

Show TranscriptSummarize Video

The exciting world of training generative AI models is about teaching computers to create new content, like text or images, by learning from huge datasets. This training helps AI to understand and recreate the complex patterns found in human language and visual arts. The process is intricate but immensely rewarding, leading to AI that can generate amazingly realistic outputs.

Technical Terms Explained:

**Large Language Models (LLMs):** These are AI models specifically designed to understand and generate human language by being trained on a vast amount of text data.

**Variational Autoencoders (VAEs):** A type of AI model that can be used to create new images. It has two main parts: the encoder reduces data to a simpler form, and the decoder expands it back to generate new content.

**Latent Space:** A compressed representation of data that the autoencoder creates in a simpler, smaller form, which captures the most important features needed to reconstruct or generate new data.

**Parameters:** Parameters are the variables that the model learns during training. They are internal to the model and are adjusted through the learning process. In the context of neural networks, parameters typically include weights and biases.

**Weights:** Weights are coefficients for the input data. They are used in calculations to determine the importance or influence of input variables on the model's output. In a neural network, each connection between neurons has an associated weight.

**Biases (not mentioned in the video):** Biases are additional constants attached to neurons and are added to the weighted input before the activation function is applied. Biases ensure that even when all the inputs are zero, there can still be a non-zero output.

**Hyperparameters:** Hyperparameters, unlike parameters, are not learned from the data. They are more like settings or configurations for the learning process. They are set prior to the training process and remain constant during training. They are external to the model and are used to control the learning process.