(/)

Swizec Teller
I fix the messy code startups produce.

So far I have written two books and attracted over a million readers to this blog.

About me (/)
(mailto:swizec@swizec.com) (https://github.com/Swizec) (https://twitter.com/Swizec) (http://swizec.com/blog/feed)

Published
November 9, 2011
in A tech a day (http://swizec.com/blog/category/coding/a-tech-a-day-coding).

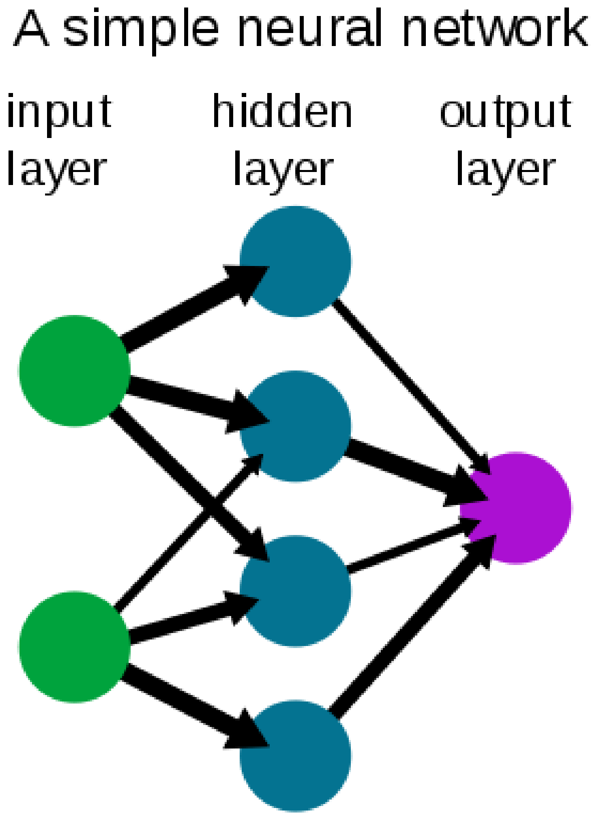
I think I finally understand what a neural network is

Last night I was consumed with watching lessons from the online machine learning class at Stanford (http://ml-class.org). The topic was neural networks (http://en.wikipedia.org/wiki/Neural_network), or rather the finer points of forward and back propagation (http://en.wikipedia.org/wiki/Backpropagation).

Every time I have wanted to learn how neural networks work it just didn't click. Anyone I asked mostly just waved their hands and said that something magical happens and … but let's start at the beginning.

The basics of neural networks is that you have neurons that are connected to other neurons. At one end you enter your input data and on the other end the neural network produces some numbers according to what it has learned. Everyone can imagine this much and it's not really difficult to visualise.

It looks like a bunch of circles with arrows:



(http://commons.wikipedia.org/wiki/File:Neural_network_example.svg)
Image via Wikipedia

Anyone who's studied this a tad further can tell you the connections between neurons are very important and the weights associated with them are somehow used in calculating stuff. Not a hard concept to grasp – every neuron outputs a number that is multiplied with the weight on each connection before being fed as an input into the next neuron.

Where it always got a bit hairy for me was trying to understand *what do the neurons do?* The most I could get out of anyone supposedly knowing this stuff is that "it calcualtes stuff". Yes but *how? What* does it do? What exactly?

Nobody knew.

Last night I finally figured it out! Neurons don't do anything. They don't even exist per se. In fact a neural network looks pretty damn odd inside a computer, it's really just a matrix of weights.

What happens when you're doing forward propagation (using a learned network) is simply this:

1. Take the outputs from the previous layer (a vector of numbers)
2. Multiply with a vector of weights (the arrows)
3. Apply the cost function (this becomes the new layer)

Then you just repeat this for all the layers and that's that. That is literally all that happens.

In the end you are left with a vector of numbers representing the output layer, which you then just have to correctly interpret.

The part I don't have completely figured out yet is the backpropagation. This is the bit where neural networks learn how to do their magic. Basically backpropagations sets those weights from step 2 via a simple hill climbing algorithm (http://en.wikipedia.org/wiki/Hill_climbing) … it is essentially a way to calculate the gradient of the cost function so that you can correctly change the weights to achieve ever lower differences between what you're supposed to know and what you actually know. Eventually you hope to achieve a global minimum (http://en.wikipedia.org/wiki/Maxima_and_minima), but you are guaranteed to at least achieve a local minimum and not being able to tell whether it's global.

That's it. That is really all there is to it. Neural networks are just a nice way to visualise a sequence of matrix multiplications. And I guess it's easier to get grants for "neural networks" than "sequence of matrix multiplications" …

Related articles

- Tradeoffs and constraints on neural representation in networks of cortical neurons. (http://www.citeulike.org/user/fbaroni/article/7490581) (citeulike.org)
- Training of Neural Network (http://vql89.wordpress.com/2011/10/26/training-of-neural-network/) (vql89.wordpress.com)
- Unique Neural Networks Make Leo Trader Pro A Unique And Profitable Forex Trading System (http://joeyjardins.wordpress.com/2011/10/22/unique-neural-networks-make-leo-trader-pro-a-unique-and-profitable-forex-trading-system/) (joeyjardins.wordpress.com)
- Artificial Intelligence Finds Fossil Sites (http://www.scientificamerican.com/article.cfm?id=artificial-intelligence-finds-fossil-sites) (scientificamerican.com)

(http://www.zemanta.com/)
You should follow me on twitter, here (https://twitter.com/swizec).

(https://twitter.com/share?text=I think I finally understand what a neural network is by @Swizec)

5 Comments A geek with a hat Sagar Giri

Sort by Best Recommend Share

Join the discussion…

Rafael Pinto · 3 years ago

Just a small correction: where you wrote "Apply the cost function", it should be "Apply the activation/transfer function". The cost function is the error function (J) used by backprop at the end of the network.

About backprop, what it is basically doing is "credit assignment", using the cost function derivatives it can tell you how much to blame each weight for the output errors and correct them accordingly. Also it propagates the errors to the hidden neurons (just a weighted average) and then you can repeat the corrections where otherwise you wouldn't have any error signals (you don't have target values for the hidden neurons).

And finally, about what the neurons do, it's more of a point-of-view... I, personally, wouldn't say they don't do anything. I prefer to see them as applying a function to a weighted average / scalar product. Actually, there is recent news in neuroscience about researchers who discovered that biological neurons compute weighted averages!
As with all artificial intelligence techniques, it doesn't seem AI after you understand it. "this is just a brute force search!" "these are just probability calculations!" "these are just kinematic equations!" "these are just matrix multiplications!"BTW, forgot to say in my other comment in another post: just found your blog reading about that tiny turing machine in JS, subscribing :)

1 ⌃ | ⌄ • Reply • Share ›

Swizec Mod ➦ Rafael Pinto • 3 years ago

Hey thanks! That's even clearer :)

Happy to hear you're subscribing, hope I don't disappoint.

⌃ | ⌄ • Reply • Share ›

Shashi Sathyanarayana • 5 months ago

You wrote "The part I don’t have completely figured out yet is the backpropagation."

I hope my tutorial will change that for you. Please see http://numericinsight.com/uplo...

⌃ | ⌄ • Reply • Share ›

sep332 • 3 years ago

Backpropagation is for training, right? So all the ouptuts that were "correct" increase the weight of the matching connections from the previous layer and decrease the others, while the outputs that were "wrong" decrease the weight of matching connections and increase the others.

⌃ | ⌄ • Reply • Share ›

Swizec Mod ➦ sep332 • 3 years ago

Yes I think that's basically the result of doing backprop matrix multiplication stuff ... I don't specifically understand that part because I haven't implemented it yet in octave and I'd rather not claim I understand something until I have written code that does it :)

⌃ | ⌄ • Reply • Share ›

ALSO ON **A GEEK WITH A HAT** WHAT'S THIS?

When the universe tries to communicate, or how I almost didn’t get home 2 comments

Code completion is bad mmkay 7 comments

My new favourite Javascript trick 2 comments

Javascript debugging [slightly] beyond console.log 7 comments