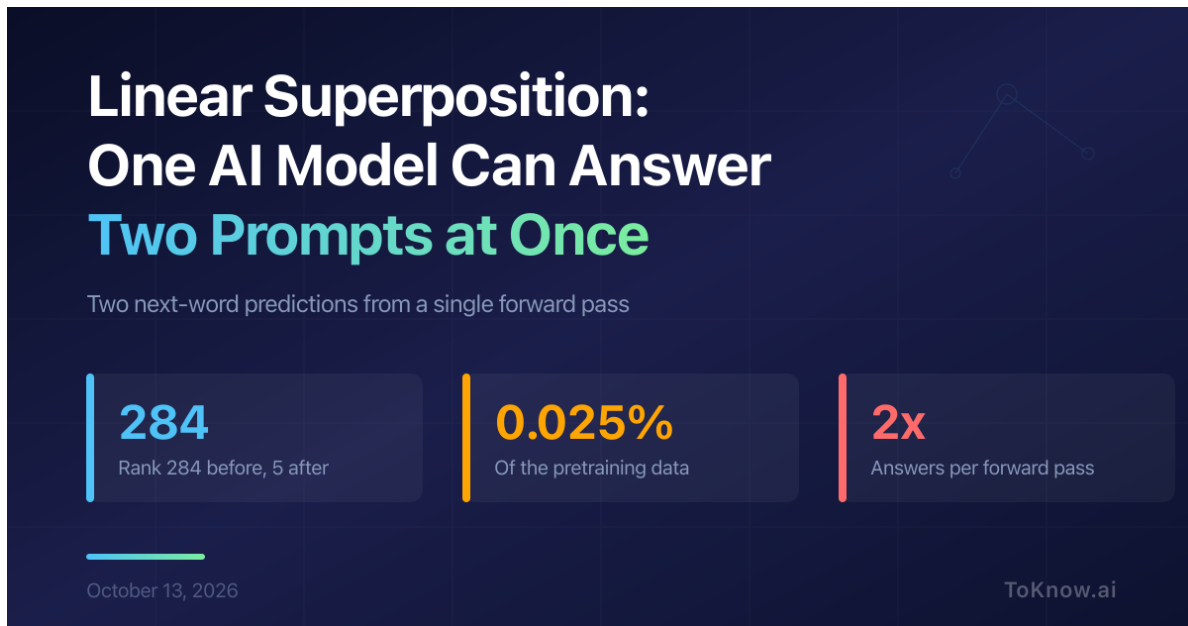


Linear Superposition: One AI Model Can Answer Two Prompts at Once

Kabui, Charles

2026-10-13

[Read at ToKnow.ai](#)



A team of researchers has found that a normal language model can hold two different texts at once. Average the [word embeddings](#), the number lists a model uses for words, from two documents, feed the blend through an unchanged model, and the output keeps the right next word alive for both texts. They call it the [Superposition Linearity Hypothesis](#). Across the

Pythia, Qwen, Llama, OLMo and Gemma models, each stream's correct next word often lands in the top 10.

The effect is strongest before training and fades as a model learns, pointing to the architecture, not to learning. A short fine-tuning run brings it back using under 0.025% of the original training data. On Pythia-2.8B that cut the gap between the blend and the true mix, measured as KL divergence, a score for how far apart two distributions sit, from 1.86 to 0.27. On Qwen2.5-3B, harder words moved from a median rank of 284 to 5. Two answers from one pass could double throughput and roughly halve memory per conversation.

Sampling straight from the blend hops between the two texts, because averaging the raw next-word scores pushes probability toward words both streams share. The team's Joint Contrastive decoder, using a small helper model to steer the streams apart, partly separates them, scoring 0.43 on LAMBADA against a 0.54 single-stream baseline. Mixing streams to get several outputs is [prior art](#), but this work ties the effect to the architecture and shows it can be trained back. Tests stop at 512-token contexts, English text only.

Read More: [Thinking to Recall: How Reasoning Unlocks Parametric Knowledge](#)

Sources:

- [Your Transformer Can Hold Two Thoughts at Once: Evidence of Linear Superposition in LLMs \(arXiv\)](#)
- [Full paper \(arXiv HTML\)](#)
- [Paper page \(Hugging Face\)](#)
- [Superposed Decoding: Multiple Generations from a Single Autoregressive Inference Pass \(arXiv\)](#)

Disclaimer: For information only. Accuracy or completeness not guaranteed. Illegal use prohibited. Not professional advice or solicitation. **Read more:** [/terms-of-service](#)