COMP 421X HW02: Discrimination by Regression

Deadline: March 25, 2019, 11:59 PM

In this homework, you will implement a discrimination by regression algorithm for multiclass classification in R, Matlab, or Python. Here are the steps you need to follow:

1. Read Section 10.8 from the textbook.

2. Load the iris data set executing the following lines:

```
data("iris")
X <- as.matrix(iris[,1:4])
y <- as.numeric(iris[,5])
```

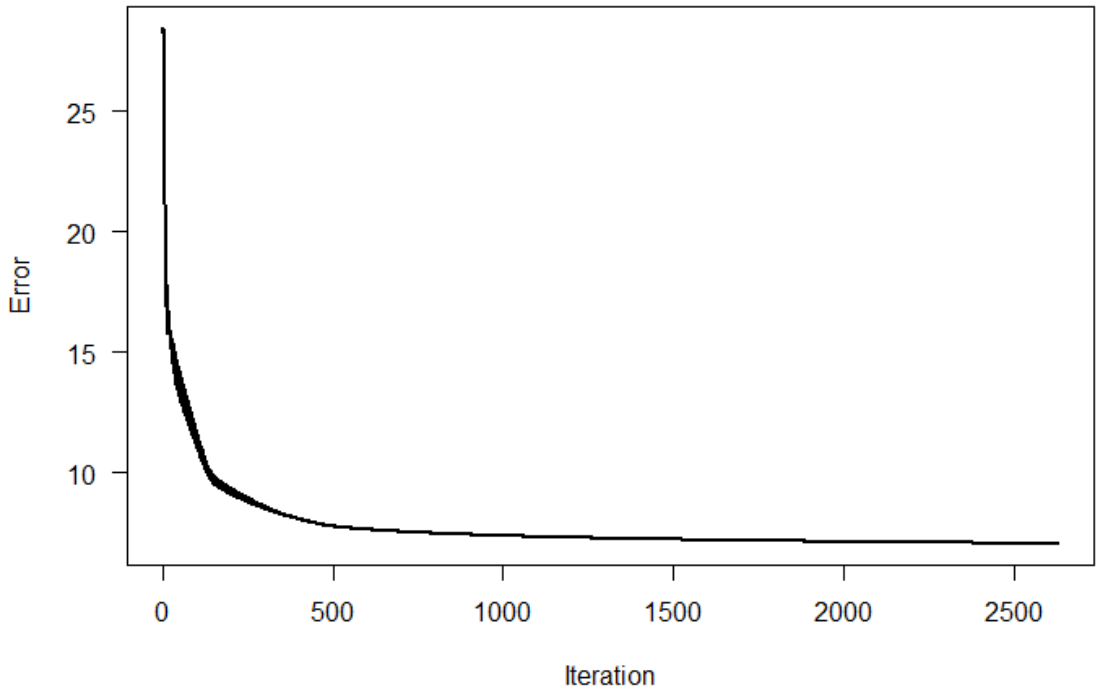
   You are given a multivariate classification data set, which contains 3 different species of iris flower (Setosa, Versicolour, and Virginica). There are 50 observations from each class with 4 features (sepal length, sepal width, petal length, petal width).

3. Divide the data set into two parts by assigning the first 25 samples from each class to the training set and the remaining 25 samples to the test set. You will have 75 data points in your training data set and 75 data points in your test data set.

4. Learn a discrimination by regression algorithm using the sigmoid function for this multiclass classification problem. You can use the following learning parameters.

```
eta <- 0.01
epsilon <- 1e-3
set.seed(521)
```

5. Draw the objective function values throughout the iterations. Your figure should be similar to the following figure.



6. Calculate the confusion matrix for the data points in your training set using the discrimination rule you will develop using the learned parameters. Your confusion matrix should be similar to the following matrix.

```
            y_train
y_predicted  1  2  3
          1 25  0  0
          2  0 24  0
          3  0  1 25
```

7. Calculate the confusion matrix for the data points in your test set using the discrimination rule you will develop using the learned parameters from the previous step. Your confusion matrix should be similar to the following matrix.

```
            y_test
y_predicted  1  2  3
          1 25  0  0
          2  0 23  2
          3  0  2 23
```

What to submit: You need to submit your source code in a single file (.R file if you are using R, .m file if you are using Matlab, or .py file if you are using Python) and a short report explaining your approach including your results but not any part of your code (.doc, .docx, or .pdf file). You will put these two files in a single zip file named as ***NAMESURNAME.zip***.

How to submit: Upload the zip file you created to the corresponding Blackboard Assignment. Please follow the exact style mentioned and upload a zip file named as ***NAMESURNAME.zip***. Submissions that do not follow these guidelines will not be graded.

If you have any questions please email cak14@ku.edu.tr with the subject line ***Intro2MachineLearningHW02.*** You could also use Blackboard forum for your questions.

Late submission policy: Late submissions will not be graded.

Cheating policy: Very similar submissions will not be graded.