## Generative AI

# What you will learn?

* Generative AI
* Large Language Model (LLMs)
* OpenAI
* LangChain
* Vector Database
* Llama Index
* Open Source LLM model
* End to End Project

# 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.

So generative AI is a very huge topics,

* Generative Image Model 🡪 Gan img to img 2018 - 2019
* Generative Language Model 🡪 LLM text to img, text to text, img to img

## Where Generative AI Exists.

* Machine Learning is a subset of Artificial Intelligence
* Deep Learning is a subset of Machine Learning
* Generative AI is a subset of Deep Learning

Generative AI is a subset of deep learning and Generative models are trained on huge amount of data. While training the generative model we don’t need to provide a label data, It’s is not possible when we have a huge amount of data, So, It’s just try to see the relationship between the distribution of the data. In Generative AI we give unstructured data to the LLM model for training purpose.

# What is LLMs?

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

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

## What is LLMs?

A large language model is a trained deep learning model that understands and generate text in a human like fashion.

LLMs are good at understanding and generating human language.

## Why we can it Large Language Model?

Because of the size and complexity of the neutral 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 makes LLM so Powerful?

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

**Text generation, Chatbot, Summarizer**, **translation, code generation & so on …**

So, LLM is a subset of deep learning & it has some properties merge with Generative AI

## LLMs Model Architecture

Large Language Model are based on transformer a type of neutral network Architecture invented by Google.

## Few milestone in large language model

* **BERT:** Bidirectional Encoder Representations from Transformers (BERT) was developed by Google
* **GPT:** GPT stands for “Generative Pre-trained Transformer”. The model was developed OpenAI
* **LXM:** 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 Research team at NVIDIA.
* **M2M-**100: multilingual encoder-decoder (seq-to-seq) model researchers at Facebook.

## Open AI Based LLM models

**MODELS DESCRIPTION**

GPT-4 A set of models that improve on GPT-3.5 and can understand as well as generate natural language or code

GPT-3.5 A set of models that improve on GPT-3 and can understand as well as generate natural language or code

GPT base A set of models without instruction following that can understand as well as generate natural language or code

DALL-E A model that can generate and edit images given a natural language prompt

Whisper A model that can convert audio into text

Embeddings A set of models that can convert text into a numerical form

Moderation A fine-tuned model that can detect whether text may be sensitive or unsafe

GPT-3 Legacy A set of models that can understand and generate natural language

## Other Open Source Models

* BLOOM
* Llama 2
* PaLM
* Falcon
* Claude
* MPT-30B
* Stablelm

So on …

## What can LLMs be used for?

* Text Classification
* Text Generation
* Text Summarization
* Conversation AI like Chatbot, Question Answering
* Speech recognition and Speech Identification
* Spelling Corrector

So on …

## How ChatGPT was trained?

Internally using a LLM which is gpt-3.5 or gpt-4

It has trained on a large amount of data which is available all over the internet.

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

## Prompt Designing

All the text that we feed into an LLM as input is called a prompt and this whole art is known as prompt design, which is about figuring out how to write and format prompt text to get LLMs to do what you want.

# What is OpenAI API?

This OpenAI API has been designed to provide developers with seamless access to state of art, pre-trained, artificial Intelligence models like gpt-3 gpt-4 dall e whisper, embeddings etc so by using this openai api you can integrate cutting edge ai capabilities into your applications regardless the programming language.  
So, the conclusion is by using this OpenAI API you can unlock the advance functionalities and you can enhance the Intelligence and performance of your application.

# Function Calling

Learn how to connect large language models to external tools.

# LangChain

LangChain is an open-source framework that simplifies the development of applications powered by large language models (LLMs). It provides tools and abstractions that make it easier to chain together different components and create complex, context-aware applications. LangChain is available as a Python and JavaScript library.

## Prompt Template

Prompt templates help to translate user input and parameters into instructions for a language model.

## Agent

In LangChain, an agent is an AI entity that leverages a Large Language Model (LLM) to decide on a course of action and then execute that action by interacting with external tools. Essentially, it's an autonomous system that uses an LLM as a reasoning engine to determine the best way to achieve a specific task, and then uses tools to carry out the necessary steps.

## Chain

Central to LangChain is a vital component known as LangChain Chains, forming the core connection among one or several large language models (LLMs). In certain sophisticated applications, it becomes necessary to chain LLMs together, either with each other or with other elements.

## Document Loader

Document loaders are specialized components that retrieve and format data from various sources into a standardized Document object, which LangChain can then process. They act as data connectors, bringing in information from files, websites, databases, and more, and converting it into a format suitable for LangChain applications. Each loader is designed for specific data formats and sources, providing flexibility and versatility for different use cases.

## Memory

Memory refers to the ability of a language model to retain and access information from past interactions, enabling it to maintain context and provide more coherent and relevant responses in a conversation or chain of operations. Essentially, it allows the model to "remember" what was said earlier, unlike stateless LLMs that treat each input independently.