

Generative AI

Group 6 Period 4

Lucas, Akash, Rajdeep, Santiago, Daniel, Ishaan



What is generative AI?

Generative AI is a type of artificial intelligence that can create original content. The content that generative AI can make can be anything, such as videos, audios, text, code, and pictures. ChatGPT is an example of generative AI that generates text.

Unlike other types of AI, generative AI does not have a predefined set of rules that it uses to create results and process information. It is because of its focus on creating new original ideas and results that it works like this.



“Generative AI enables users to quickly generate new content based on a variety of inputs.” (Nvidia)



NVIDIA



OpenAI

Application

- Language
- Essay generation, code development, translation
- Audio
- Development of songs, creating realistic noises and sound effects
- Visual
- Produces images, videos, and graphs
- Synthetic data

How text generative AIs (such as ChatGPT) work



Step One: Tokenization

Before the AI can generate a response that you want, it first needs to break your prompt into smaller parts, called **tokens**. Most words are mapped to a single token. Each token is associated with a number, called a **token ID**, which is what the AI will use.

OpenAI's tokenizer:

<https://platform.openai.com/tokenizer>

GPT-3.5 & GPT-4 GPT-3 (Legacy)

Many words map to one token, but some don't: indivisible.

Unicode characters like emojis may be split into many tokens containing the underlying bytes: 🍌

Sequences of characters commonly found next to each other may be grouped together: 1234567890

Clear

Show example

Tokens

Characters

57

252

Many words map to one token, but some don't: indivisible.

Unicode characters like emojis may be split into many tokens containing the underlying bytes: 🍌🍌🍌🍌🍌

Sequences of characters commonly found next to each other may be grouped together: 1234567890

Text

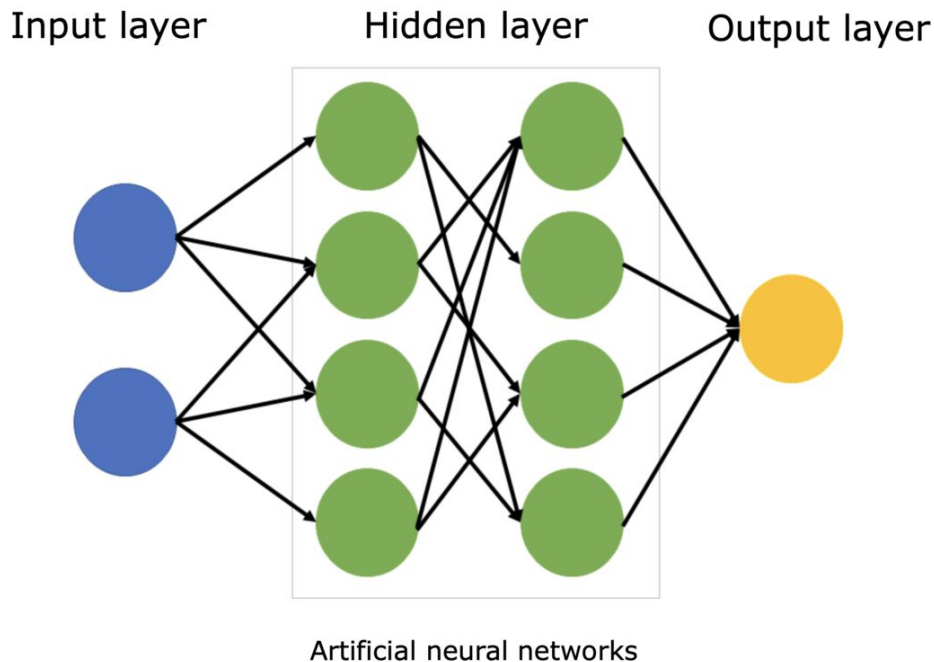
Token IDs

[27247, 16926, 3041, 499, 709, 198, 27247, 16926, 1095, 499, 1523, 198, 27247, 16926, 1629, 2212, 323, 24521, 499, 198, 27247, 16926, 1304, 499, 16106, 198, 27247, 16926, 2019, 47555, 198, 27247, 16926, 3371, 264, 10457, 323, 13194, 499]

Btw this is totally not rickroll

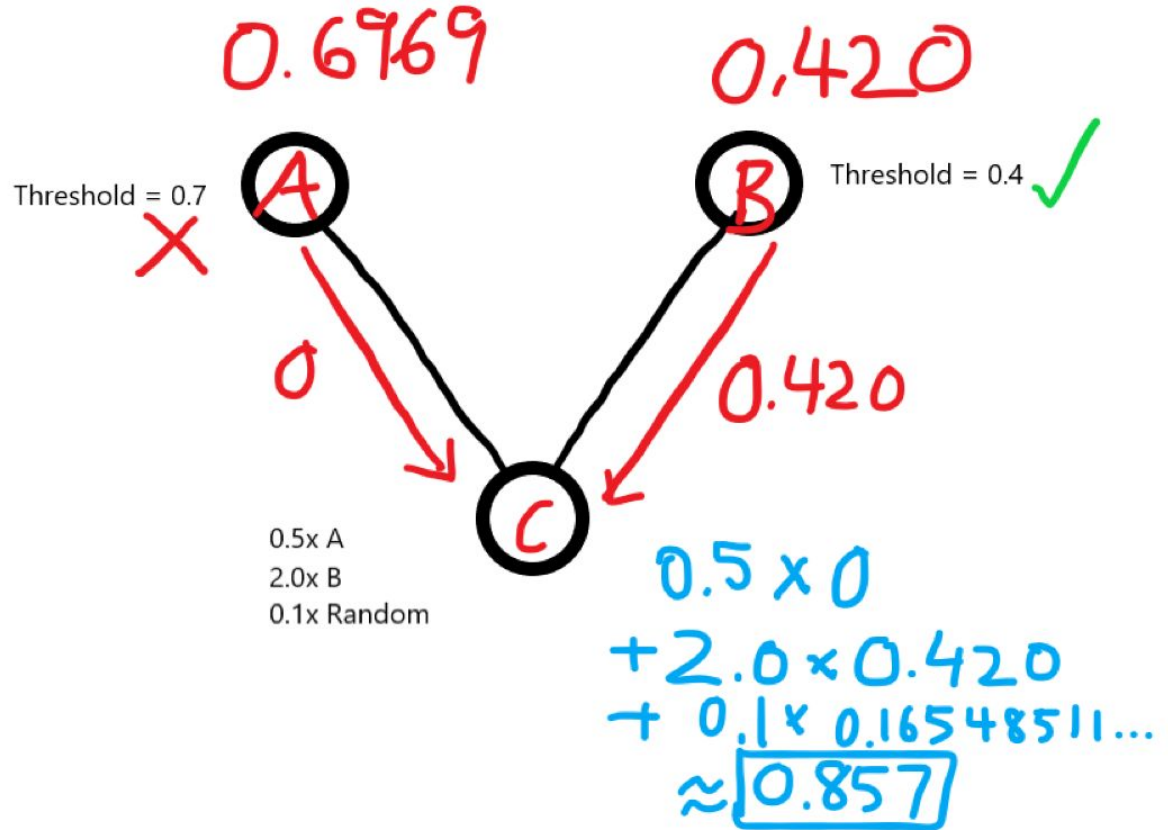
Step Two: Neural Network

A neural network is a machine learning model that is based off of how neurons are actually connected in brains. This is the part that is responsible for generating the response. The tokens from the previous step are fed into the input layer, and go through many layers of nodes called hidden layers until it reaches the output layer. Advanced models such as GPT-3 use upwards of 175 billion nodes. The exact parameters that it uses depends on how it was trained.



Step Two: Neural Network

When data passes through the hidden layers, each node's value is determined by a weighted average of the nodes connected to it as well as some randomness, with the exact weights determined during the training process. However, data is only passed through if the value meets a certain threshold.



Step Three: Output

Once the output layer is reached, the AI spits out words that it thinks is the best to begin the response (based on the values of the output nodes). Then it goes back and and spits out the second part, then the third part, etc.

You can think of it as if the AI were guessing the answers to a series of multiple choice questions on what word should come next in the response.

| 1. First word | 2. Second word |
|---------------|----------------|
| A. The 5% | A. sorry 80% |
| B. Of 5% | B. a 10% |
| C. It 20% | C. going 5% |
| D. I'm 70% | D. not 5% |

| 3. Third word |
|---------------|
| A. it 5% |
| B. I 10% |
| C. but 60% |
| D. that 25% |

...

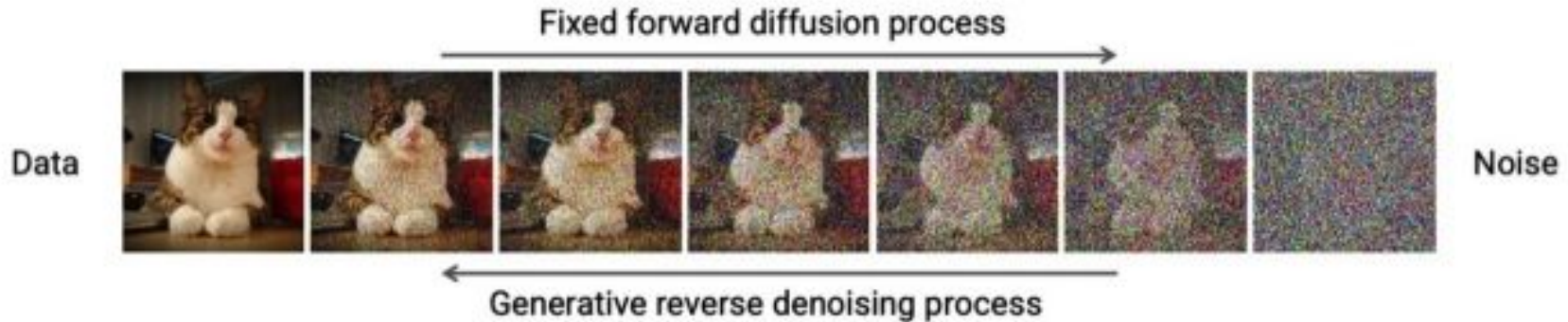


I'm sorry, but as an AI language model,

How does AI image generation work?

Diffusion models are the main image generating model for AI.

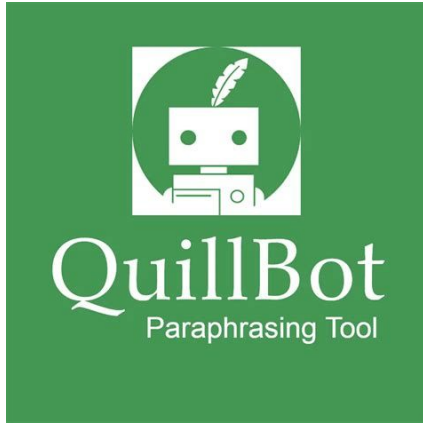
The neural network learns how to predict the addition of “noise”, and learns how to reverse the process. You can create whole new images by adding random noise and telling it to work backwards from there.



How AI image generation works (continued)



Generative AI in Action



QuillBot (Paid)

An online writing tool that can automatically paraphrase text and change it to adopt a particular style (narrative, academic, etc.)



Microsoft Copilot (Free)

A chatbot developed by Microsoft as part of Bing Chat.



DALL-E (Free/ Paid)

An OpenAI product that generates detailed images based on descriptions.

Challenges and Limitations

Bias

- The data the AI is trained on can be biased, leading to results that don't accurately reflect what it should reflect.

Outdated Data

- If data is outdated, data produced after it was collected by the AI will not be represented. (Ex. ChatGPT only collects data from up until September 2021, so you can't ask it about events that have happened afterwards.)

Automated Theft

- Generative AI can use automatically steal artists' works without permission

Future Trends

- Improved Model Capabilities
- Multimodal Generative Models
- Fine-Tuning and Customization
- Explainability and Interpretability
- Ethical AI and Bias Mitigation

<https://www.youtube.com/watch?v=PVVUPUJeezo>



Generative AI Kahoot

Go to Kahoot.it and wait for us to pull out the code.

Works Cited

<https://www.nvidia.com/en-us/glossary/generative-ai/>

<https://zapier.com/blog/how-does-chatgpt-work/>

<https://www.ncl.ac.uk/academic-skills-kit/good-academic-practice/artificial-intelligence/use-and-limitations/>

<https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai>

<https://www.techtarget.com/searchenterpriseai/definition/generative-AI#:~:text=Generative%20AI%20vs.%20AI,-Generative%20AI%20focuses&text=Generative%20AI%2C%20as%20noted%20above,neural%20networks%20and%20reinforcement%20learning.>