

University of Tehran

College of Engineering

School of Electrical and Computer Engineering

**Intelligent Systems Course**

**Computer Assignment #1**

**Under the Supervision of: Dr. Hosseini**

**Golmehr Khosrokhavar**

**Student ID: 810198507**

Khosrokhavar.g@gmail.com

**Jan 2023**

Contents

[**Question 1:** 3](#_Toc147270763)

[Part A) 3](#_Toc147270765)

[Part B) 4](#_Toc147270765)

[Part C) 4](#_Toc147270765)

[**Question2:** 8](#_Toc147270763)

[Part A) 8](#_Toc147270765)

[PartB) 9](#_Toc147270765)

[PartC) 10](#_Toc147270765)

[**Question 3:** 11](#_Toc147270764)

[PartA) 13](#_Toc147270765)

[Part B) 15](#_Toc147270765)

[**Reference** 15](#_Toc147270763)

**Question 1: Binary logistic regression**

It is suggested to write a summary of the process of solving the question, its purpose and its requirements in this section. In your report, write Persian texts using B Nazanin font and English texts using Times New Roman font.

**Part A:**

Rewrite:

|  |
| --- |
|  |

**Part B & C:**

Calculation of descending gradient and calculation of accuracy:

Considering the initial values ​​of zero and Learning rate = [0.005, 0.01, 0.1, 1], the graph of changes of J in terms of the number of repetitions per training data is drawn below.

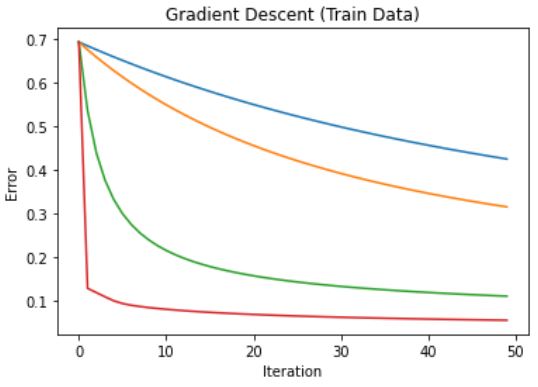


Figure 1-1

The percentage of correctness and accuracy of this implementation for different learning rates will be as below:

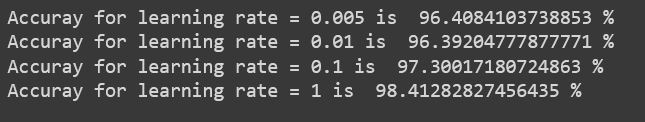


Figure 2-1

The graph of changes of J according to the number of repetitions per test data is drawn below:

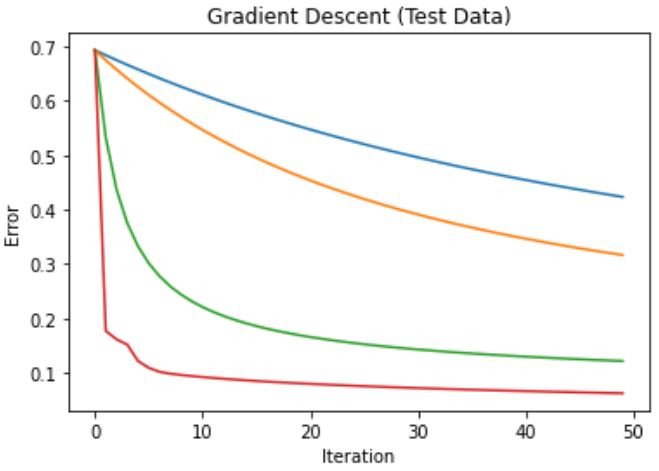


Figure 3-1

The percentage of correctness and accuracy of this implementation for different learning rates will be as follows:

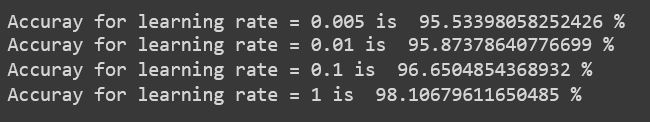


Figure 4-1

**Part C:** Stochastic Gradient Descent

At first, for batch size = 1, we will have:

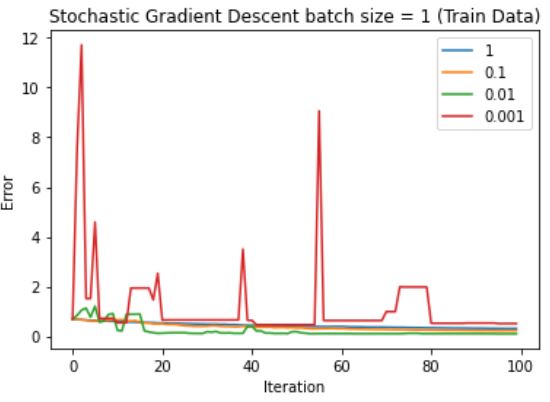


Figure 5-1

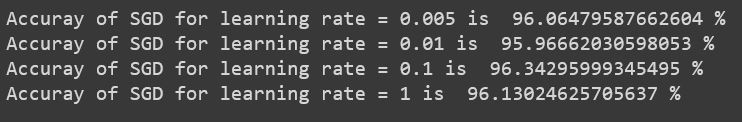


Figure 6-1

Then we repeat the same steps for the test data.

The percentage of correctness and accuracy of the answers of the test data are as follows:

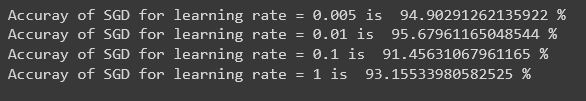


Figure 7-1

Next, we repeat the steps for batch size = 100:

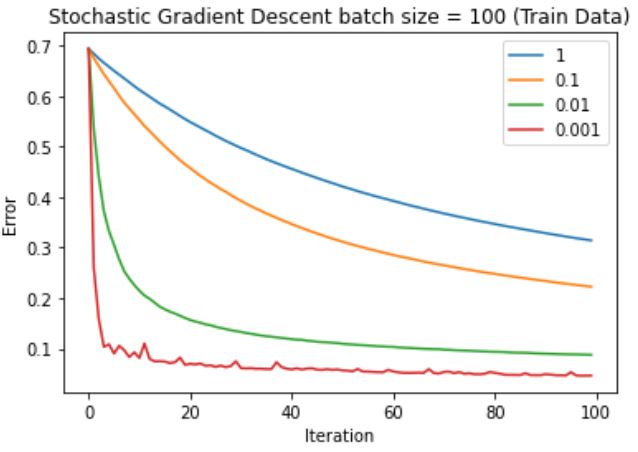


Figure 8-1

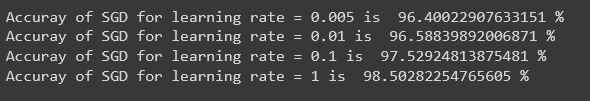


Figure 9-1

**Question 2: Optimization in non-convex functions**

**Part A:** Newton's method (analytical)

|  |
| --- |
|  |

We start from the point (0,0) and according to Newton's method we will have:

|  |
| --- |
|  |

**Part B:** Newton's method (simulation)

In this section, with the help of Newton's method and genetic algorithm, we want to find the minimum point for the given function.

The coordinates of the minimum point are found and the value of the function at that point is determined in the following figure:

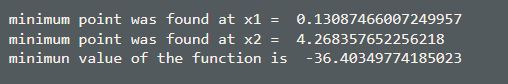
**

Figure 2-1

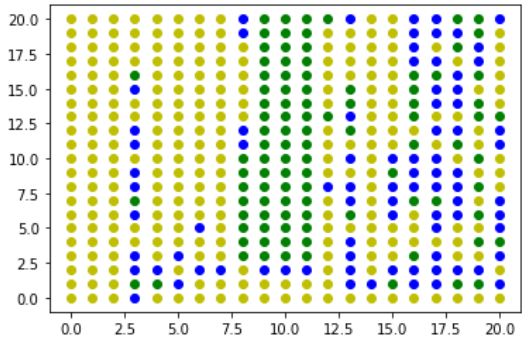


Figure 2-2

**Part C:** meta-heuristic method

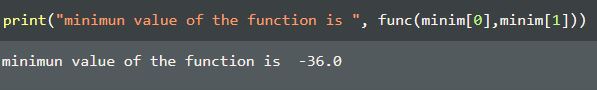


Figure 2-3

As can be seen from the image above, the minimum value of this function calculated by the genetic algorithm method is close to the answer of part b with an acceptable approximation.

**Question 3: Support vector machine**

**Part A:** Analytical

1. Consider the following formulation for the support vector machine classification:

|  |
| --- |
|  |

What result can be obtained if the value of becomes zero for a sample?

To minimize the margin, we must satisfy the following condition:

|  |
| --- |
|  |

which is equivalent to the following condition:

|  |
| --- |
|  |

Practically,  becoming zero means that when calculating

, the value of the term was less than zero Is. This means that the above condition is not fulfilled.

This means that some data values ​​are placed between the marginal lines and

Suppose we have three support vectors

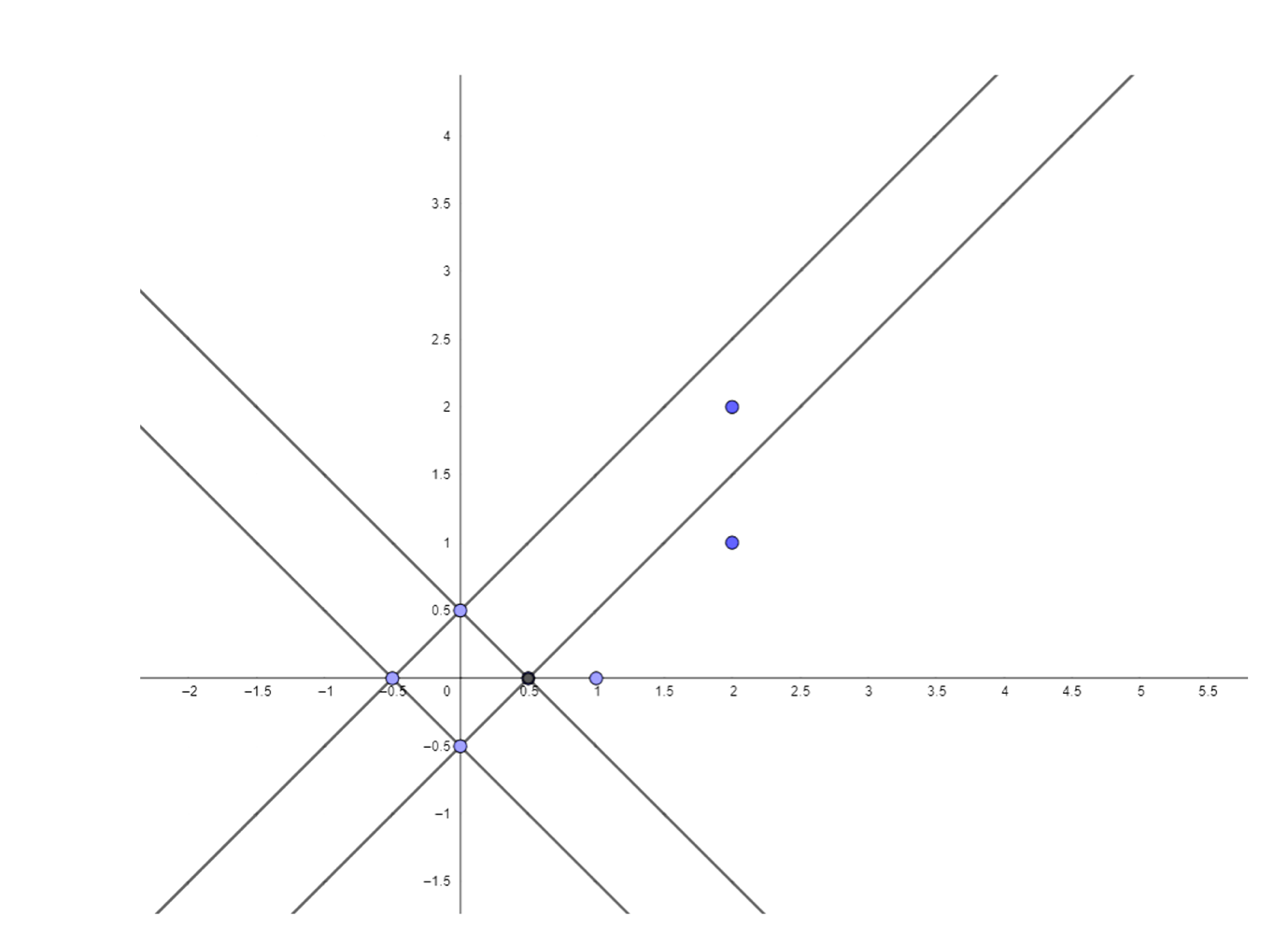


Figure 2-4

As it is clear from the above coordinates, the optimal decision matrix for separating the positive and negative classes from each other is the third case.

The line conveniently separates positive and negative class points from each other.

In the case of other lines, all the defects fall on one side of them, and in Saman, there is not even a proper margin on both sides of the dividing line.

**Part B:** Implementation

In this part, we are going to classify a data set with the help of support vector machine.

The data distribution will initially be as follows:

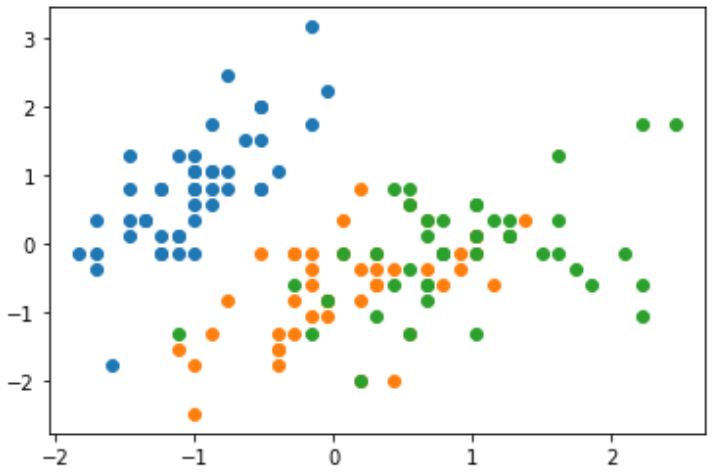


Figure 2-5

Then by implementing the support vector machine using method one in comparison with the others, we define different areas. The accuracy and details of this classification are as follows:

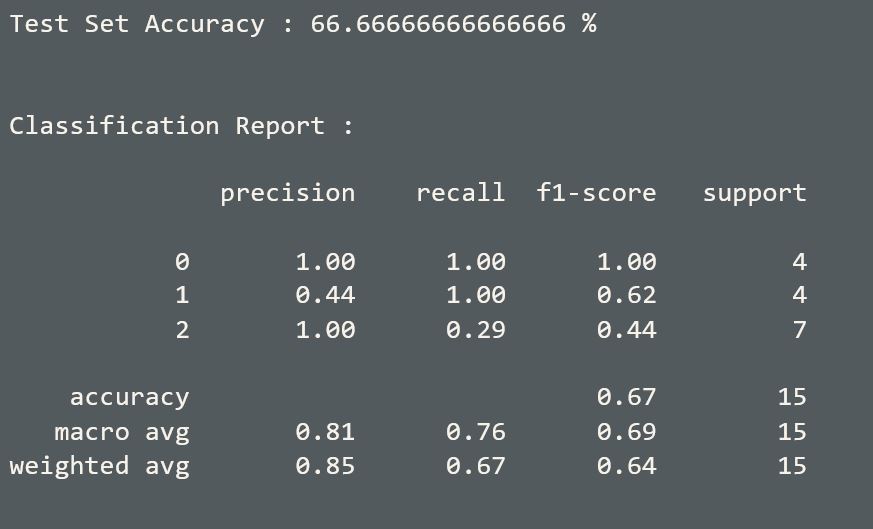
**

Figure 2-6

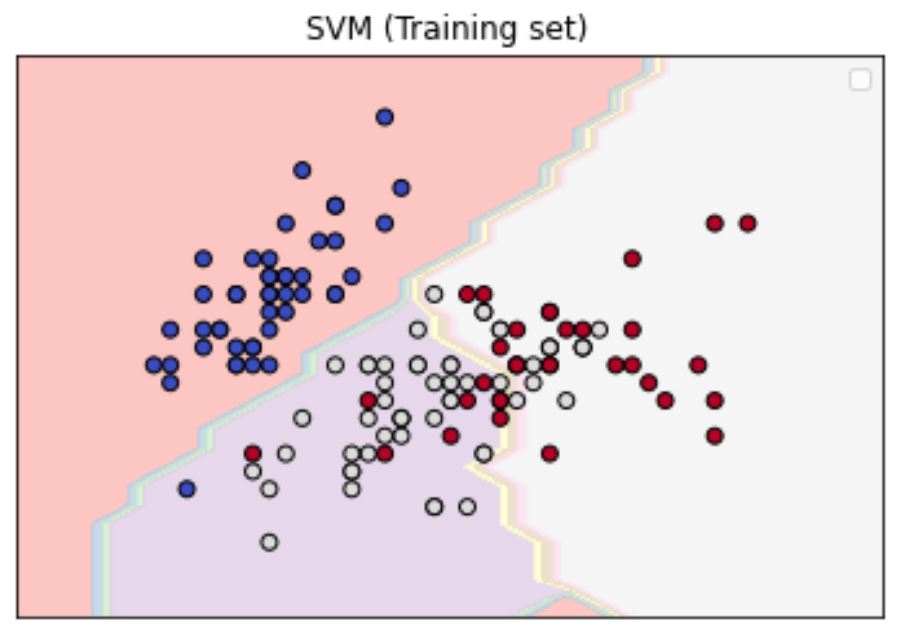
**

Figure 2-7

The confusion matrix will be as shown below:

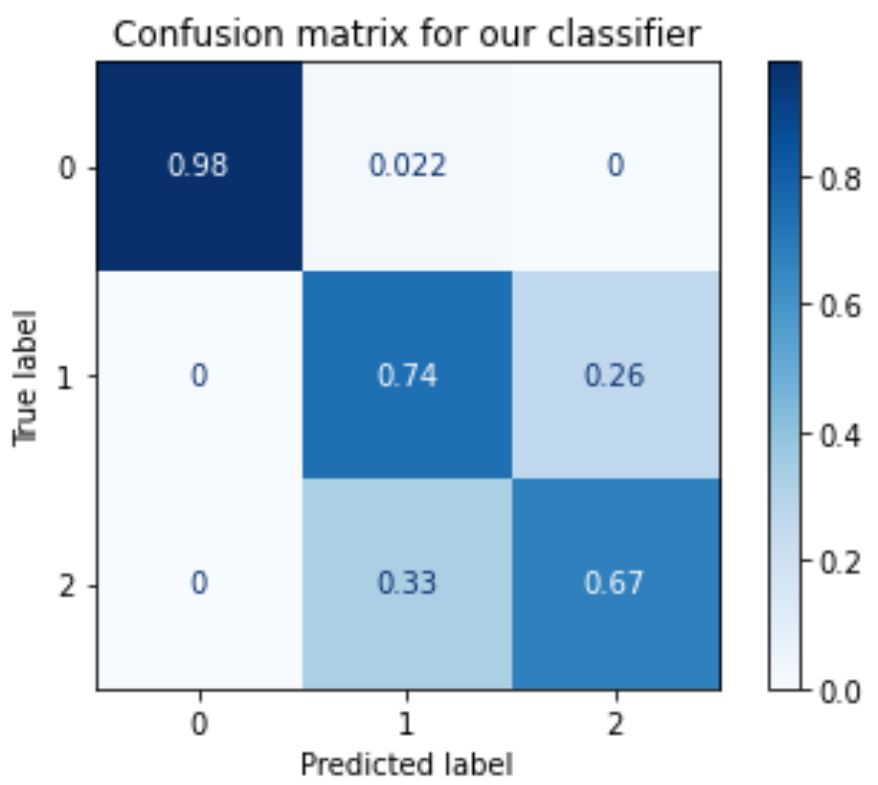


Figure 2-8

**References:**

* <https://scikit-learn.org/stable/modules/sgd.html>
* <https://towardsdatascience.com/complete-step-by-step-genetic-algorithm-from-scratch-for-global-optimization-6fee5c55dd3b>
* <https://towardsdatascience.com/an-overview-of-the-gradient-descent-algorithm-8645c9e4de1e>