Yogesh Haribhau Kulkarni • You
AI Advisor (Helping organizations in their AI journeys) | PhD (Geometric Modeling) | Tech Colum...
now •

Frankly, I would have been happier if the topic in the paper below, focused on "embedding of geometry" (rather than "geometry of embedding"!!), especially since my own research involves using neural networks for geometric transformations.

But, never mind. This paper takes a bold step further. It explores whether there's a 'shared geometric structure' across embeddings, regardless of how they're created or even what data they were trained on. Scary, right? 👀

Introduces 'vec2vec', a method that aligns text embeddings from different language models without access to the models or any labeled data. Almost like finding common underlying geometric structure. Talks about 'Platonic Representation Hypothesis', suggesting that large models, trained on different datasets, still learn embeddings that can be aligned through transformation.

What's more? The findings raise serious concerns for vector database privacy, as it shows that sensitive content can be reconstructed from just embeddings. Hmmnnn... again ... Scary, right? 👀

📄 Attached is the original paper with my personal annotations/highlights on top. Would love to hear your thoughts!

Rishi Jha Collin Zhang

#AI #MachineLearning #NeuralNetworks #Embeddings #VectorPrivacy #LLMs #GeometricDeepLearning #AIResearch #DataSecurity #LinkedInScience #vec2vec Antaripa Saha