Adnan Boz
in/adnanboz

GPT, Bard, LLaMA, Diffusion, GAN, what???

Generative AI Algos in Plain English:

Swipe ➡

Adnan Boz
in/adnanboz

1. Transformer-based Models

The power behind OpenAI **ChatGPT**, OpenAI **DALL-E**, Google **Translate**, Google **Bard**, Microsoft Turing-NLG, Meta OPT, Meta **LLaMA**.

Imagine you have a long text, like a book, and you want to find the main ideas or translate it into another language. It would be a time-consuming task to read everything and understand the context. Transformer-based models can look at the entire text at once and focus on the most relevant parts to make sense of it.

Swipe ➡

Adnan Boz
in/adnanboz

For example, if the model is translating a sentence from English to French, it can look at the whole sentence and figure out the best way to translate each word, considering the context and how the words are related to each other.

During the training process, the model learns to assign different importance to the input data. For example, let's consider "The cat sat on the mat." When translating the word "sat," the model needs to comprehend the relationship between "sat" and the other words in the sentence.

Swipe ➡

Adnan Boz
in/adnanboz

Otherwise word-to-word translation would be very bad. Therefore, it calculates the relevance of each word in the context of “sat”. Compared to the words “on” and “the”, the words "cat" and "mat" are more relevant because they are directly involved in the action “sat”.

After training on a lot of text, it learns to recognize patterns and relationships between words, such as the fact that "cat" and "mat" are related to the action “sat". This contextual information lets the model produce a more accurate translation of the sentence to French. Or, predict the correct next word in your ChatGPT response.

Swipe ➡

Adnan Boz
in/adnanboz

2. Diffusion models

The power behind OpenAI **DALL-E 2**, **Midjourney**, Google **Imagen** and OpenAI GLIDE.

Imagine you're at a weird party with people wearing red-shirts gathered in one corner, and people with blue-shirts in another corner. Now, let's say you want to mix them up so that everyone is mingling and the colors are evenly distributed throughout the room. Diffusion models help us understand how the red and blue shirts would mix in this situation.

Swipe ➡

Adnan Boz
in/adnanboz

They describe how things spread out, or "diffuse," over time. For example, how long it would take for the red and blue-shirted people to mix evenly across the room.

To create a diffusion model for our party, we would consider several factors. For example, we'd look at how fast people move, how many people there are, and the size of the room. By taking these factors into account, we can come up with a model that tells us how long it will take for the colors to be evenly distributed.

Swipe ➡

Adnan Boz
in/adnanboz

In real life, these models are used to study a wide range of scenarios, from how information spreads on social media to how pollutants disperse in the environment helping us make predictions and understand complex situations.

Stable Diffusion is one of the well known diffusion models and allows for generating new images from scratch using text prompts that describe elements to include or exclude. The model can also redraw existing images with added elements based on text prompts.

Swipe ➡

Adnan Boz
in/adnanboz

3. Generative Adversarial Networks (GANs)

The power behind NVIDIA **StyleGAN**, and DeepMind **BigGAN**.

Imagine you're an art enthusiast and want to create an AI that can generate amazing artwork. GAN, or Generative Adversarial Network, is a technique that can help you achieve that. Think of it as a game between two friends, an artist and a detective. The artist tries to create fake paintings that look like famous works of art, while the detective tries to tell if a painting is real or fake.

Swipe ➡

Adnan Boz
in/adnanboz

The artist learns to create more convincing forgeries, and the detective improves at spotting them. GAN works in the same way and consist of two parts, a generator and a discriminator. The generator is like the artist, creating new images, while the discriminator is like the detective, determining if the images are real or fake. As they compete, the generator becomes better at producing realistic images. Eventually, the AI can create impressive artwork that's difficult to distinguish from the real thing.

Swipe ➡

Adnan Boz
in/adnanboz

Once the model is trained, it can be deployed in various applications, such as image synthesis, data augmentation, or even improving the quality of low-resolution images. GANs can be fine-tuned or adapted to specific tasks, making them versatile and powerful tools in the field of artificial intelligence.

Swipe ➡

Adnan Boz
in/adnanboz

4. Variational Autoencoders (VAEs)

Used by Adobe, NVIDIA, Google and OpenAI research projects.

Imagine you're trying to describe the key features of a cat to someone who has never seen one before. You might mention things like pointy ears, whiskers, and fluffy tails. Variational Autoencoders (VAEs) work in a similar way. They understand the important features of different objects or images, like cats, and then use that understanding to create new, similar objects or images.

Swipe ➡

Adnan Boz
in/adnanboz

First, we provide the VAE a bunch of cat pictures. Then it finds the most important features of these cat pictures, simplifying them by focusing only on these key features to learn what a cat looks like.

Then, it can create new cat pictures that we've never seen before by taking the simplified cat features and expanding them back into full images, like describing a cat in more detail to someone who now understands the basics. This can be really useful for tasks like image generation, data compression, and improving recommendation systems.

Swipe ➡

Adnan Boz
in/adnanboz

For more AI topics follow me on https://www.linkedin.com/in/adnanboz/

Or, join one of my AI programs for product managers and business professionals at

AI PRODUCT INSTITUTE
SILICON VALLEY

https://aiproductinstitute.com