Gradient Descent

Towards Neural Networks

Justin Stevens

Outline

Example Images

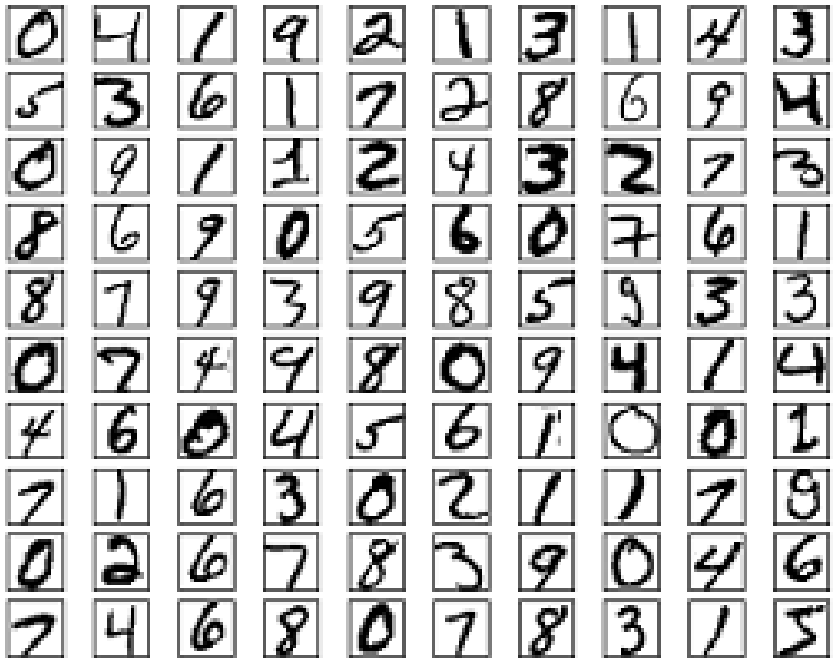

Figure 1: How would you devise a system for a **computer** to classify the digits? What assumptions do we have to make about the data set, known as MNIST?

Assumptions

- The MNIST database contains thousands of handwritten digits.
- Each data-point contains both an image, and the desired digit.
- The images are 28×28 pixel arrays, with each pixel ranging from 0 to 255.
- $60,000$ of these are designated for training purposes, and $10,000$ for test purposes.
- We'll build a model from these, that will learn to classify digits!

What we're building towards

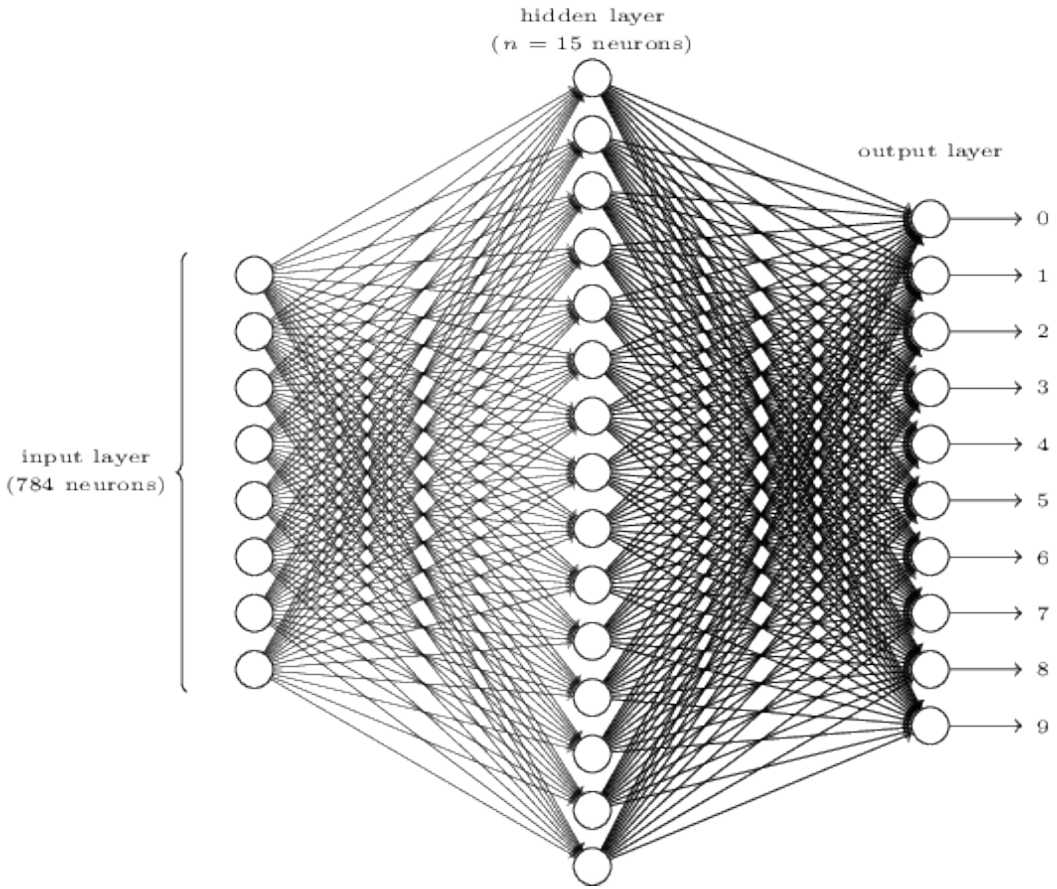

Figure 2: A simple neural network structure. The input vectors on the left hand side have $28 \times 28 = 784$ inputs for each pixel, and the output layer has 10 digits, one for each number from 0 to 9.

References

- Stewart Calculus: Early Transcedentals, 6th Edition
- Professor Leonard Calculus 3 (Full Length Videos)
- Paul's Online Math Notes, Calculus III