

Q6/7: Paper Review

0-Main Contribution

The main contribution of the paper was to challenge the notion that learning weights for neural networks is the only effective way of producing high performing networks. If this notion is removed, then the process for solving neural networks would be reduced to only designing the structure of the network; the weights can be randomly initialized and the network is ready for use. The authors discuss their method for learning the network structure automatically.

1-Strengths

Weight Agnostic Neural Nets (WANN) are immediately useful, and do not require tuning of weights, whereas traditional parameterized models require extensive training to tune good/useful weights.

2-Drawbacks

Looking at results table, we see that there is more variance in performance metrics for WANN than fixed topology models. This might be just a natural by-product of random/shared weights. This variance goes away, however, with weight tuning. Even with tuning, WANN lose to traditional models on benchmark problems like Biped, CarRacing, MNIST(2 layer CNN). However, the authors do note that this is somewhat of an unfair comparison. Many fixed topology networks(they focus on CNNs) are products of years of research and experimentation.

Another important weakness of the paper is that it fails to discuss the time tradeoff between searching for the WANN and tuning optimal parameters. The authors did not show that using WANN produced significant speedup compared to training traditional fixed topologies to find optimal weights. This is not exactly a red flag, but it would have provided a stronger argument for researching WANN.

3-Personal Takeaways

This was very interesting to me, as I had read about NEAT as part of my VIP work. The VIP team works on a genetic programming autoML framework, and some members were interested in incorporating NEAT into the framework. The evolutionary style of searching for WANN makes me think that this is also a good feature to implement.

Conceptually, I find that this provides a nice balance in thinking about neural network design and optimization. On one end of the spectrum is the traditional optimization of parameters and the other end is WANN.