Bryson Meiling Assn 3 - Paper Summary

1) a brief statement of the problem addressed in the paper;

Neural networks are linkages of nodes that activate when their internal biased in reached and send output to other nodes. Some nodes affect others more than others. These linkages are put together with input layers that take information from a system, disseminate the information through to the last layer of nodes, called the output layer.

Because of the massive amount of biases and weights that can be adjusted in neural networks, it appears as though they are useless sense there is no rhyme or reason to what bias/ weight will produce a desired output. Adjusting these values by hand is inefficient and no possible due to the lack of effect it will have on the output.

This paper addresses the means by which Rumelhart et al will systematically adjusts these network values, so the system produces a desired output from any given input. By comparing the squares of the differences between the desired output and actual output, one can see what node and weight made a most impact. Then moving to the forward layers, the system will see what other nodes were most responsible for making the output node activate. This process is repeated until all the nodes and their values are adjusted. By repeating this entire process for many inputs and outputs, the system is ‘trained’ to produce a desired output.

2) what you liked in the paper and why;

I liked that for the time, this was a difficult task that had to be solved. It turned a useless mathematical model into something that could be trained and adjusted to mimic a neurological system. From this, the whole network turns from a interesting theory into a practical model that could be trained to do nearly anything with enough training data.

I also like that the math expressions used to represent the system were much more manageable and practical than the McCulloh and Pitts paper. This was terribly difficult to understand since it used PH. D level discrete math and was not easily read, less understood.

3) what you did not like in the paper and why;

I didn’t like that they used unconventional XOR network to convey as an example throughout the paper. This system would be more difficult to model in just a linear algebra model because the input layers connect to a hidden and output layers. So, computation must be done to complete the next step, making the process asynchronous. However, they make up for this by showing a more typical computer model that one would we now with distinct layers between the input and outputs.

4) any inspirations you found in the paper
They used the system to characterize if the input was a T or a C. This is the first step to handwritten letter recognition and also image recognition. It’s inspiring that the base idea of

even complicated NNs today are still based on simple math and linear algebra that was discussed year ahead. Many times, genius is not recognized in their times.

References

https://www.academia.edu/2520405/Learning_representations_by_back-propagating_errors