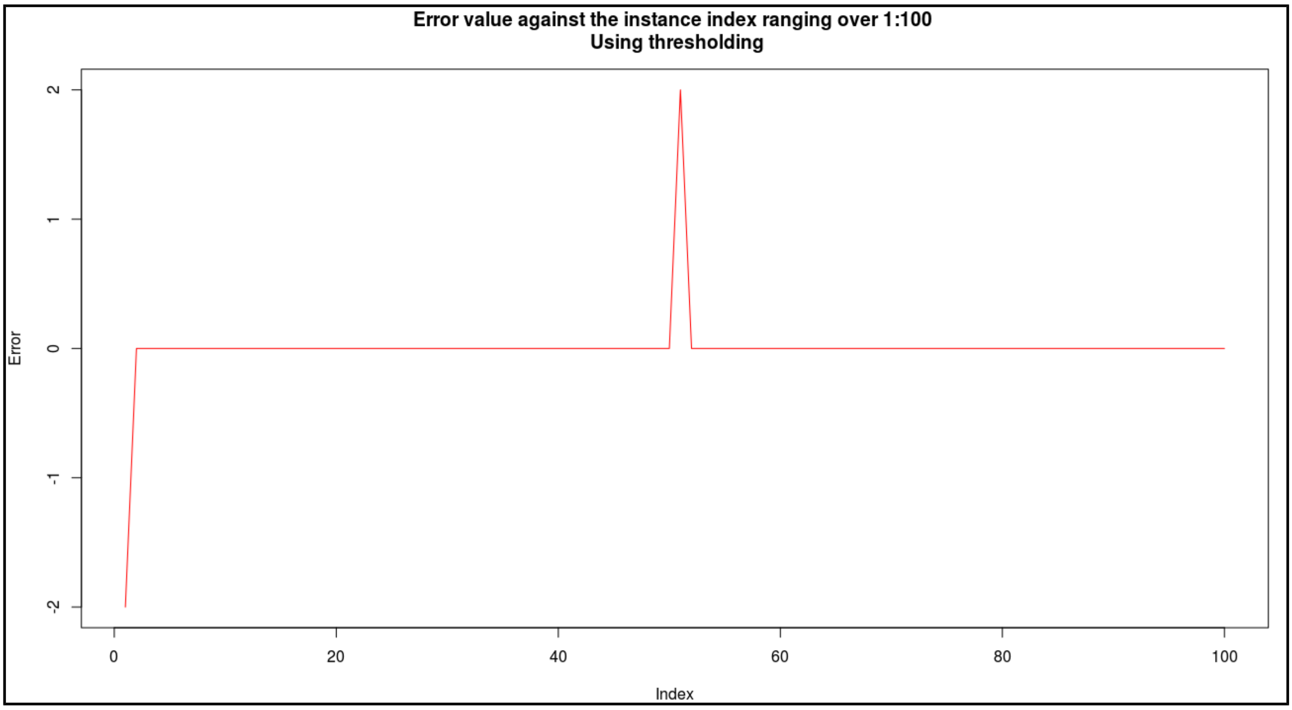

Figure 3

b. <u>Observing how the error value changes throughout the training process:</u>

As mentioned above, the values decreased along the training process as a result for continuously adjusting the weight of the inputs. The previous two figures were meant to highlight this enhancement in the performance, however their representation was weak because it did not represent the whole training process. The following two figures can significantly show the enhancement of the performance of the algorithm.

Figure 4 represents the error value against the instance index during the complete training process which consists in this case of 6 iterations on the data in order to reduce the error values. During the first 4 iterations the value of error was large compared to the the last 2 iterations. In the last iteration the value of error was so small that when the threshold is applied this small error is eliminated that is why the learning process stopped after the 6th iteration.

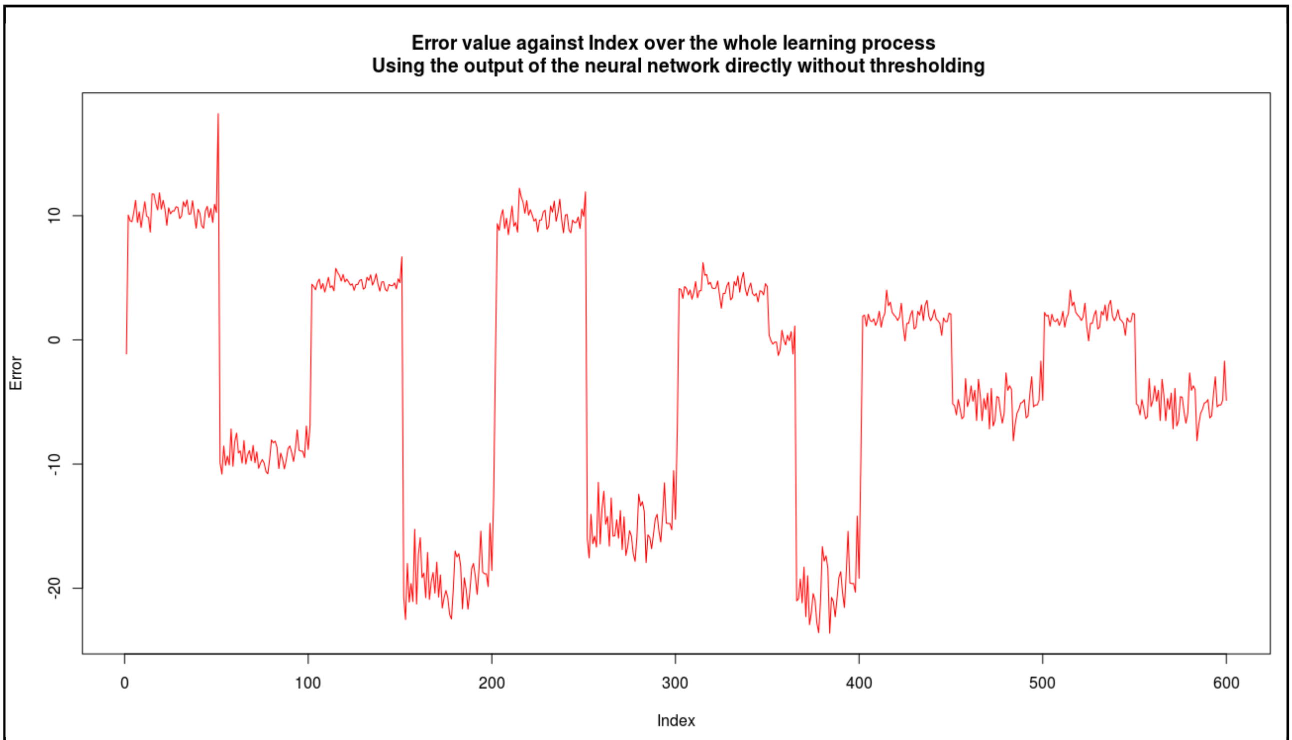

Figure 4

The following figure represents the same comparison but after applying threshold to the error values. The figure shows that the 6th iteration was completely free of errors which means the the learning process was successful which caused the algorithm to stop the learning process.

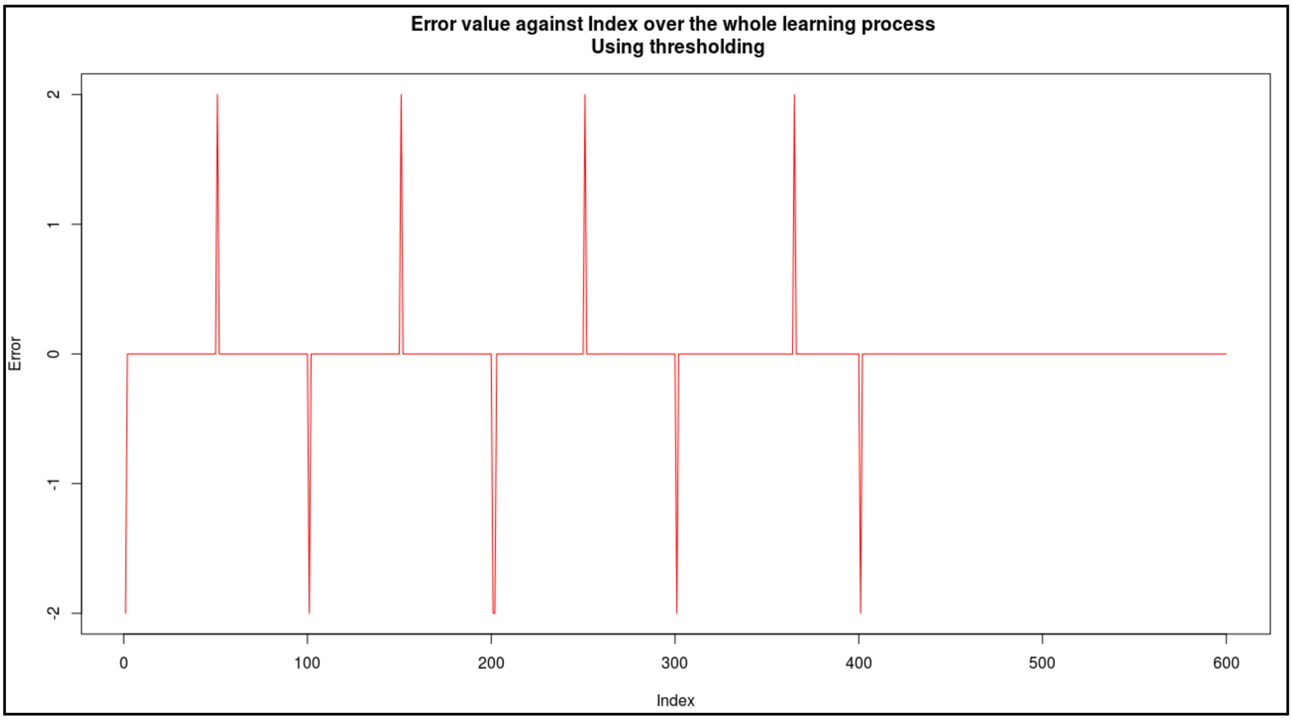

Figure 5

3. Modifying the learning rate:

In my implementation, changing the learning rate from 0.2 to 1.0 did not bring about any changes in the number of iteration required to complete the learning process as the algorithm kept looping for 6 times so that the error is eliminated. Generally, changing the value of learning rate affects the number of iteration required to complete the training process as the less the learning rate is the more iteration it takes to complete the learning. Changing the learning rate affected the values of the weights of the inputs as the learning rate controls how much the weights change after each error. It is worth to mention that the smaller the learning rate is the more stable the neural network. This can be observed by comparing the plot of the error in case of using 0.2 and 1.0. The reason why smaller values are more stable is because of the gradual change in the values of the weights which takes longer time to train the network, but this results in a more stable and noise resisting network.
To illustrate the effect of learning rate, the network was tested with learning rate = 0.0001 which required 25 iterations to complete the learning process.

4. Using test mode to classify test cases:

a. Yes, the trained algorithm managed to classify the provided 3 test cases successfully.

b. In the scope of the provided test cases, the algorithm was accurate and classified each case to its correct class. However, the database used to train the model is not large enough compared to the size of databases that are usually used to train neural networks. This means the algorithm may fail to classify another test cases correctly. But as mentioned, in the scope of the provided cases it was accurate.

c. NA, the algorithm did not make any wrong classifications.

5. Exploratory data analysis:

5.1 Petal length analysis:

a. Figure 6 represents a plot to shows the relation between the Petal length on the x-axis and the Iris class on the y-axis.

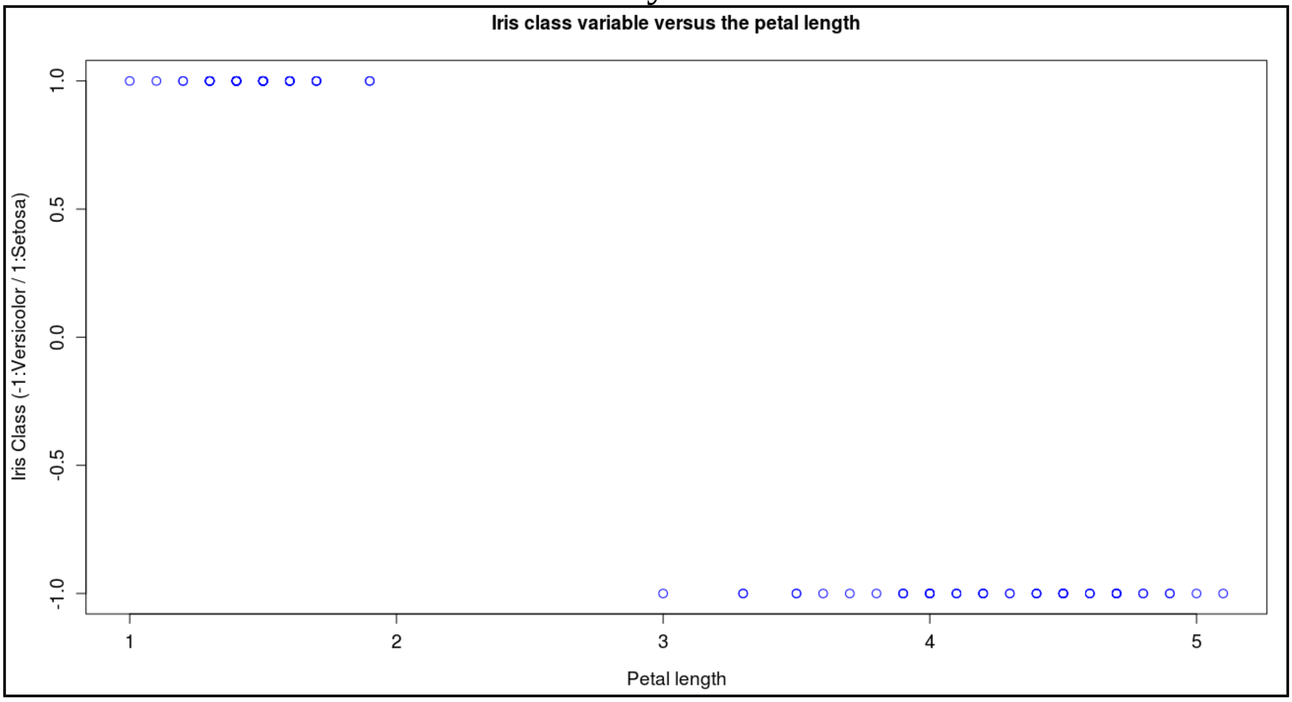

Figure 6

b. R has two built-in functions that can be used to calculate the covariance and the correlation between variables. These functions are cov(x, y) and cor (x, y). Using these functions the output was as follows:

Covariance = -1.413131
Correlation = -0.969902

c. Figure 7 represents a plot that shows the relation between the Sepal length on the x-axis and the Iris class on the y-axis.

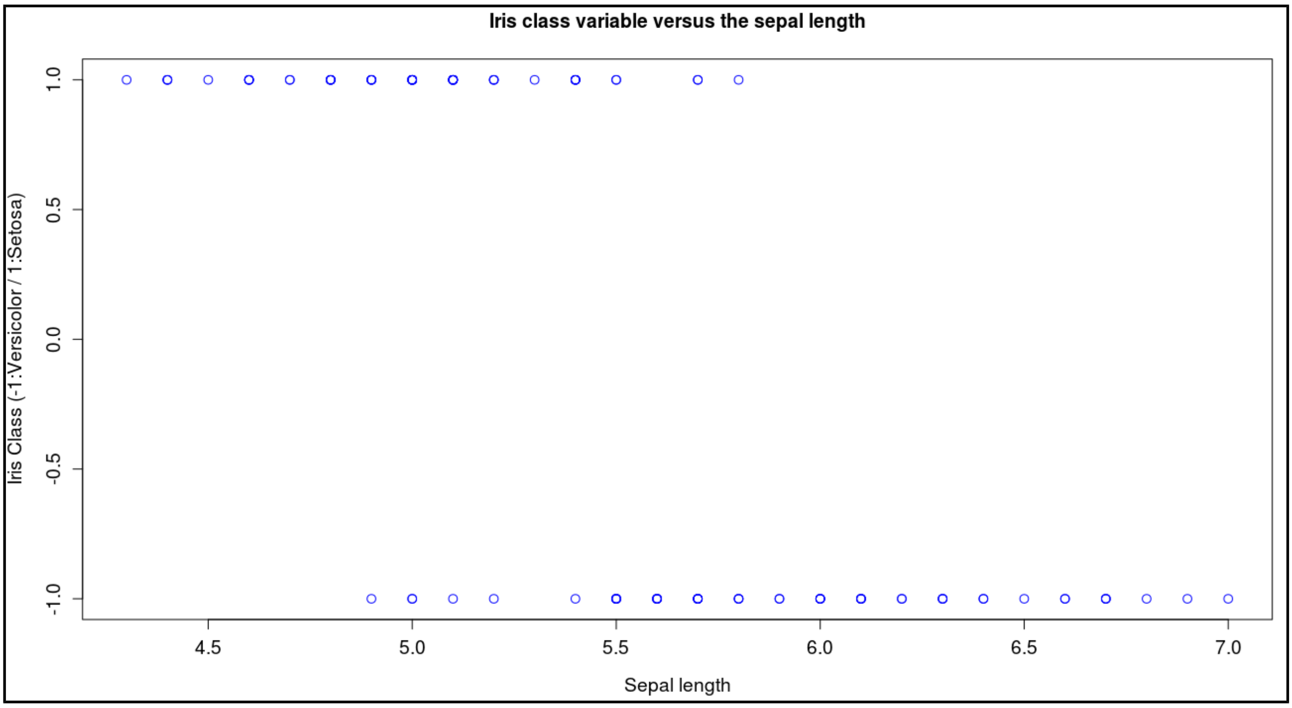

Figure 7

Covariance = -0.469697
Correlation = -0.72829

d. It would be meaningful to explain the meaning of covariance and correlation in order to define the relationship between the variables.
The covariance is a measurement of how much 2 variables are dependent on each other. Correlation also represents the strength of the relation between variables, however it guarantees that the value is limited between 1 and -1 whatever units used.

From figure 6 and 7 and also the result of the covariance and correlation of the 2 relationships, it can be deduced that petal length has more effect on determining the class of the plant than sepal length. But generally, since the value of the correlation of both relations is close enough to -1.0, this leads to making the process of training the perceptron much easier.