Generative AI Basic

Presented by
Md Abdullah Al Hasib
CSE(Islamic University)

What we will learn today

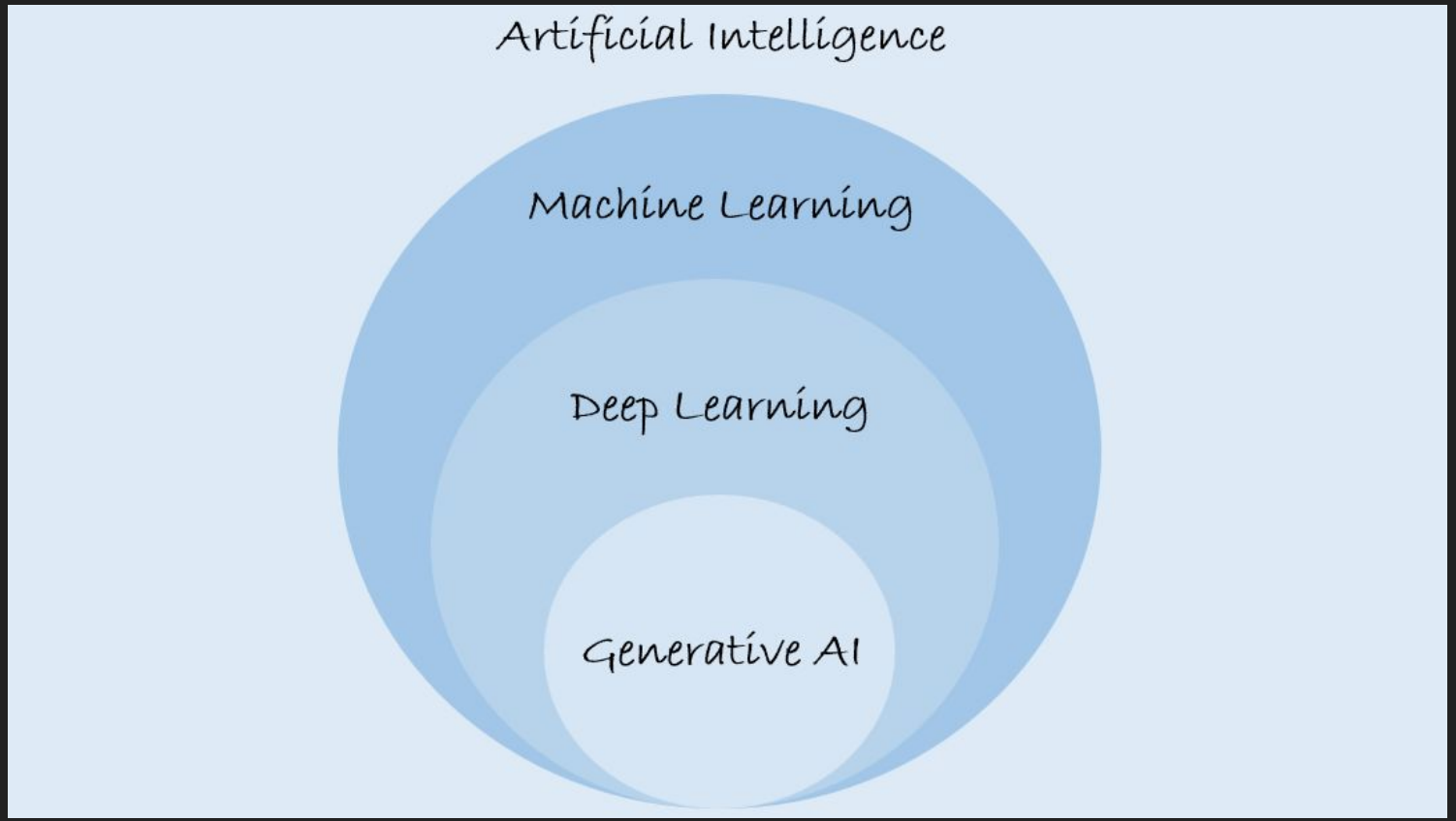
- What is Generative AI?
- Relation among AI,ML, DL & Generative AI
- Importance of Generative AI
- Types of Generative AI
- Generative Image Models
- Large Language Models
- Multimodal LLMs
- Improved the performance of LLMs
- Libraries of LLMs
- Applications of Generative AI
- Limitations
- Future of Generative AI

What is Generative AI?

Generative AI is artificial intelligence capable of generating text, images, videos, or other data using generative models, often in response to prompts.Generative AI models learn the patterns and structure of their input training data and then generate new data that has similar characteristics.

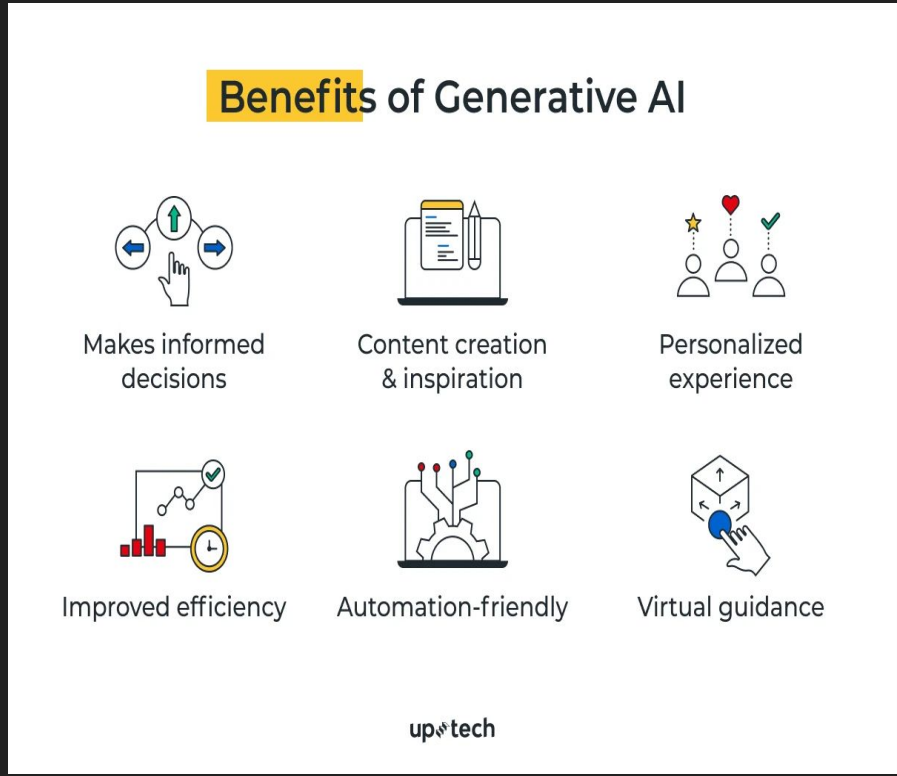
Most popular Applications are ChatGPT, DALL·E, GitHub Copilot, Mid Journey, Jasper, Synthesia etc.

Relations among AI, ML, DL & Generative AI



Importance of Generative AI

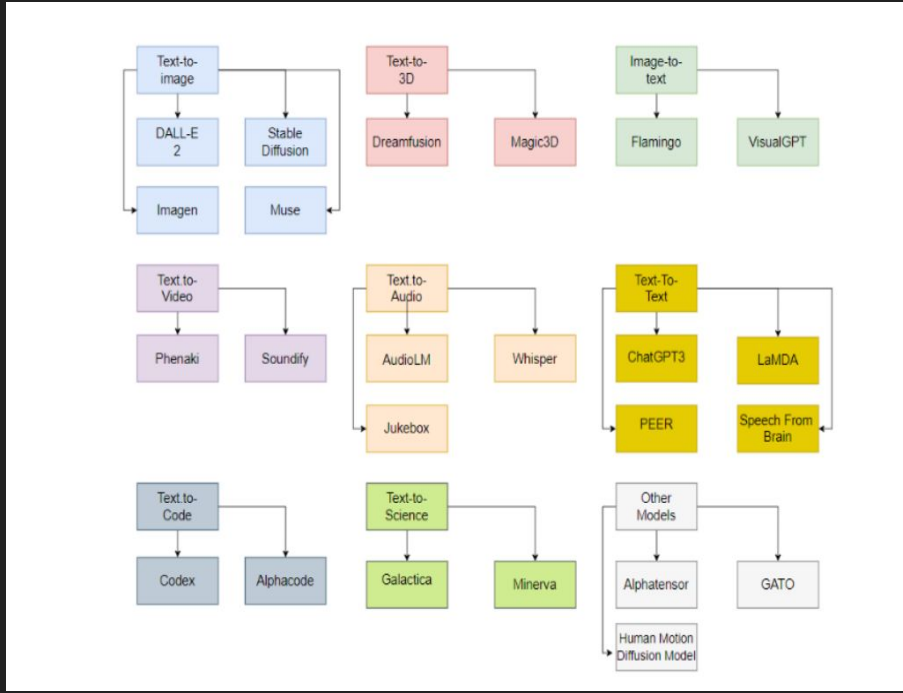
- Data Augmentation
- Increase creativity
- Content creation
- Efficiency & Cost reduction
- Healthcare advancements
- Enhancing human AI collaboration
- Increase productivity
- Increase job opportunities
- Automation of work



Types of Generative AI

Generative AI can be categorized based on the type of data it generates, such as images, text, audio, or combinations of multiple modalities.

- Generative image models
- LLMs
- Audio synthesis models
- Generative Video Models
- Multimodal Generative AI
- Generative Code models



Generative Image Models

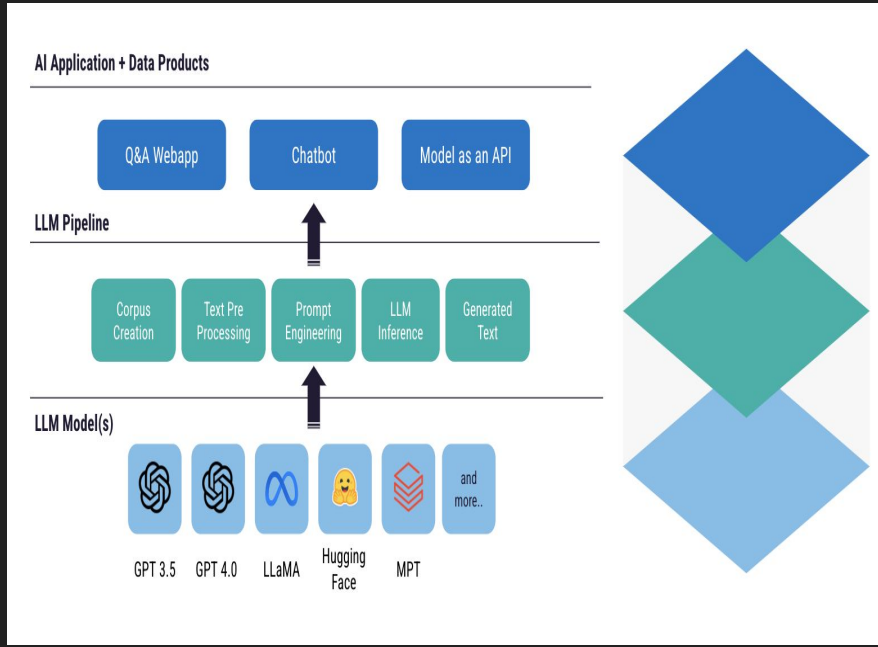
Generative Image Models are designed to create or generate new images from input data, often based on patterns and features learned from large datasets.Some models are GANs, VAEs, ,Diffusion Models(DALL·E, Stable Diffusion) etc.

Applications are Art Generation, Style Transfer, Video Game Development(Generate characters, landscapes, and environments in gaming, Cartoon Character generation, Generate new clothing designs and styles, Create unique room layouts or furniture designs, Generate synthetic medical images,Generate higher-quality images from low-resolution inputs, Create additional synthetic data to enhance datasets and improve model training.

Large Language Models (LLM)

Large Language Model is designed to understand and generate human-like text. LLMs are trained on vast amounts of text data and can perform a variety of language-related tasks such as language translation, summarization, text completion, content recommendation, conversation and others.

GPT-4 , BERT, T5,LLaMA, Claude, Gemini are some examples of LLM models.



Multimodal LLMs

Multimodal LLM refers to a type of AI model that processes and generates multiple types of data, or modalities such as text, images, audio,video etc. Unlike traditional LLMs, which typically focus on uni modality, multimodal LLMs integrate various forms of input and output, allowing for more versatile interactions and use cases. CLIP (OpenAI), Gemini pro vision, Flamingo are multimodal llms.

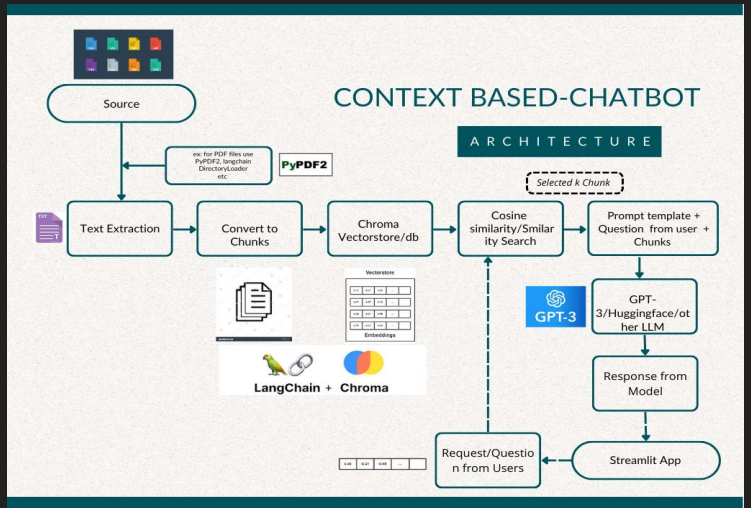
Image Captioning, Visual Question Answering (VQA), Interactive AI Systems, Content Creation, integrate medical imaging data (like MRIs) and text data (medical reports) to assist in diagnosis are some applications.

Libraries to perform LLMs

- Hugging face
- Open API
- Langchain
- LlamaIndex
- Ollama
- Heystack
- Gemini
- Pinecone
- Mistral AI and others

Applications of Generative AI in Industry

- Healthcare (reconstruct medical images, design personalized treatment plans, generate potential new drug compounds etc)
- Manufacturing (generate product designs, identify defects etc)
- Retail & Ecommerce(Product Recommendations, content creation, Chatbots & Virtual Assistants etc)
- Education (Personalized Learning, create educational materials like quizzes, tutorials etc)
- Legal & Compliance
- Fashion, Agriculture, Architecture etc…

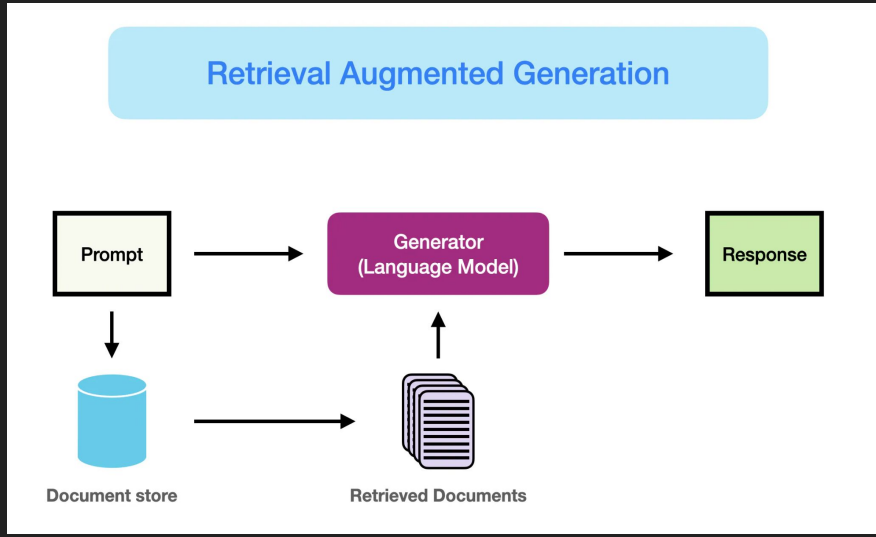


Limitations

- Require high computational resources
- Hallucinations
- Requires massive amount of data
- Lack of explainability of LLM models
- Requires high skilled persons
- Require high cost
- Ensuring fairness and mitigating biases is a critical concern.
- Determining ownership and attribution poses legal and ethical challenges.
- Misuse for spreading misinformation, generating deep fakes, and other malicious activities.
- Unemployment

Improve the Performance of LLMs

- Fine-Tuning and Domain Adaptation
- Model Pruning, Quantization
- RAG (Retrieval-Augmented Generation)
- Prompt Engineering
- Reinforcement Learning with Human Feedback (RLHF)
- Knowledge Distillation



Future of Generative AI

- Enhanced Creativity and Innovation
- Increase Multimodal applications
- 3D content creation
- Improved Human-AI Collaboration
- Bias Mitigation
- Increase explainability & Interpretability
- Increase No-Code/Low-Code AI Platforms etc

Next: History of Generative AI

Thanks For Watching