Homework 3, Machine Learning

Multilayer Perceptrons

1 Data

In this task we use the MNIST dataset of handwritten digits from `http://yann.lecun.com/exdb/mnist/`. This dataset has 784 features (each sample is an image of 28x28 pixels) and 10 classes (each image represents a decimal digit). You can download the training dataset in CSV format from the course web page. You'll have to split it into training and validation set yourself.

2 Multilayer Perceptrons

The task is to implement the multilayer perceptron and the backpropagation training algorithm (described here: `http://courses.cs.ttu.ee/w/images/5/54/Bp_math.pdf`). You can either:

1. Implement the model and learning algorithm by your self using your favourite programming language or
2. Use some toolkit or library and implement only the code to execute experiments.

You can do both of course if you like to.

2.1 Toolkits

For python the recommended toolkit is `PyBrain`: `http://pybrain.org/`. Some relevant tutorials are: `http://pybrain.org/docs/tutorial/fnn.html` and `http://pybrain.org/docs/tutorial/datasets.html`.

You can also use any other toolkit that is available for your favorite programming language.

2.1.1 Experiments

If using a toolkit, experiment with different options provided by the toolkit. For example:

- different learning rates;
- different hidden layer sizes;
- with and without biases;
- different number of hidden layers;
- logistic and hyperbolic tangent hidden activations.

For a self-implemented network, the first two experiments (learning rate and hidden layer size) are sufficient. Record the validation errors of different values of hyperparameters or options.

2.2 Implementation

Your implementation must be able to save and load model parameters. Testing a previously saved model should be straightforward (e.g. on command line: `test <model filename> <test dataset filename>`). Testing should output the classification error in addition to cross entropy error.

2.3 Evaluation

Evaluate the results of different experiments using the validation data (you need to split it from the training data yourself).

3 Report

The homework submission must include the code of the programs, a short write-up (preferably in LaTeX) and your best model; The write-up should include a self-contained description of both tasks and the solution:

- What is the problem;
- Short description of the used method;
- Details of how to use your program and how to evaluate your best model on the test set;
- Descriptions of the experiments;
- Results of the experiments including the listing of the options/hyperparameters that resulted in your best model.