Bio-Intelligent Algorithms

Exercise 2

Overview

In this exercise you will implement from scratch, without using any external machine learning related library, a convolutional neural network for performing classification.

Note: You can use any programming language, but if you use C or C++ you will receive a bonus.

Data

Download the data.zip file together with the exercise. It contains three files:

a. train.csv, containing 8,000 samples

b. validate.csv, containing 1,000 samples

c. test.csv, containing 1,000 samples

Each row in each file represents a single sample, as follows:

The first value is the target class, between 1 to 10 (thus, the output layer of your neural network must have 10 neurons, one for each class), and then all the sample values appear one after the other, for a total of 3072 values, each a floating point number between 0 and 1 (thus, the input layer of your neural net must have 3072 neurons). These 3072 values represent an RGB image of size 32x32. That is, 1024 values for each channels. The 1024 values for each channel start from top-left corner of the image, and provide the values from left to right, row after row.

For the test.csv file, instead of the target class, a question mark (?) appears. This is because your model has to predict the values and you submit your predictions (see instructions below).

Training and Validation

Use the 8,000 samples (or any part of those) for training. Do not use the validation set for training! Use the validation set only for measuring the accuracy of your model.

You are allowed to use any neural network architecture, any preprocessing, regularization, etc. that you would like.

Test

After you finish training and validation, run your model on test.csv file, and for each sample store your prediction (a number between 1 to 10 denoting the predicted class) in a file named output.txt (make sure your output is adjusted to be between 1 to 10, not 0 to 9!).

What to Submit

You must submit the following:

- Your code and instructions for running it. If I run your code and the results are substantially different from what you have reported, the exercise score will automatically be zero.
- A report file (DOC or PDF) which contains the following, in the following precise order:
  - Your name and ID
  - Validation score: the accuracy you got on the validation set. Note that if the number you report here will be significantly different from the number I obtain measuring your test accuracy, then most probably you reported a false validation score, and the exercise score will be zero.
  - Detailed explanation of the various experiments you conducted with various hyperparamters. Include tables, graphs, etc.
- A text file names output.txt, containing exactly 1,000 rows, and in each row only a number between 1 to 10, denoting the prediction for the corresponding row in the test file.

Put everything in a ZIP file and use your ID for naming the file, e.g., 012345678_012345678.zip (ID of both submitters).

How to Submit
Upload your zip file by clicking on Upload link.

Exercise Scoring
Your final score will mostly depend on the test accuracy you obtain (I wll of course check your code, your report, number of experiments conducted, etc., but those would be secondary to test score).