

Understanding How Neural Networks See (and Read)

Michael Guerzhoy

University of Toronto
`guerzhoy@cs.toronto.edu`

I introduce a slide deck I use in high-school hackathons and (with minor modifications) in first-year data science (non-CS audiences) and computer science classes to introduce neural networks. The slide deck assumes strong mathematical background from high school, but not more than that. The slide deck is usually particularly well-received in high-school hackathons where enthusiastic high school students with good math/programming background at the high school level are first exposed to neural networks. The slide deck is readable in about 10 minutes. Different versions of the talk can be 20 to 50 minutes.

Slide Deck Structure

The structure is as follows

- The idea of supervised learning
- One-nearest neighbor as the simplest and most intuitive way to do supervised learning
- The notion of distance between images (as necessary for one-nearest neighbor), walking up to the notion that a dot product is a reasonable way of measuring image similarity
- A zero-hidden-layers feedforward neural network (i.e., basically multinomial logistic regression), interpreted as k dot products of the input with weight sets, where k is the number of outputs
- Interpreting sets of weights used to compute the outputs of a zero-hidden-layer feedforward neural network as “templates” to which the input is compared
- Connection between matching to templates in a neural network and nearest-neighbor
- One-hidden-layer neural network as computing “templates over templates”, visualization of templates corresponding to hidden units
- Deep neural networks intuition
- Deep neural networks as a model of computation
- From images to predicting the network word

Key Features

- No calculus required – but linear algebra is
- A coherent story building up from the intuitive to one intuition for deep neural networks
- Bonus content for fans of computation: neural networks are programs