Assignment 1: Simple feedforward

Welcome to your first assignment. Upload the completed script task1.py and task2.py to the VLE before the deadline.

Standard criteria for valid assignments

In general, an assignment is considered valid if the following criteria are met:

- The delivered program uses Python and Tensorflow.
- The delivered program code is organised and easy to understand.
- Only one code file is used per task, which is a modified version of the provided scripts, with the same file names.
- The delivered program is deterministic (use fixed seeds for random numbers).
- When the delivered program is run, the model is trained from scratch using gradient descent (or an extension of gradient descent).
- The way the final test error is calculated is left as is.
- Only the provided dataset is used and without modification (although the data may be transformed for use by the model).
- The test set does not influence training in any way.
- The delivered program code is not excessively modified. I will be correcting your program by checking what was changed from the original provided scripts and needless modification makes corrections slower.
- Your aim should be to deliver the fastest, smallest, and best performing (according to test error) program possible within the given task constraints.

Task 1: Single-layer neural networks

Summary: Learn a single-layer neural network that correctly classifies MNIST handwritten digits.

You are provided with a file “dataset.txt” that contains a collection of flat bitmap images of hand written digits together with a half-finished Python script “task1.py”. Your task is to complete the script in order to learn a single-layer neural network using gradient descent with Tensorflow which accurately classifies the given digits in the dataset. You may split the dataset into separate validation sets if needed and you may use any technique discussed during the lectures.

The script also includes code to perform sensitivity analysis performed on a random test image for each digit. In the corner is the digit that the image is classified as by the neural network. The red blobs show parts of the image that when changed will be the most to change the classification.

Constraints

This task in particular is considered valid if the following criteria are met:

- The model is a single-layer simple feedforward neural network.
- The output of the model is a vector of 10 numbers such that the number at position 0 quantifies how likely the given digit is a 0, the number at position 1 quantifies how likely the given digit is a 1, etc.
- The final test error is less than 20% (extra marks if it is less than 13%).
- The total running time is less than 3 minutes.
- The total number of parameters is less than or equal to 7,850.

Task 2: Multi-layer neural networks

Summary: Learn a multi-layer neural network that correctly classifies MNIST handwritten digits.

As the previous task but with as many layers as you want. Script name is “task2.py”.

Constraints

This task in particular is considered valid if the following criteria are met:

- The model is a multi-layer (more than one) simple feedforward neural network.
- The output of the model is a vector of 10 numbers such that the number at position 0 quantifies how likely the given digit is a 0, the number at position 1 quantifies how likely the given digit is a 1, etc.
- The final test error is less than 15% (extra marks if it is less than 12%).
- The total running time is less than 5 minutes.
- The total number of parameters is less than or equal to 800,000.