grokking the system design interview

Grokking (machine learning) - Wikipedia

In ML research, "grokking" is not used as a synonym for "generalization"; rather, it names a sometimes-observed delayed-generalization training phenomenon in which training and held-out performance do not improve in tandem, and in which held-out performance rises abruptly later.

[2201.02177] Grokking: Generalization Beyond Overfitting on ...

In this paper we propose to study generalization of neural networks on small algorithmically generated datasets. In this setting, questions about data efficiency, memorization, generalization, and speed of learning can be studied in great detail.

Grokking: A Deep Dive into Delayed Generalization in Neural ...

One of the most intriguing is the phenomenon of grokking, where neural networks exhibit surprisingly delayed generalization, achieving high performance on unseen data long after they have...

GROKKING Definition & Meaning - Merriam-Webster

Grok may be the only English word that derives from Martian. Yes, we do mean the language of the planet Mars. No, we're not getting spacey; we've just ventured into the realm of science fiction. Grok was introduced in Robert A. Heinlein 's 1961 science fiction novel Stranger in a Strange Land.

What is Grokking? From Rote to Revelation, overfitting ...

Grokking forces us to reconsider established practices in training neural networks. It challenges the validity of early stopping criteria and suggests that a model appearing to overfit might actually be on the verge of discovering deeper patterns.

Grokking - GitHub Pages

Grokking, or delayed generalization, is a phenomenon where generalization in a deep neural network (DNN) occurs long after achieving near zero training error. Previous studies have reported the occurrence of grokking in specific controlled settings, such as DNNs initialized with large-norm parameters or transformers trained on algorithmic datasets.

Grokking in Large Language Models: Concepts, Models, and ...

Definition of Grokking: Grokking refers to a surprising phenomenon of delayed generalization in neural network training. A model will perfectly fit the training data (near-100% training accuracy) yet remain at chance-level on the test set for an extended period.

Grokking in Neural Networks: A Review - SN Computer Science

One such phenomenon is grokking. According to the Oxford English Dictionary, "to grok something" means "to understand something completely using your feelings rather than considering the facts."

Carlisia Campos - Grokking

Grokking implies experiential, embodied learning, something beyond surface-level exposure. It hints of an orientation towards fluid intuition, rather than rigid knowing or memorization.

Do Machine Learning Models Memorize or Generalize?

When Does Grokking Happen? It's important to note that grokking is a contingent phenomenon — it goes away if model size, weight decay, data size and other hyper parameters aren't just right.