Why Large Language Models can understand the world

compiled by D. Gueorguiev, foreword by Paul Thompson, 9/23/2025

# Introductory Notes

Why do Deep Neural Nets with billions of parameters generalize well, instead of overfitting?

Statisticians often say you need 10–30 training samples per predictor, so how do these massive models can still perform well on unseen data?

[1] presents a theory with explanation:

In well trained neural networks, only a small subset of the nodes really matters. The optimal subnetwork is often one whose weights happened to be “lucky” - initially random but set near a better local optimum in the objective function and also converge faster. Large networks contain "many chances to be lucky". As network size grows, the number of possible small subnetworks increases super-exponentially, making it more likely that at least one is well-initialized and generalizes well. This "lottery ticket hypothesis", has parallels in human collaborations and game theory (think of the maximally competitive sub-alliance in a collaborative network).

In AI-human collaborations, is deep learning “choosing” teams (from the "8 billion-Choose-N" = 8000000000!/(N! (8000000000 - N)!) combinations of humans) who happen to try out the best-performing methods and get their papers into NeurIPS?

A collage of multiple images of a network

AI-generated content may be incorrect.

# References

[1] [This is why large language models can understand the world, Algorithmic Simplicity, March 29, 2025](https://youtu.be/UKcWu1l_UNw?si=5OrnLXCewvdrOeXL)

[2] [Reconciling modern machine learning practice and the bias-variance trade-off, Mikhail Belkin et al, 2018](https://github.com/dimitarpg13/deep_learning_and_neural_networks/blob/main/literature/articles/Reconciling_modern_machine_learning_practice_and_the_bias-variance_trade-off_Belkin_2018.pdf)

[3] [The Lottery Ticket Hypothesis: Finding Sparse, Trainable Neural Networks, Jonathan Frankle, Michael Carbin, 2019](https://github.com/dimitarpg13/deep_learning_and_neural_networks/blob/main/literature/articles/The_Lottery_Ticket_Hypothesis-Finding_Sparse_Trainable_Neural_Networks_Frankle_2019.pdf)