

How does a neural net, work?

Philipp Zettl

24. Dezember 2020

At some point one can write a preamble to explain what a NN is, what to use it for and why it's actually better for us to use them then old numerical methods like the euler method to calculate the root(s) of a function

Kapitel 1

FFNN

Feed forward neural networks are a great entry point to start discovering neural networks. The following chapter will introduce the idea of general neural networks and use this motivation to discover the first type of neural networks.

1.1 Using NNs to recognize handwritten digits

Unlike us humans computers have rather a hard time recognizing handwritten characters, digits or even full words. Humans taught themselves over millions of years of evolution how to recognize things.

Since handwriting is mostly unique to a person one will have a hard time implementing an algorithm to detect digits from handwritten text, using an enormous amount of if/else conditions will be the least problem, but figuring out how to provide a stable algorithmically solution to this problem seems impossible. Nowadays we can use neural networks to mock this behavior and create models to do that task for us. In the following chapter I will describe how to create such model as well as how to train it using real live examples.

To train a NN one uses a so called "Training-Set" of data to "teach" the network what prediction to give for a given input. In some way we can assume achieving similar results using equal training data sets.

But wait, how does a neural net, work?

Now let us construct a basic neural network.

1.2 General NN

To define the structure of Neural Networks we need to define a bunch of technical terms. In biology we differentiate between perceptrons and neurons. Perceptrons are basically binary decision makers having n input values and a single output value. We can illustrate them using the following notation.



When we talk about NNs we differentiate between perceptrons and neurons.

1.3 Backpropagation

1.3.1 Notation

Das Gewicht w , welches die Verbindung zwischen dem k -ten Neuron des $l-1$ -ten Layers und dem j -ten Neuron des l -ten Layers wird geschrieben als

$$w_{jk}^l \tag{1.1}$$