Foundations of Data Mining ($2IMM20$)
Homework Assignment $3A$
MLP

Decebal Mocanu

Release date: October 9, 2019.
Guided Self-Study session: October 16, 2019.
Deadline: November 1, 2019.

Preamble

This homework has to be solved using Python. Please submit your solutions in a *.zip file[1] through Canvas. The solution needs to contain a report in the pdf format to discuss your approach and a Python file which can be easily run. State clearly in the report the group number and the group members.

Problem

In this homework you are requested to implement from scratch a MultiLayer Perceptron (MLP). This MLP model has to perform binary classification on the dataset uploaded in Canvas. You can train the model on the data from the file "HW3Atrain.xlsx" and validate the model on the file "HW3Avalidate.xlsx"[2]. In the files, each row represents a data point. The first two columns (x_o, x_1) represent the input features, while the third column (y) represents the class label (0 or 1). Further on, x_o and x_1 are collected in the vector $\mathbf{x}$

The implemented MLP model has to have:

- 1 input layer ($\mathbf{x}$) with 2 neurons
- two hidden layers ($\mathbf{h^{(0)}}$, $\mathbf{h^{(1)}}$) with 10 neurons each of them.
- 1 output layer of nerons ($\mathbf{o}$) with 1 or 2 neurons (please note that this depends on your choice on how to perform the classification)

The relation between layers are given by:
$\mathbf{h^{(0)}} = \phi^{(0)}(\mathbf{x})$
$\mathbf{h^{(1)}} = \phi^{(1)}(\mathbf{h^{(0)}})$
$o = \phi^{(2)}(\mathbf{h^{(1)}})$
, where $\phi^{(0)}$, $\phi^{(1)}$, and $\phi^{(2)}$ represent the activation functions of that specific layer.

The training has to be done using Backpropagation and Stochastic Gradient Descent.

Assignment

Please solve each of the following subproblems and give clear explanations for your solutions.

[1] no rar, no 7z, no tar, no tgz, no nonsense, just send a zip.
[2] Please note that we do not have and we do not use a testing dataset.

I. MLP (50 points)

1. (10 points) *Activation and Loss functions.* Please choose suitable activation functions $\phi^{(0)}$, $\phi^{(1)}$, $\phi^{(2)}$ and a suitable Loss function to perform the task. Report and justify your choices in the report.

2. (10 points) *Learning rate, batch size, initialization.* Please choose a suitable learning rate, batch size, initialization of the parameter values, and any other setting you may need. Discuss and justify your choices in the report.

3. (10 points) *Training.* Make plots with the loss function computed over the training set and over the validation set. Stop the training when the error is small enough. Justify your stopping criterium. Report the final accuracy obtained and the confusion matrix on the validation dataset.

4. (20 points) *Implementation.* We will run and check the uploaded Python file. To obtain the points for this subproblem, the Python file has to run (no errors) and the MLP model and the Backpropagation algorithm have to be implemented completely from scratch by you. You are not allowed to use any library which implements MLP models, but you are allowed to use auxiliary libraries, e.g. Numpy, Matplotlib, Pandas.

II. Peer Review paragraph (0 points)

Finally, each group member must write a single paragraph outlining their opinion on the work distribution within the group. Did every group member contribute equally? Did you split up tasks in a fair manner, or jointly worked through the exercises. Do you think that some members of your group deserve a different grade from others?

Success!