Introduction to neural networks

Homework No. 1

Elad Tolochinsky Part

Part 1

We implemented the following variants:

* A simple network with no hidden layer. The network architecture was as follows:
  + An input layer comprised of 11 neurons.
  + An output layer of 10 neurons.
* A network with a single hidden layer.
  + An input layer comprised of 11 neurons.
  + A hidden layer of 20 neurons.
  + An output layer of 10 neurons.
* A network with two layers
  + An input layer comprised of 11 neurons.
  + A hidden layer with 20 neurons.
  + A hidden layer with 30 neurons.
  + An output layer of 10 neurons.

The number of neurons in each layer was determined after running repeated experiments with different number of neurons.

We implemented a grid search to determine the optimal learning rate and number EPOCHS.

The variant with no hidden layers exhibited the worst generalization capabilities.

The training time was the fastest for the simple network and increased as hidden layers were added.

Part 2

Expanding the dataset

To expand the given dataset, we randomly flipped 25 bits. We did so 4 times for each image. Furthermore, we cyclically translated the images by one row in all 4 directions. Thus we managed to generate additional 8 images from every given image.

We know from the previous part and from theory that a neural network with no hidden layers generalize poorly, thus we implemented only the following variant:

* Two hidden layers
  + An input layer comprised of 256 neurons.
  + A hidden layer with 150 neurons.
  + A hidden layer with 70 neurons.
  + An output layer of 10 neurons.

As before we ran a grid search to determine the best learning rate.