



Efficient Training of Deep Learning Models

Understanding the Lottery Ticket Hypothesis (Frankle & Carbin, 2018)

Examination Exercise for class dlmdsdl01 Deep Learning

Koen Bothmer | Student Matriculation number: 92014567



Outline

- Pruning in Decision Trees and Neural Networks
- The Lottery Ticket Hypothesis Explained
- Finding Winning Tickets
- Conclusions

What is Pruning?

Pruning in decision trees

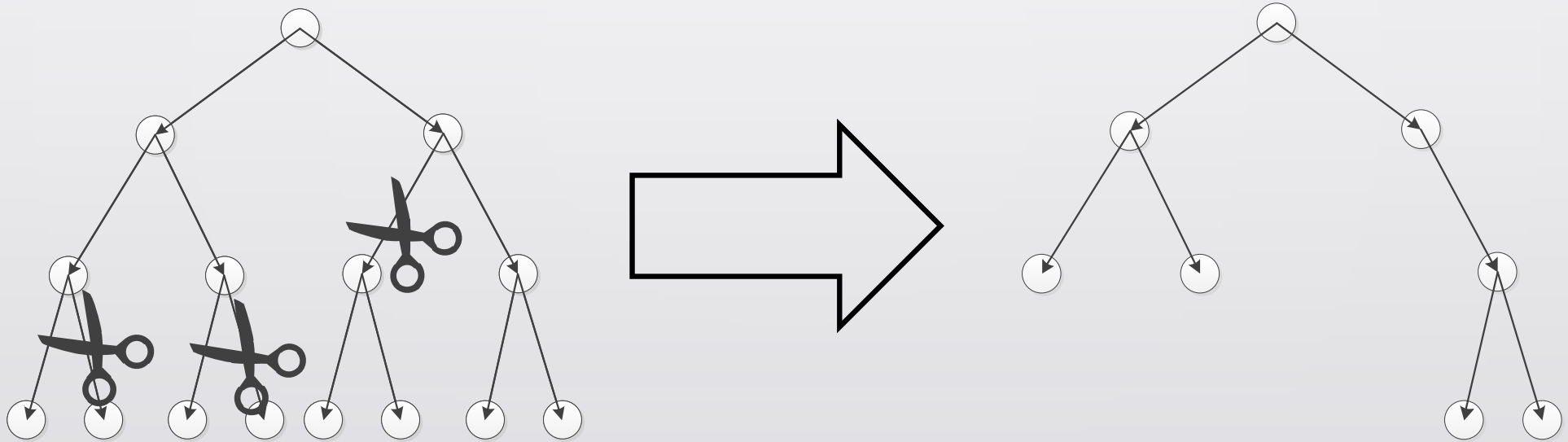


Figure 1. Pruning Decision Trees.

Note: Loosely after the explanation of pruning (Mingers, 1989)



Pruning Neural Nets

- State of the art Neural Networks are big
- In general: Networks are pruned to run more effectively in inference stage

Pruning Neural Nets

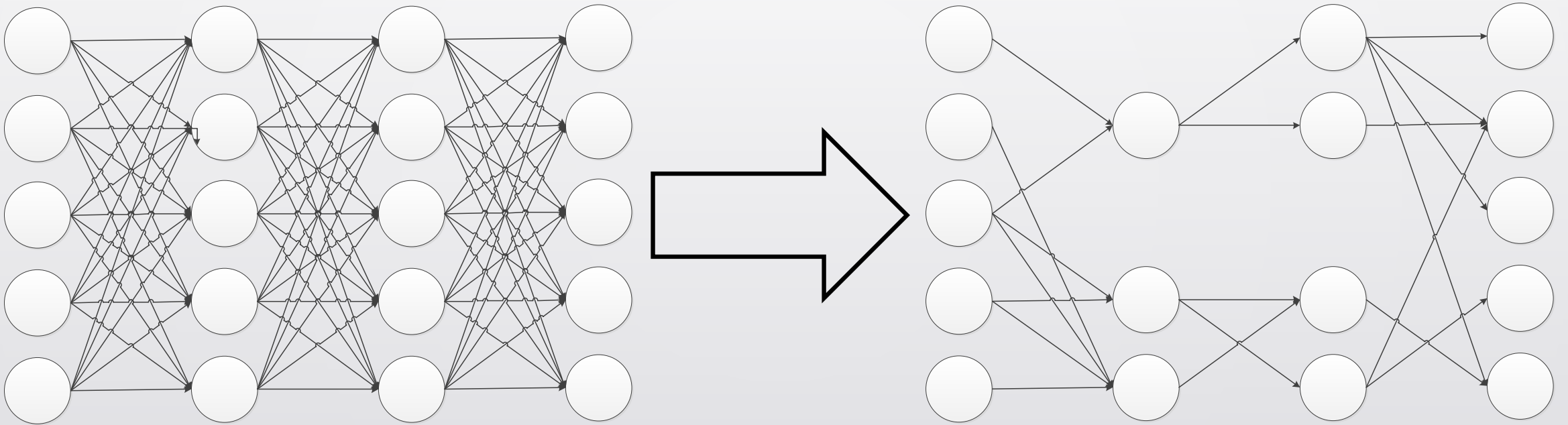


Figure 2. Pruning Neural Networks.

Note: Visualization based on techniques described by Karnin (Karnin, 1990)

Pruning Neural Networks

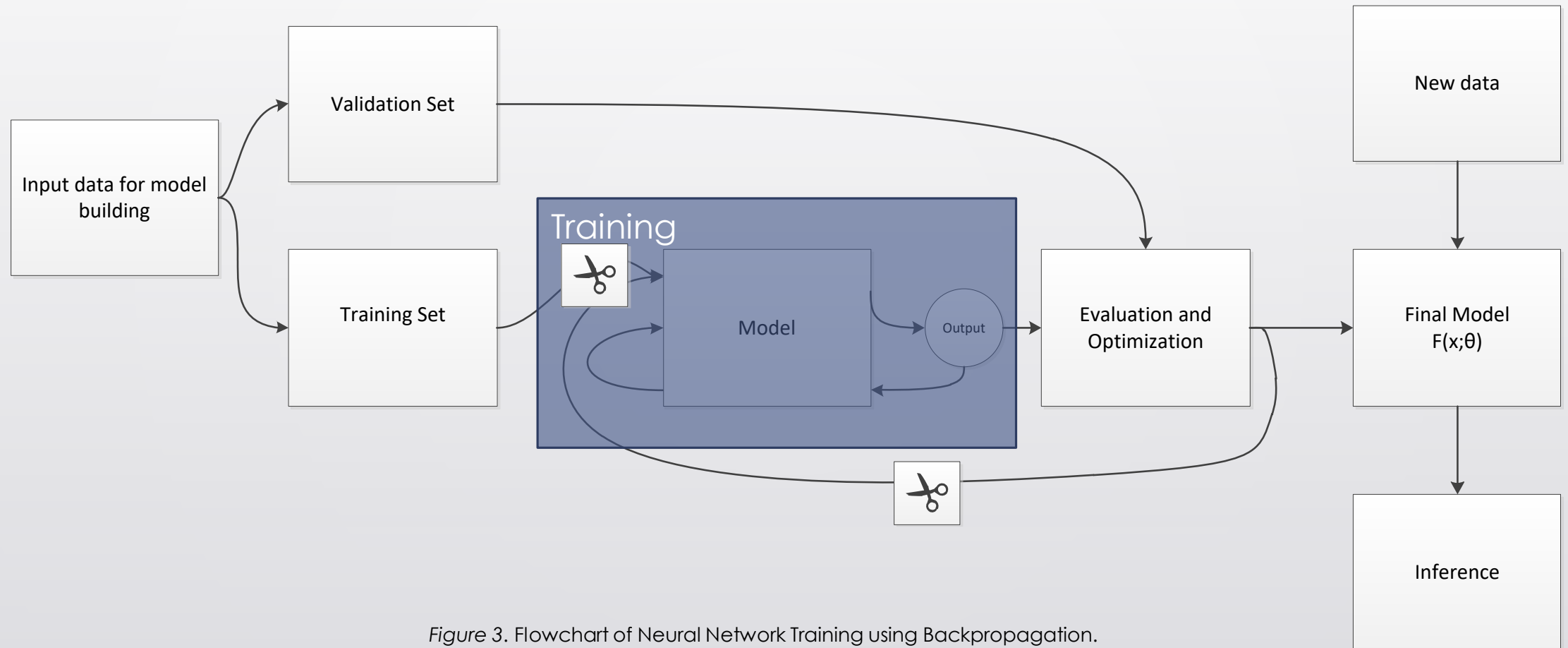


Figure 3. Flowchart of Neural Network Training using Backpropagation.

The Lottery Ticket Hypothesis

- “A randomly-initialized, dense neural network contains a subnetwork that is initialized such that—when trained in isolation—it can match the test accuracy of the original network after training for at most the same number of iterations” (Frankle & Carbin, 2018, p. 2).

- In a dense network
 - Which has not seen data yet
 - A sparse subnetwork exists
 - Which outperforms the original

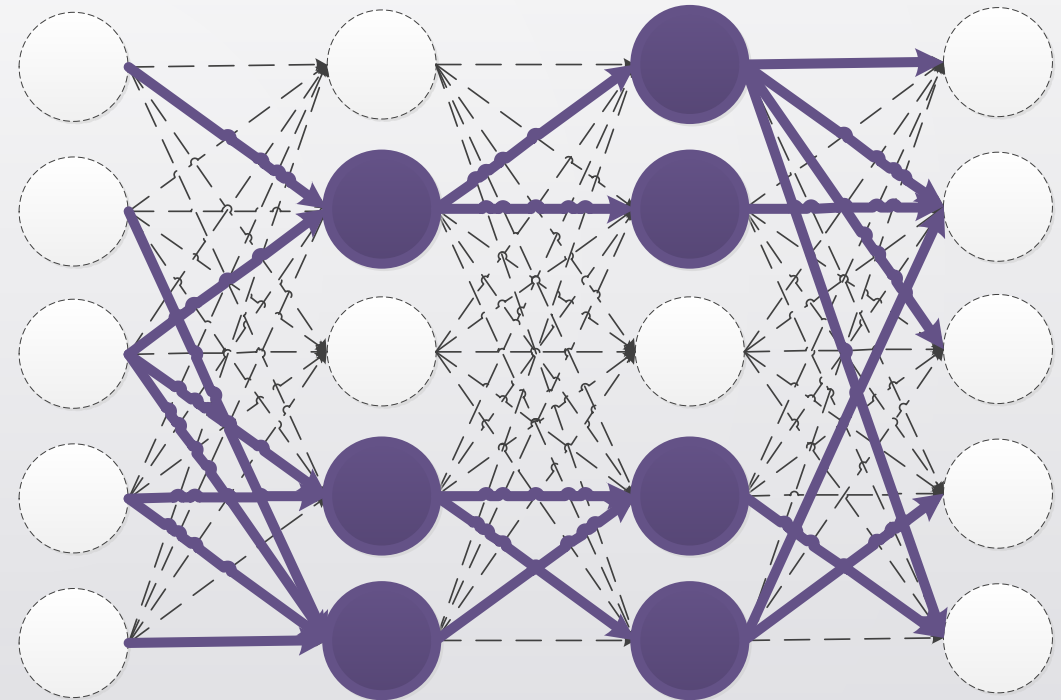


Figure 4. Sparse Network within Network.

Note: Visualizing Lottery Ticket ideas (Frankle & Carbin, 2018),

The Lottery Ticket Conjecture

- Untested explanation (Frankle & Carbin, 2018)
- Overparameterization leads to many subnetworks (Frankle & Carbin, 2018)
- Landing in a 'valley' of the loss landscape (Machine Learning Street Talks, 2020)

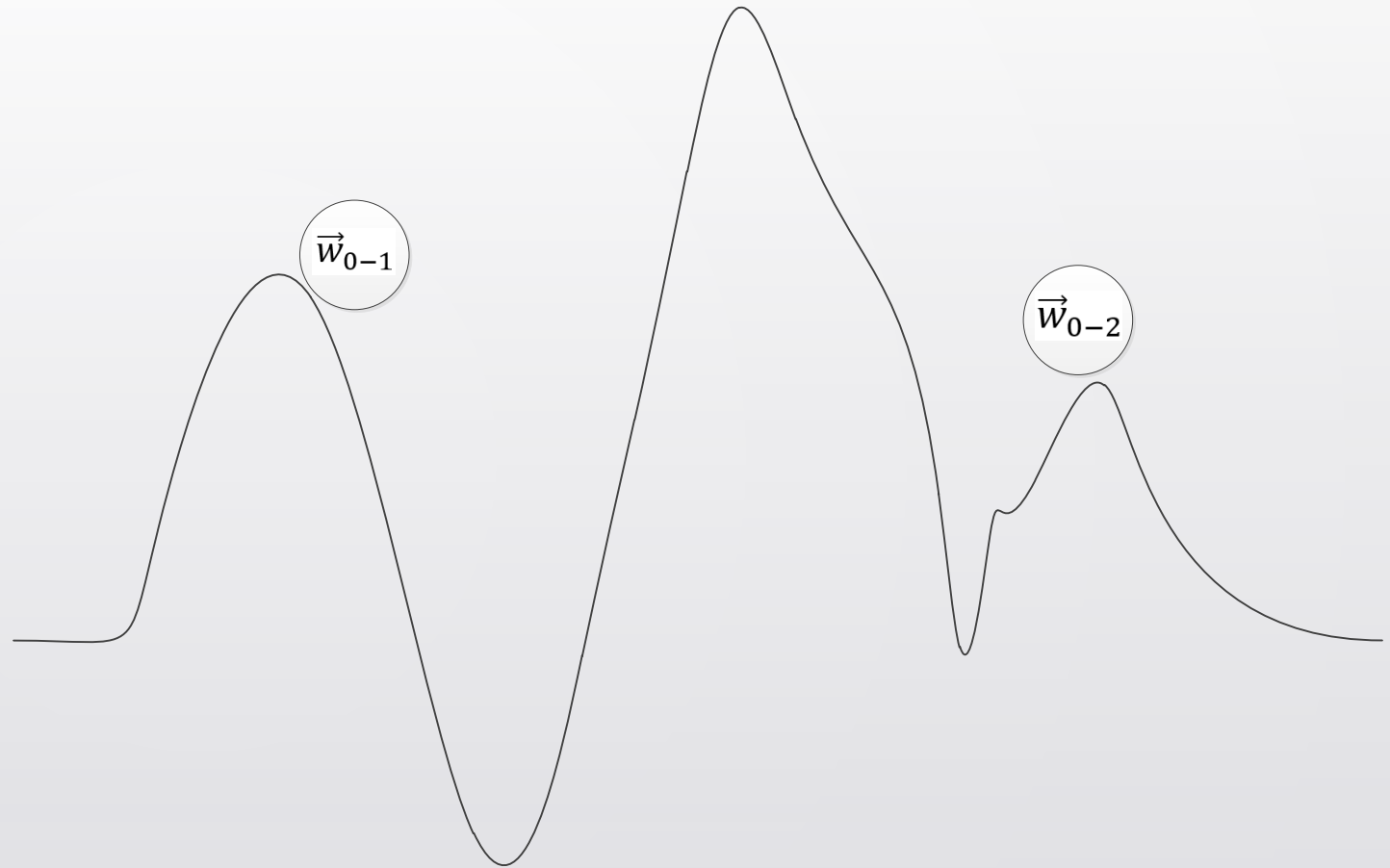


Figure 5. Simplified Representation of the Loss Landscape.

Note: Visualizing ideas Frankle speaks about (Machine Learning Street Talks, 2020)

Finding a winning Lottery Ticket

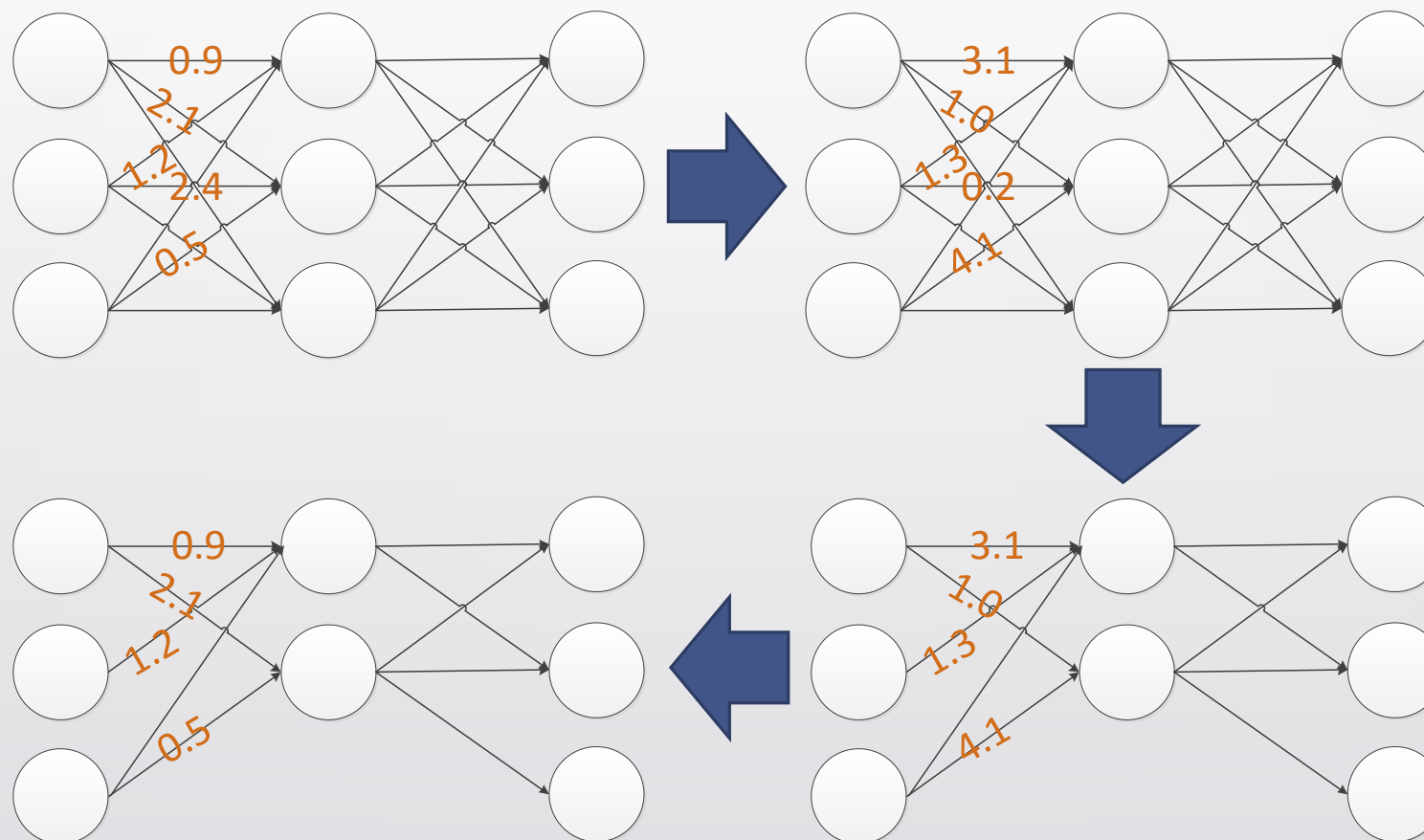


Figure 6. Finding a Winning Lottery Ticket in Dense Networks.

Note: Visualizing Lottery Ticket ideas (Frankle and Carbin, 2018), visualization method inspired by Statquest (Starmer, 2020)

Iterative Pruning

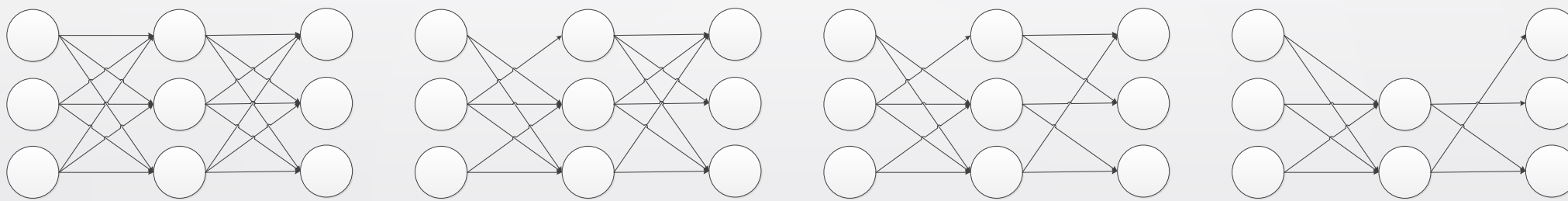


Figure 7. Iterative Pruning to find Winning Tickets.

Note: Visualizing ideas Lottery Ticket ideas (Frankle & Carbin, 2018)

Conclusion

- Main source (Frankle & Carbin, 2018) shows a lot of evidence for it's Hypothesis
- Resetting pruned networks to initial parameters is the ground breaking part
- Step towards more effective training of Neural Networks
- Questions? Please reach out!
Koen.Bothmer@iubh.de
+31 646 59 46 49

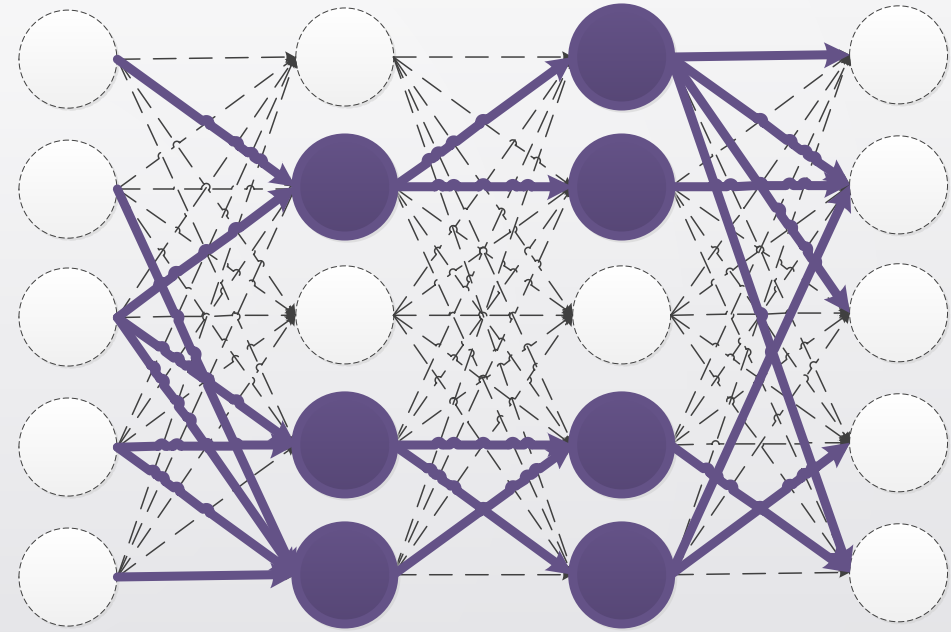


Figure 4. Sparse Network within Network.

Note: Visualizing Lottery Ticket ideas (Frankle & Carbin, 2018),



List of Figures

	Slide
<i>Figure 1.</i> Pruning Decision Trees.	3
<i>Figure 2.</i> Pruning Neural Networks.	5
<i>Figure 3.</i> Flowchart of Neural Network Training using Backpropagation.	6
<i>Figure 4.</i> Sparse Network within Network.	7
<i>Figure 5.</i> Simplified Representation of the Loss Landscape.	8
<i>Figure 6.</i> Finding a Winning Lottery Ticket in Dense Networks.	9
<i>Figure 7.</i> Iterative Pruning to find Winning Tickets.	10



Bibliography

- Frankle, J., & Carbin, M. (2018). The lottery ticket hypothesis: Finding sparse, trainable neural networks. *arXiv preprint arXiv:1803.03635*.
- Karnin, E. D. (1990). A simple procedure for pruning back-propagation trained neural networks. *IEEE transactions on neural networks*, 1 (2), 239-242.
- Mingers, J. (1989). An empirical comparison of pruning methods for decision tree induction. *Machine learning*, 4(2), 227-243.
- Machine Learning Street Talk. (2020, may 2). *The Lottery Ticket Hypothesis with Jonathan Frankle* [Video]. YouTube. URL <https://www.youtube.com/watch?v=SfjJoevBbjU>
- Starmer, J. (2020, October 19.). *Neural Networks Pt. 2: Backpropagation Main Ideas* [Video]. YouTube. URL <https://www.youtube.com/watch?v=IN2XmBhILt4&t=639s>