

Project 4, Handwritten Digit Recognition using Neural Networks (Tensorflow)

CSC540 Introduction to Artificial Intelligence
Spring 2021, Max Score: 100

Objectives

For this assignment you will implement a feedforward neural network with one hidden layer that learns how to recognize handwritten digits. You are encouraged to use Tensorflow library. Tensorflow Installation: <https://www.tensorflow.org/install>. This problem is important in applications such as automatic zip code reading on letters, which is in current use by the U.S. Postal Service.

The Data

The data file, called optdigits-3.tra, is an ASCII file containing 1535 examples, one per line. Each example is a comma-separated (without any white space) list of 65 integer values, the first 64 specifying the input and the last value specifying the digit which is the desired output. The input values are integers in the range [0..16]. You should first normalize the input values by converting them to reals in the range [0..1] by dividing every value by 16.0. This is useful because the derivative of the sigmoid function is often very close to 0, which can cause the network to converge very slowly to a good set of weights.

You may convert the desired output digit to a target output vector for the four output units (0, 1, 2, 3). For example, if the digit is a “3” then create the target vector [0.1 0.1 0.1 0.9]. Using this set of teacher output values is preferred because the sigmoid function cannot produce the exact output values of 0 and 1 using finite weights, and so the weight values may get very, very large causing overflow problems.

Training

Please follow the tutorial that I demonstrated in class for the training process. <https://www.tensorflow.org/tutorials/keras/classification>

Testing

Test your network using the examples in the test set. Report the percentage correct classification on the test set. Define the output digit computed by the network as the corresponding output unit with maximum activation (i.e., output) value. In case you are interested, the complete dataset that includes examples of all ten digits and includes over 5,000 examples is available at <ftp://ftp.ics.uci.edu/pub/machine-learning-databases/optdigits/>.

Experiments with Varying Numbers of Hidden Units

Repeat the training and testing steps given above after varying the number of hidden units in your network. Use values of 5 and 50. Show the training plot using the same training and testing sets. Comment on how performance changes with the number of hidden units, both in terms of the percentage correctly classified on the training set and the test set, and also in terms of the number of iterations that seem necessary for the network to “converge.” Specify what you consider to be the “best” set of parameters (alpha, number of hidden units, number of iterations) for this problem based on your experiments.