Introduction

• Neural Networks
• Recognize handwritten digits
Data Set

• MNIST data set: http://yann.lecun.com/exdb/mnist/
• 60,000 training images and 10,000 testing images
• Benchmark data set for machine learning
Write a program that can do this?

- Normalized to fit into 28 x 28 pixel bounding box
- Grey scale
Neural Network

• Create Neural Network (NN) that can learn and recognize hand-written digits
• As name says, inspired by brain
• Many NN variants exist
• Understanding basic NN is requirement to understand more complex ones
• Classic example => “Hello World” of Machine Learning
Reducing Loss

- Gradient descent
  - the derivative of the loss function with respect to the model parameters.

- Learning rate
  - Multiply the gradient by a scalar

![Diagram showing the effect of small and large learning rates on reducing loss.](Image)
Optimizing computation

• In large data sets, calculating the gradient descent on billions of data points can be very computationally intensive.

• Computing the gradient descent in a small fraction of the data set produces similar results. This is called **Stochastic Gradient Descent**.

• An intermediate solution, is computing the gradient descent in a small batch of data. This approach is called **Mini-Batch Gradient Descent**.