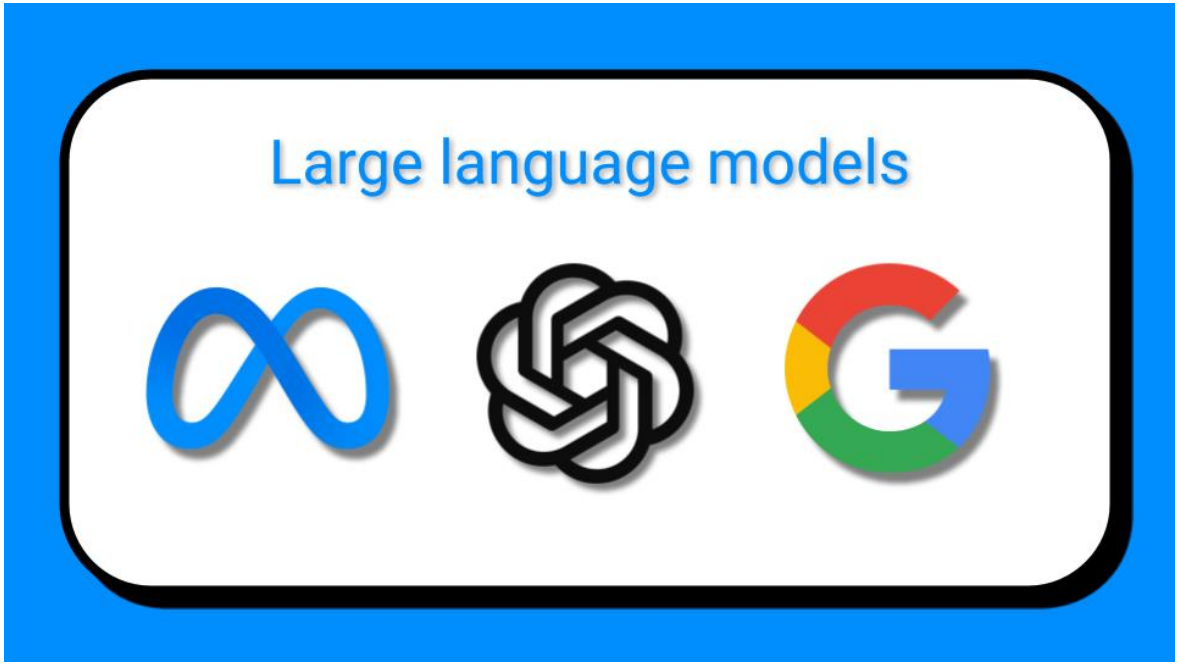
Generative AI

A Quick Overview of GenAI

What you will learn?

- What is Generative AI?
- Why Generative AI Booming?
- Where Generative AI Exists?
- Discriminative vs Generative AI
- What is LLMs?
- Domain based Applications of Generative AI

Generative AI



- ChatGPT
- Google Bard
- Meta Llama 2
- Perplexity
- Claude

Generative Model

Questions

Responses

What is Generative AI?

Generative AI generate new data based on training sample. Generative model can generate Image, Text, Audio, Videos etc data as output.

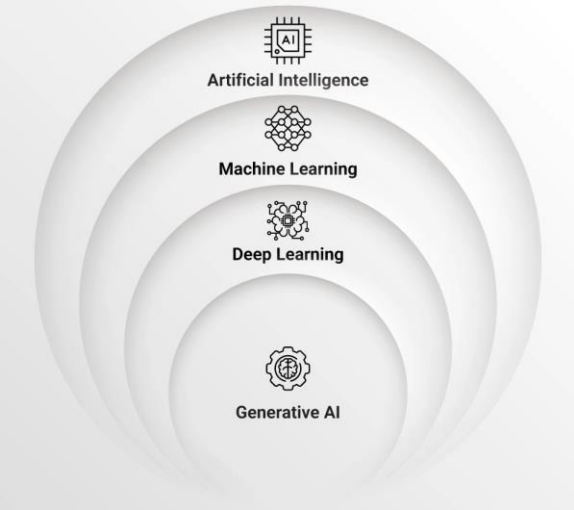
Generative Model you can divide into two major segment

- Generative unimodal
- Generative Multimodal

Why is Generative AI Booming

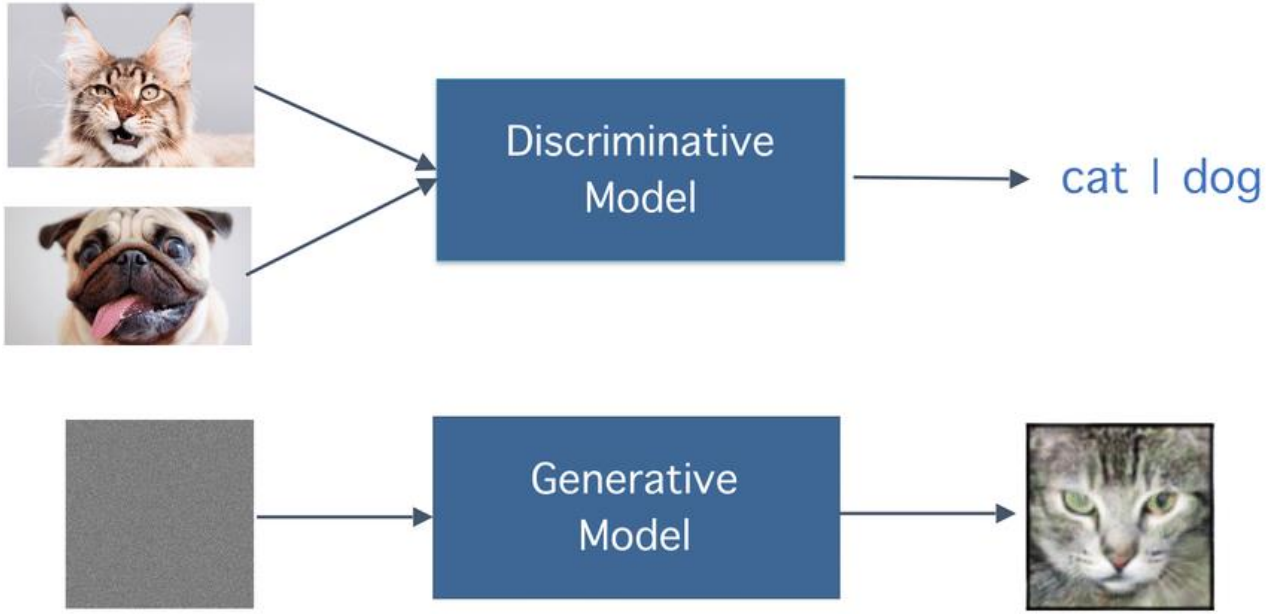
- Transforming creativity and automation.
- Increased Investment and Research
- Applications across various industries

Where Generative AI Exists

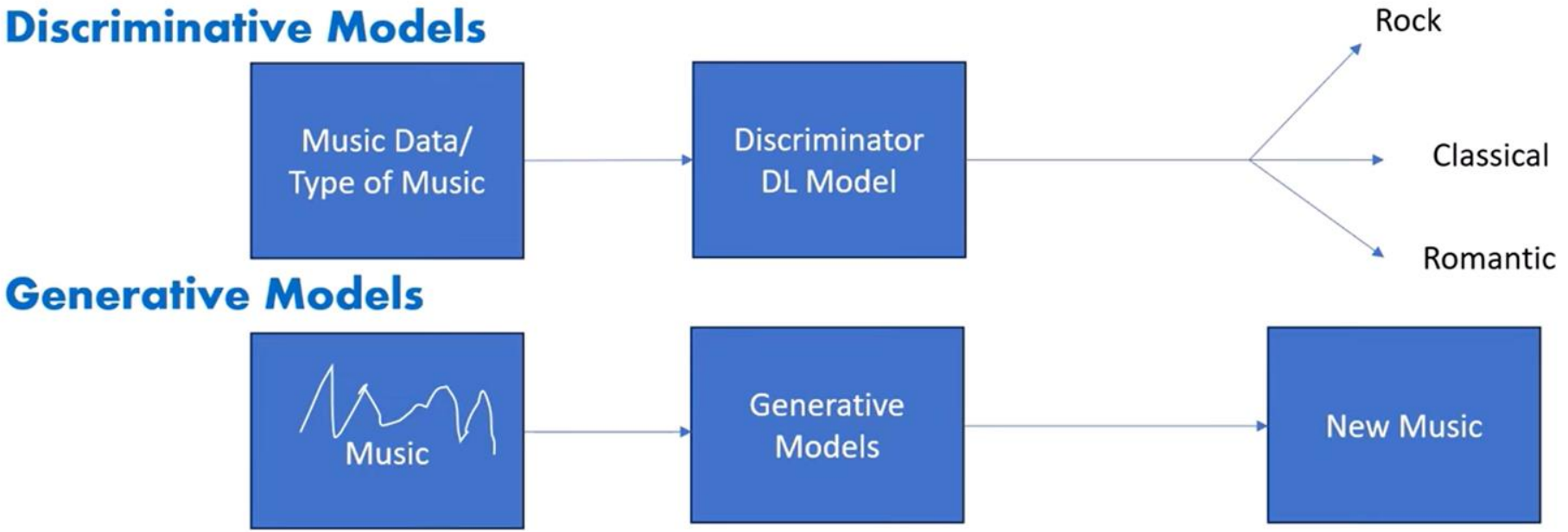


- Machine Learning is the subset of Artificial Intelligence
- Deep Learning is the subset of Machine Learning
- Generative AI is the subset of Deep Learning

Discriminative vs Generative Model



Discriminative vs Generative Model



What is LLMs?

Large Language Models (LLMs) are foundational that use deep learning models like transformer to process and understand natural. These models are trained on massive amounts of text data to learn patterns and entity relationships in the language.

It is a language model which is responsible for performing task such as **text** to **text generation , text** to **image generation , image** to **text generations and image** to **image generation.**

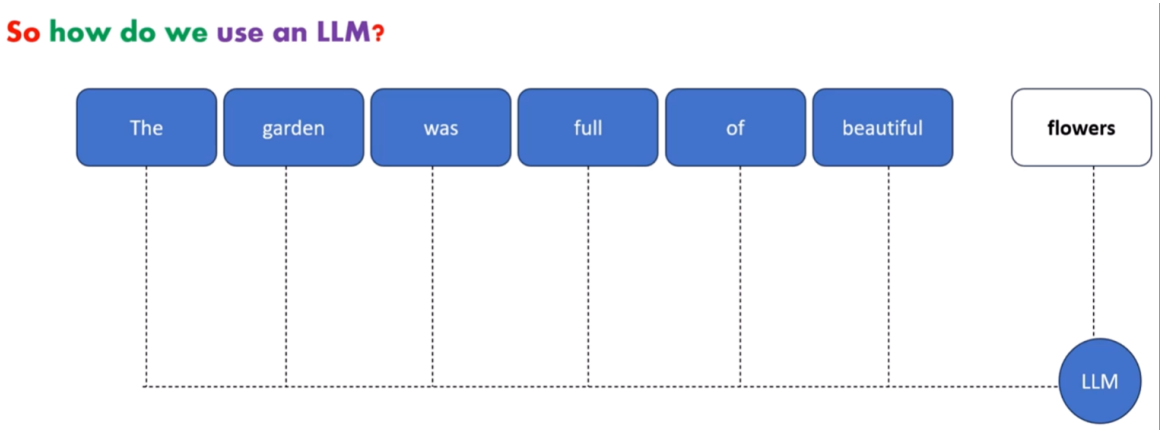
Application Areas

- Text Generation
  - Chatbots, story generation, summarization.
- Image Generation:
  - DALL-E, Stable Diffusion, MidJourney.
- Music and Audio:
  - Tools like OpenAI Jukebox.
- Video Synthesis and Animation:
  - Applications like Deepfake creation, editing and Video Summarization.

What is LLMs?

A large Language model is a trained deep learning model (Transformers) that understands and generate text in a human like fashion.

LLMs are good at Understanding and generating human language



Why we call it Large Language Model?

Because of the size and complexity of the Neural Network as well as the size of the dataset that it was trained on.

Researchers started to make these models large and trained on huge datasets

That they started showing impressive results like understanding complex Natural Language and generating language more eloquently than ever.

What LLMs Powerful?

In case of LLM, one model can be used for a whole variety of tasks like:-

generation, summarizer, translation, classification

& so on …

And now it has Multimodal capability also.

Generative Models

- **GANs (Generative Adversarial Networks)**:
  - Basic architecture: Generator aur Discriminator.
  - Common GAN types: DCGAN, CycleGAN, StyleGAN.

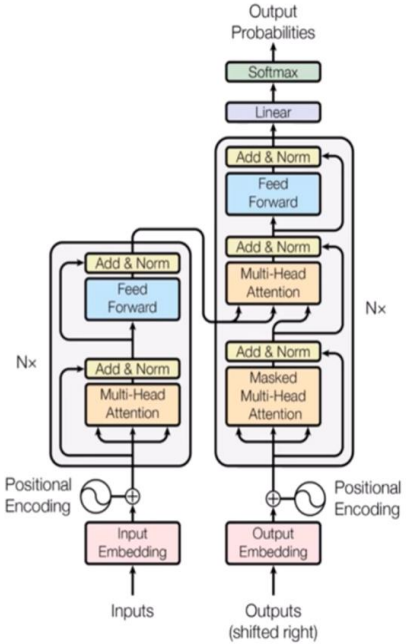
- **VAEs (Variational Autoencoders)**:
  - Encoding/Decoding and latent variable modeling.

- **Transformers in Generation**:
  - GPT series, BERT for masked token generation.
  - Applications in text generation.

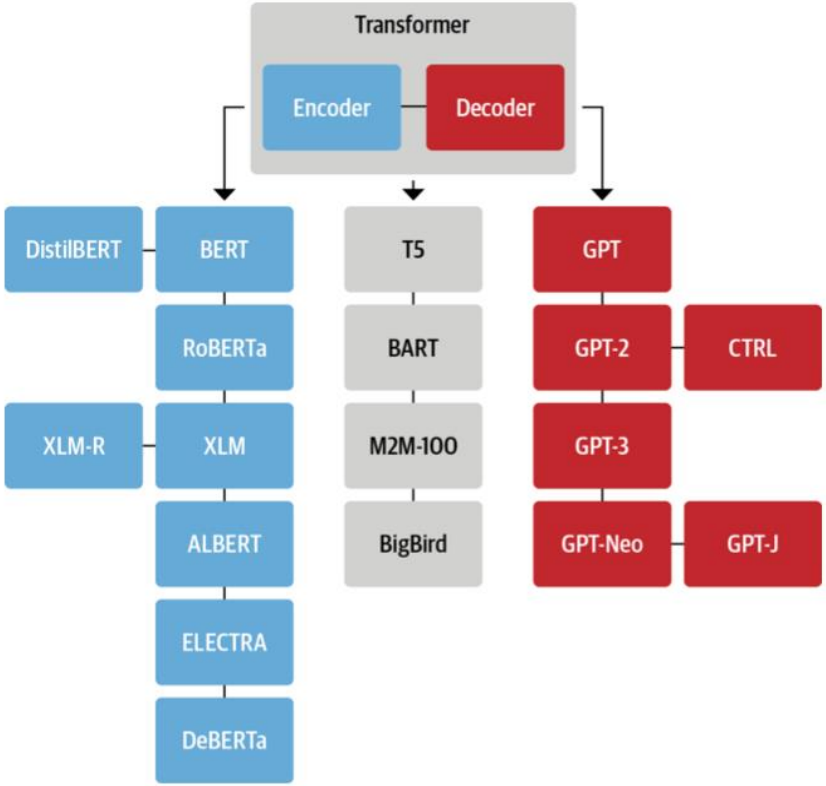
Transformers: Architecture Behind LLMs

Large Language models are based on transformer a type of Neural Network

Architecture invented by Google.



Transformer Tree



Few initial milestone in LLMs

- **BERT:** Bidirectional Encoder Representations from Transformers (BERT) was developed by Google
- **GPT:** GPT stands for "Generative Pre-trained Transformer". The model was developed by OpenAI
- **XLM:** Cross-lingual Language Model Pretraining by Guillaume Lample, Alexis Conneau.
- **T5:** The Text-to-Text Transfer Transformer It was created by Google AI
- **Megatron:** Megatron is a large, powerful transformer developed by the Applied Deep Learning Research team at NVIDIA
- **M2M-**100: multilingual encoder-decoder (seq-to-seq) model researchers at Facebook

Famous large language models

- Meta Llama
- Mistral AI
- Google Gemini
- Anthropic Claude
- OpenAI GPTs
- Nvidia AI

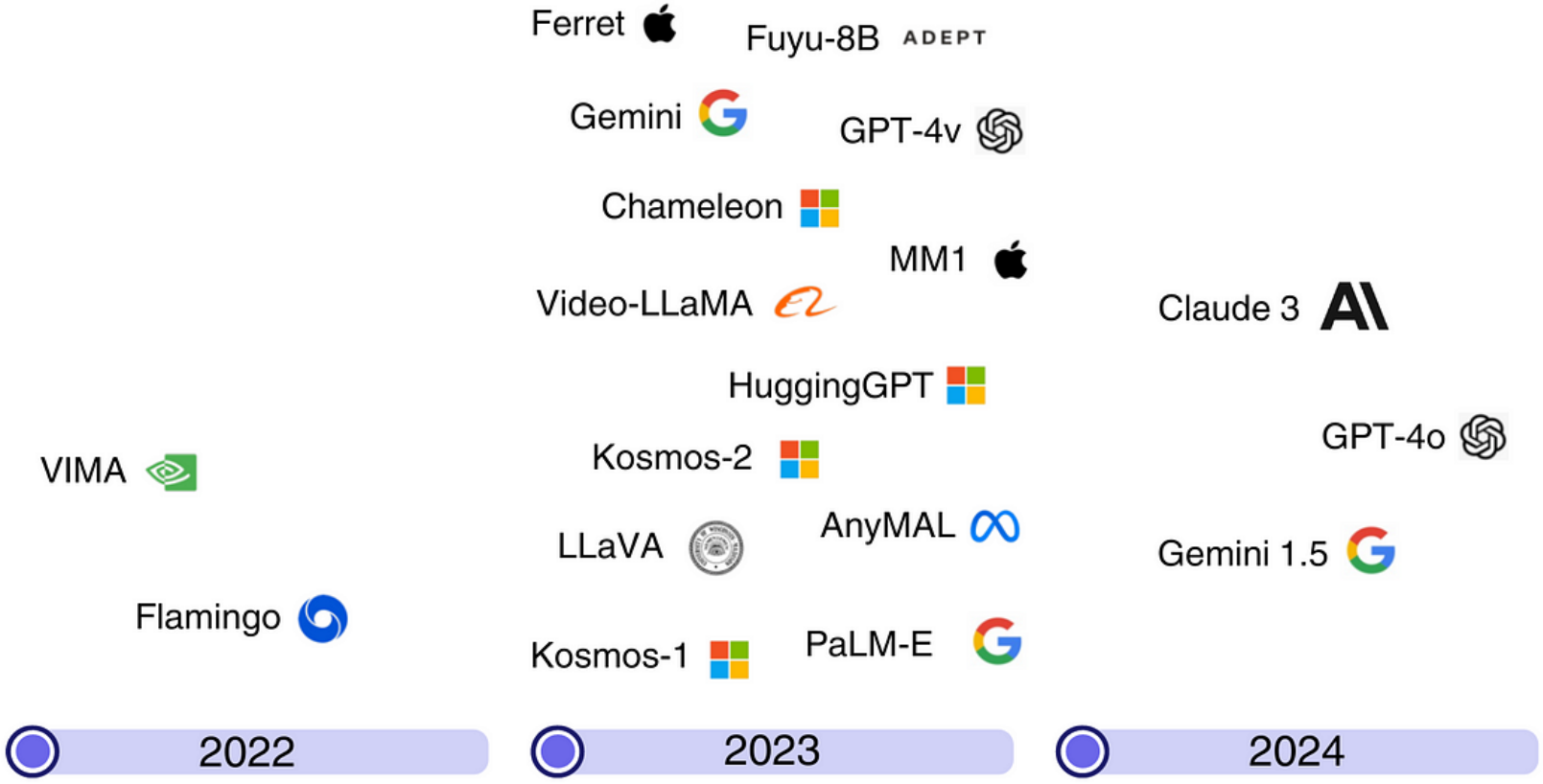
Famous Open-Source Models

- BLOOM
- Llama 2
- PaLM
- Falcon
- Claude
- MPT-30B
- StableIm
- Flan T5

What is a multimodal Transformer

Multimodal transformer models extend the transformer architecture to incorporate other modalities, such as images and audio. These models have achieved state-of-the-art performance on various multimodal tasks, including visual question answering, image captioning, and speech recognition.

Evolution of Multimodal Large Language Models (MLLMs)



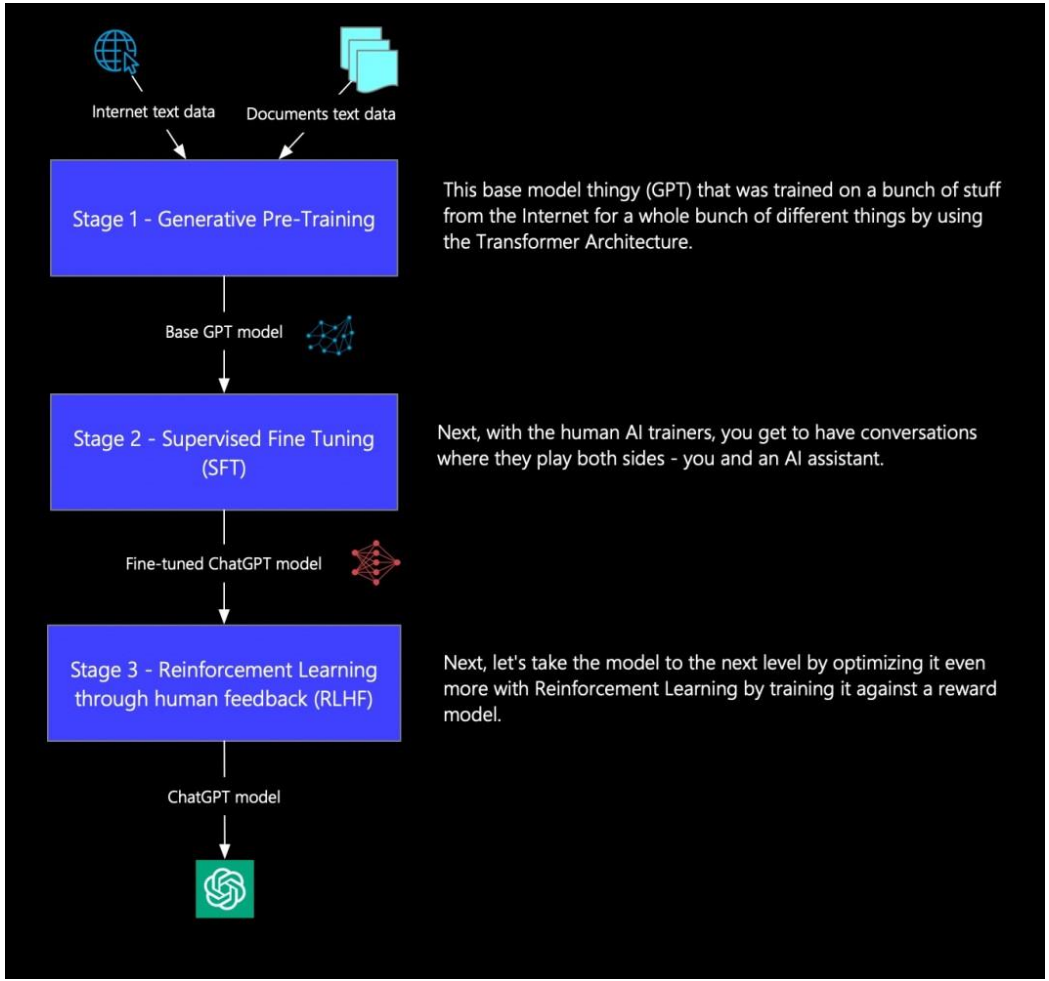
Multimodal Capable Model

- LLaVA : Research at Microsoft
- Vision Transformer(ViT): Google Research, Brain Team
- Qwen-VL : Alibaba
- Claude 3.5 Sonnet: Claude
- GPT-4v, 4o, Sora, DALL-E : OpenAI
- Stable Diffusion: Stability AI
- CLIP (Contrastive Language–Image Pre-training): OpenAI
- Google Gemini 1.5
- Runway Gen-2

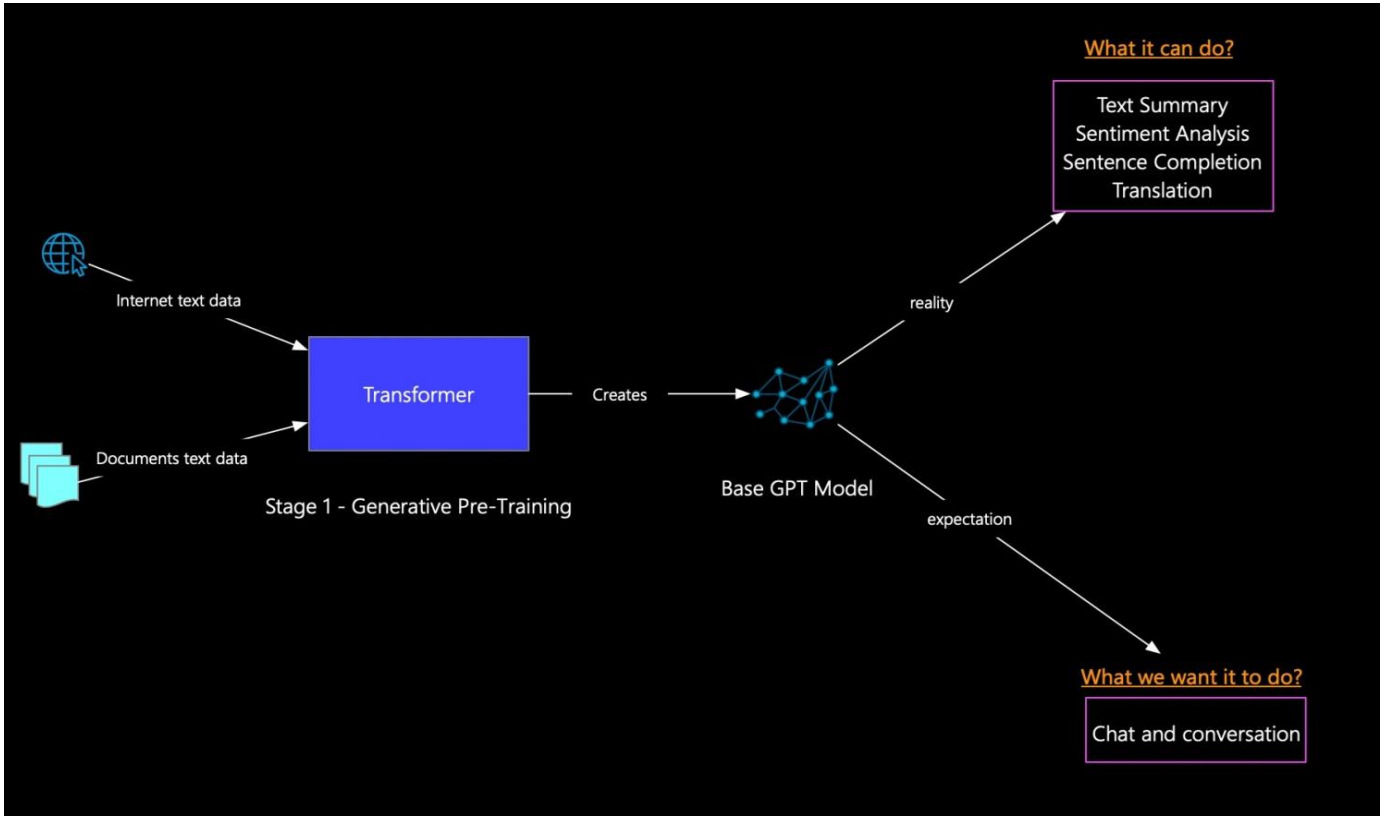
How ChatGPT was trained?

ChatGPT is a complete software which is using LLM GPT-3.5 Turbo, GPT-4 and GPT-4o. It has trained on an immense amount of data available all over the internet.

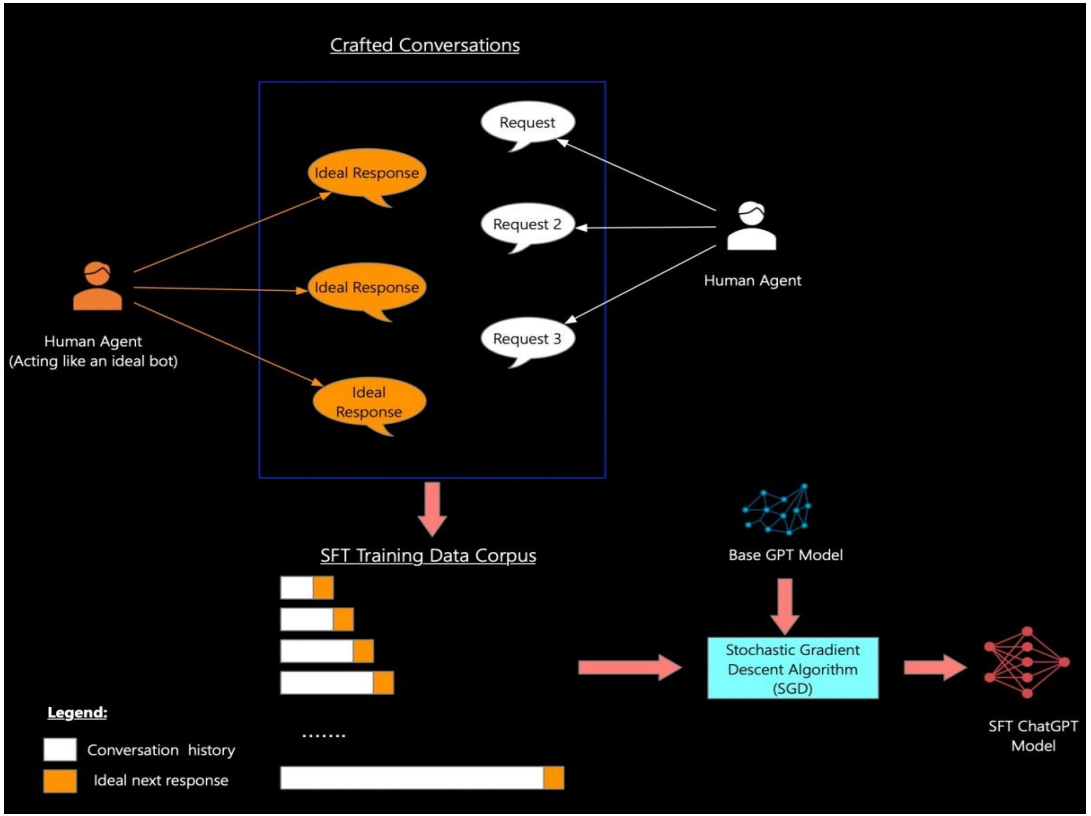
1. Generative pre-training
2. Supervised fine-tuning
3. Reinforcement learning


Internet text data
Documents text data
Stage 1 - Generative Pre-Training
This base model thingy (GPT) that was trained on a bunch of stuff from the Internet for a whole bunch of different things by using the Transformer Architecture.
Base GPT model
Stage 2 - Supervised Fine Tuning (SFT)
Next, with the human AI trainers, you get to have conversations where they play both sides - you and an AI assistant.
Fine-tuned ChatGPT model
Stage 3 - Reinforcement Learning through human feedback (RLHF)
Next, let's take the model to the next level by optimizing it even more with Reinforcement Learning by training it against a reward model.
ChatGPT model

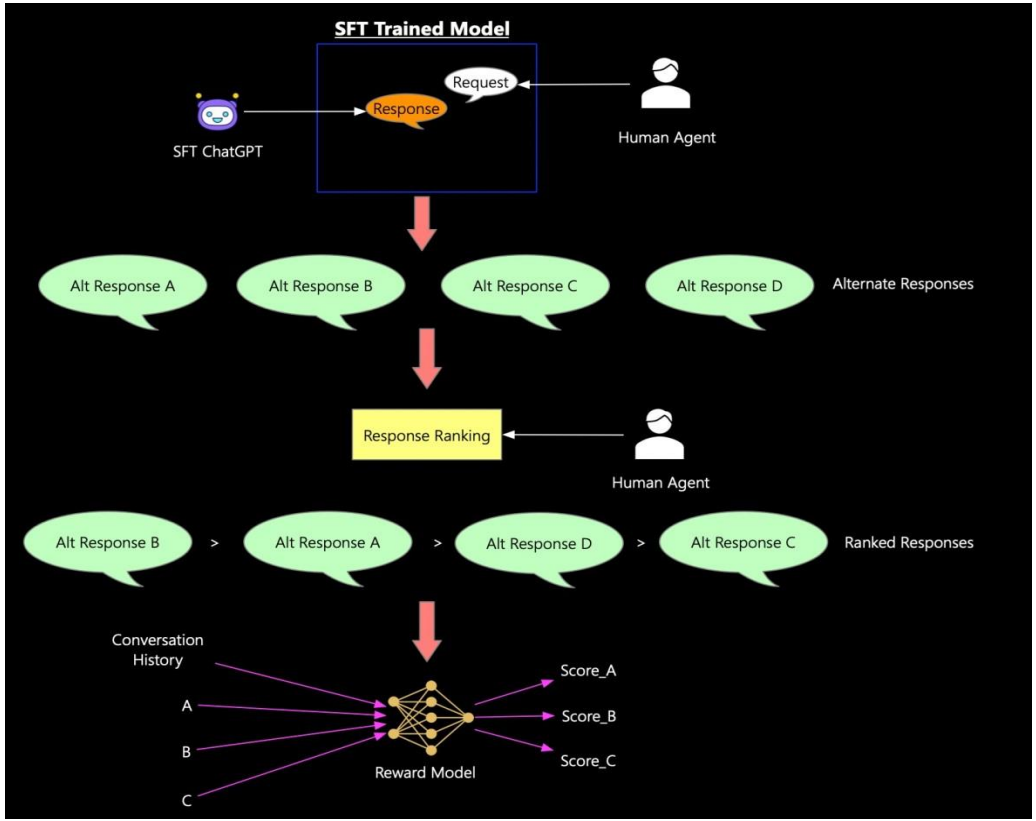
Generative Pre-Training



Supervised Fine-Tuning (SFT)



Reinforcement Learning through Human Feedback (RLHF)



Industry	Usecase
Medical and Healthcare	Primary Health Doctor, Vision For Blind
Finance	Stock Trading Bot, Investment Analysis
Business	Customer Support Chatbot, Market Research
Information Technology	Coding Assistant, Process Automation
Marketing	SEO Agent, Content Writing Agent
Human Resources	Resume ATS & JD Writer, Interviewer

Industry	Usecase
Social Media	Instagram and LinkedIn Post Content Writer
BPO	Sales Agent, Technical Support
Education	AI Teacher, Test QA Generator
Entertainment	Movie Recommendation System, Youtube Automation
AI	Low/No Code Tool For AI Architecture(Finetuning and RAG), Synthetic Data Generation
Personal	JARVIS, Rabbit R1 Clone, Email Automation, Zapier with AI App

Thank You!