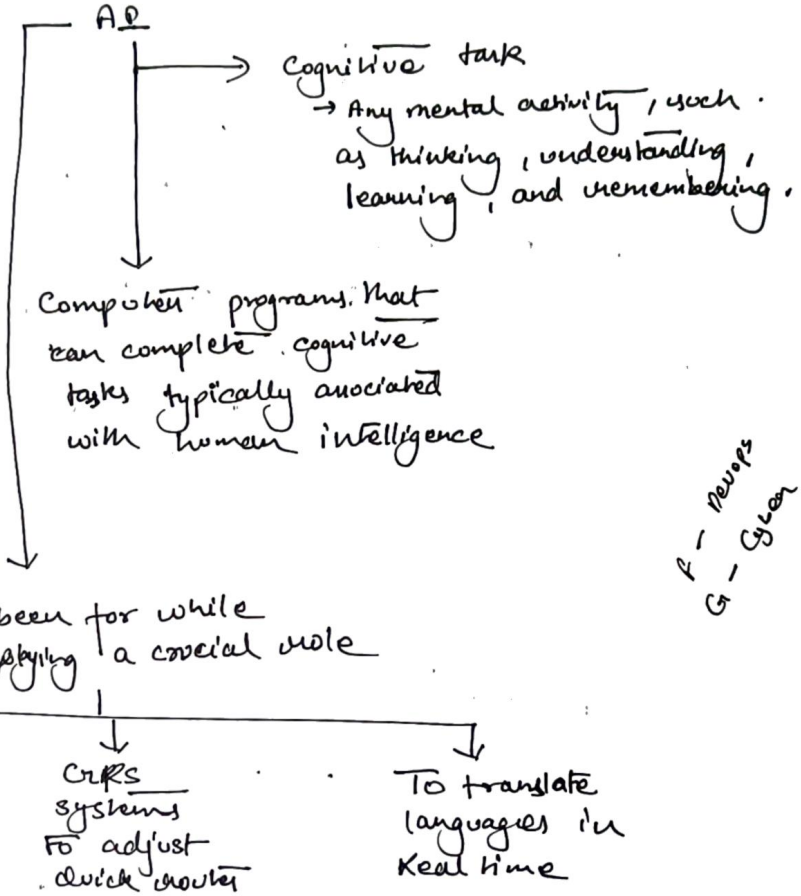
Generative AI (Overview)

(AI Essentials)

→ AI stands for Artificial Intelligence which used usually to make machine task less time consuming



P - neurips
G - Cyber

Machine Learning: A subset of AI focused on developing computer programs that can analyse data to make decision or prediction

It is developed by using a training set which is a collection of data used to teach AI

For ML programs to work effectively ~~It is impor~~ there training data matters. A fundamental issue is biasness within the training data. This could unintentionally cause an AI tool to generate inaccurate or ~~wrong~~ ~~output~~ unintended outputs.

Generative - AI

→ AI ~~that~~ that can generate new content, such as text, images, or other images.

How gen AI works with Natural Language
- Provide Input
- Data is processed
- Output is generated

Benefits of Gen AI
- Boost your productivity
- Help you avoid mistakes
- Improve your decision-making process

Conversational AI tool : A generative AI tool that processes text requests and generates text responses

Capabilities of AI
- Generate content
- Analyze information quickly
- Answer questions
- Simplify day-to-day tasks

Limitations of AI :
- → Can't learn independently
- → Can reflect or amplify biases
- → Can contain inaccuracies.

AI augmentation : The process of using AI to improve a work product, whether by making it easier to do or higher in quality.

AI automation : The process of using AI to accomplish tasks without any action on the user's part.

AI automation requires practice, successful businesses can set themselves apart by applying a people-first-approach in the AI services and jobs.

Think of AI as a collaborative workspace that thrives on diverse perspective and job

Consider a public relation manager who wants to use AI for its press releases. To make sure they get an optimal result, they should ~~app~~ give a human side to the AI tool. It's essential that they collaborate with other members of the team, such as managers, members of the editorial team, ~~gordab~~ legal department etc.

How to boost Productivity with AI:

→ An AI tool is an AI-powered software that can automate or assist users with a variety of tasks

Eg: Speak O → Helps with oos communication skills

Neuro filter → It is a built in AI feature in Adobe photoshop ~~which helps~~ which helps in photo editing.

→ A custom AI solution is an application that's tailor-made to solve a specific problem.

Eg: For instance in 2022 John Hopkins Hospital implemented an AI solution to help doctors detect sepsis, a condition where body has an extreme response to an infection.

AI model - The backbone of AI tools

→ A Computer program trained on a set of data to recognize patterns and perform specific tasks.

→ The model provides the underlying capabalities and AI tools provide you the interface to assist you with the task.

Prompt - A prompt is a text input that provides instructions to the AI model on how to generate output.

- For GenAI it is important to be clean and specific with your prompts.

Where need of human comes in

→ Although AI is useful tools to help help you accomplish tasks it requires human involvement

→ No AI tool has the depth of experience and practical knowledge, and interactive skills that we do. Maintaining human oversight over AI is critical.

An effective strategy of doing so is the human in the loop approach

↳ A combination of machine and human intelligence to train, use, verify, and refine AI models.

The approach uses human insight that is critical for practicing Responsible AI.

↳ The principle of developing and using AI ethically, with the intent of benefiting people and society while avoiding harm.

One aspect of responsible AI involves managing the limitations of AI models, such as knowledge cut-off.

Hallucination

AI output that are not true

The concept that an AI model is trained at a specific point in time, so it doesn't have any knowledge of events or information after that date.

↳ Sometime they turn out to be useful to generate creative prompts.

Eg: Generating image of a flying Castle

Here, we did hallucination intentionally.

How to ones whether to apply generative AI to a task or not:

→ Is the task generative?

→ Can the task be iterated on to acheive the best outcome?

→ Are there resources to provide adequate human oversight?

Prompt Engineering:

Prompt : Text input that provides instructions to the AI ~~outputs~~ model on how to generate output.

Prompt Engineering is the practice of developing effective prompts that elicit useful output from generative AI.

" The more clear and specific are you with your prompts the more likely you are to get useful outputs. "

" Iteration means evaluating the output and revising your prompts also helps to get your results you need. "

" ~~Few shot prompt~~

How LLMs work and their limitations:

Large Language Model (LLM) : An AI model that is trained on large amounts of text to identify patterns between words, concepts, and phrases so that it can generate responses to prompts

An LLM is trained on millions of sources of text, including books, articles, websites and more. This training helps the model learn the patterns and relationships exist in human language.

→ A factor that might affect the quality of answers of an LLM is hallucinations.

- ↦ Quality of an LLM's training data
- ↦ Phrasing of the prompt
- ↦ Method an LLM uses

How to write Prompts?

Prompt Engineering involves designing the best prompt you can to get an output you want from an LLM.

This includes writing clear, specific prompts that provide relevant context. To get a better understanding of the context LLMs need, Lets compare how a human and an LLM might respond to the same question.

~~From a Human~~: Can you find me a restraunt nearby

~~From LLM~~: ~~Can you~~

From Human!

→ There are vegetarian restraunts.

From AI:

→ There are the restraunts (veg & non veg

→ Human friend have the prior knowledge that the person is a vegetarian whereas the AI doesn't.

Prompts for various works:

Summarization : An LLM can summarize a lengthy document's main points

"Summarize the following paragraph in a single sentence"

Classification : For segregating customer reviews as positive negative and neutral.

"Customers reviewed our retail website's new design. Classify the reviews as positive, negative, or neutral."

Extraction : ~~Suppose~~ You have a report that provides information about a global organization.

~~"You can prompt Gemini to extract all mentions of cities"~~

"Extract all mentions of cities and revenue in the following report and place them in a table"

Translation : For translating different languages.

"Our organization is offering a training session. Translate the title for the training session from English to Spanish."

Editing : To change the tone of a paragraph.

"Edit the language of the following paragraph so that it's easy for a non-technical audience to understand"

Problem Solving :

"Help us solve a problem related to our copy editing service • We want to expand to other states • Provide suggestions for increasing our client base • Take into account that we already have marketing campaigns underway."

How to critically evaluate LLM output while using Iterative Approach

→ Is the output accurate ?
→ Is the output unbiased ?
→ Does the output include sufficient information ?
→ Is the output relevant to my project or task ?
→ Is the output consistent if I use the same prompt multiple times ?

Few Shot Prompting :

In prompt engineering, the word "shot" is often used as a synonym for the word "example".

Few shot prompting is technique that provides two or more examples in a prompt.

- Zero Shot Prompting : Used when we need simple answers.
- One Shot Prompting : It is used when you want your LLM to respond in a more specific nuanced way.
- few shot Prompting : When we need unique answers by providing additional context and examples.

Problem Solving :

"Help us solve a problem related to our copy editing service • We want to expand to other states • Provide suggestions for increasing our client base • Take into account that we already have marketing campaigns underway."

How to critically evaluate LLM output while using Iterative Approach

→ Is the output accurate ?
→ Is the output unbiased ?
→ Does the output include sufficient information ?
→ Is the output relevant to my project or task ?
→ Is the output consistent if I use the same prompt multiple times ?

Few Shot Prompting :

In prompt engineering, the word "shot" is often used as a synonym for the word "example".

Few shot prompting is technique that provides two or more examples in a prompt.

- Zero Shot Prompting : Used when we need simple answers.
- One Shot Prompting : It is used when you want your LLM to respond in a more specific nuanced way.
- Few shot Prompting : When we need unique answers by providing additional context and examples.

Use Case

→ Can be used to generate content in a particular style

" Write a one-sentence description of a product. It should contain two adjectives that describe the product. Review the examples and write the description of the skateboard in the same style. "

NOTE : A lot number of examples can also make the answers of the LLMs too flexible and too specific. This is bad when we need creativity.

Responsible AI

The principle of developing and using AI ethically, with the intent of benefiting people and society while avoiding harm

AI models are trained on data created by humans, so they consist of values and are subject to bias.

They can also sometimes produce inaccurate results. Because an AI model is trained on a data set to recognize patterns and perform tasks, the model is only as good as data it receives.

The output from the AI tool may be affected by both systemic bias and data bias.

↓

A tendency upheld by institutions that favours or disadvantages certain outcomes or groups

↓

A circumstance in which ~~sys tes~~ systemic errors or prejudice lead to unfair or inaccurate information, resulting in biased output.

" More the input is biased more the output will also be biased. More is the variety of data, more inclusive the output is. "

AI models are value laden

Human actions to use AI responsibly!

→ AI models require humans to take action, like retraining models on more diverse data sets and continuously to fine tune them frequently

Kind of harms AI can cause if used irresponsibly:

→ Allocative harm: Wrongdoing that occurs when an AI system's use or behavior withholds opportunities, resources, or information in domains that affect a person's well-being.

→ Quality of service harm: A circumstance in which AI tools do not perform as well for certain groups of people based on their identity.

→ Representational harm: An AI tool's reinforcement of the subordination of social groups based on their identities.

→ Social system harm: Macro-level societal effects that amplify existing class, power, or privilege disparities, or cause physical harm, as a result of the development or use of AI tools

→ Deepfakes: AI generated fake photos or videos of real people saying or doing things that they did not do.

Use Case

Interpersonal harm : The use of technology to create a disadvantage to certain people that negatively affects their relationships with others or causes a loss of one's sense of self- and agency.

" Being aware of potential harm and negative outcomes is a first step to using AI responsibly "

Privacy & Security :

Privacy → The right for a user to have control over how their personal information and data are collected, stored, and used

Security → The act of safeguarding personal information and private data, and ensuring that the system is secure by preventing unauthorized access.

Measures to protect privacy and ~~secur~~ security :

→ Be aware of an AI tool's terms of use or service, privacy policy, and any associated risks

→ Don't input personal information or confidential information

→ Stay up-to-date on the latest tools.

How Stay ahead of the AI Curve:

→ To start innovating with AI, consider what tasks you need to accomplish

→ Analyze what slows down your work

→ Think about how incorporating AI into your workflow may help you meet your needs.

Eg: MedPaLM is an Googles AI tool to answers questions related to medical.

PaLM + Training in Clinicians = Med-PaLM

[Existing google LLM]

Multi Modal Model:

→ Some machine learning and AI tools analyze just one modality or form of data in their input.

→ A multi-modal model is an AI model that can accept and learn from multiple types of input, such as images, video, or audio.

→ Refibend developed an AI system that uses sensors to identify and organize different textiles and recycling plants based on their material. This help in reducing the amount of waste in recycling plants.

"Staying knowledgeable about AI as it evolves can benefit you in the workplace in multiple ways."

Use Case

Leverage AI in your work:

→ Examine the tasks you do on a typical day.
→ Analyze your work process as a whole
→ Address challenges in your work process.

Leverage AI in your work:

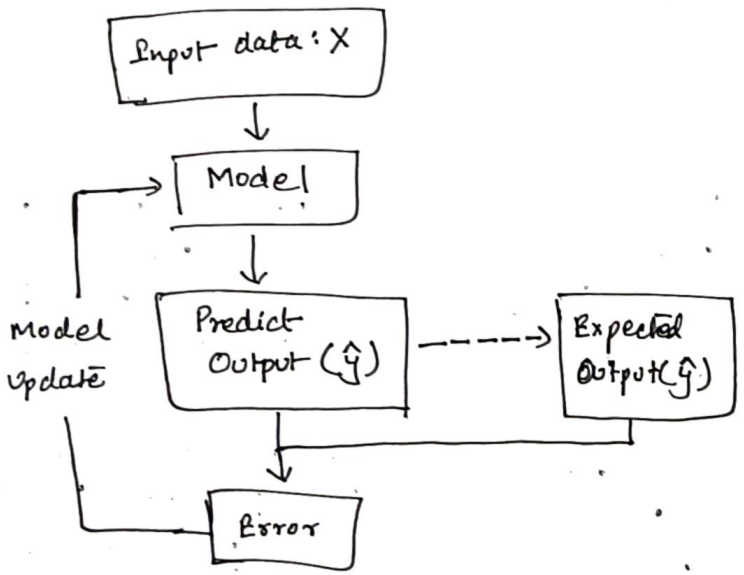
→ Examine the tasks you do on a typical day.
→ Analyze your work process as a whole
→ Address challenges in your work process.

Gen-AI (All need to Know)

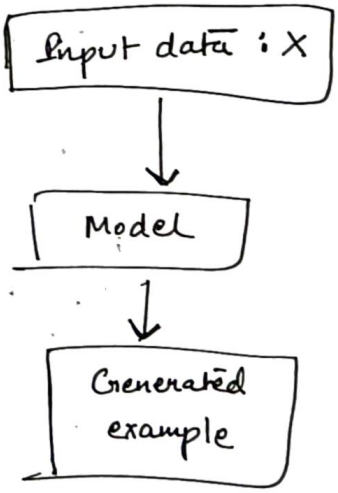
→ Generative AI is a type of artificial Intelligence technology that can produce various types of content including text, imagery, audio and synthetic data.

"AI has to do with the theory and methods to build machines that think and act like humans"
Machine learning is a sub-feild of AI. It is a program or system that trains a model from input data
Machine learning gives the computer ability to learn w/o explicit programming."

Supervised Learning



Unsupervised Learning



Deep learning AND GenAI :

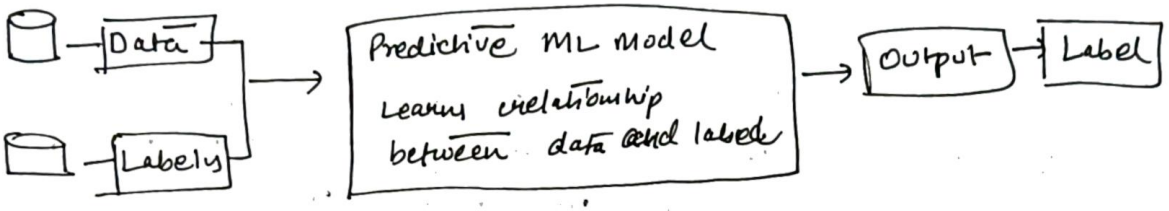
It is a type of machine learning based on artificial neural networks in which layers of processing are used to extract progressively higher level features from data. It is a kind of of machine learning which inspired from human brain.

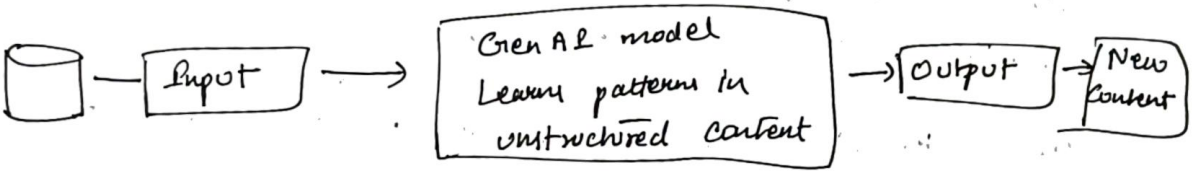
· Generative AI :

GenAI is a subset of deep learning, which means it uses Artificial Neural Networks, can procen both labeled and unlabled data, using supervised, unsupervised, and semi-supervised ~~learning~~ methods

Deep learning model or can be divided into two types
— . generative and discriminative

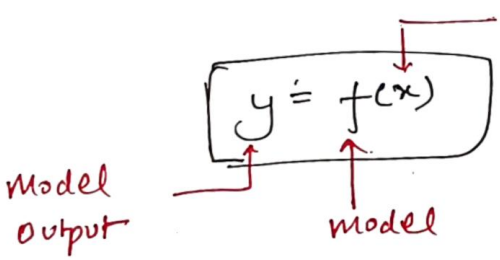
Discriminative model	Generative Model
→ A discriminative model is a type of model that is used to classify or predict labels for a data points.	→ A generative model generates new data instances based on a learned probability distribution of existing data
→ Discriminative model are typically trained on a dataset of labelled data points, and they learn the relationship between the features of data points and labels.	→ The Gen model learns the joint probability distribution (or the probability of x and y) p(x,y) and predicts the conditional probability that this is a dog and can then generate a picture of a dog





Difference btw Predictive ML model & Gen AI model

Not genAI when y is:
- → Number
- → Discrete
- → Class
- → Probability



Is genAI when y is:
- → Natural Language
- → Image
- → Audio

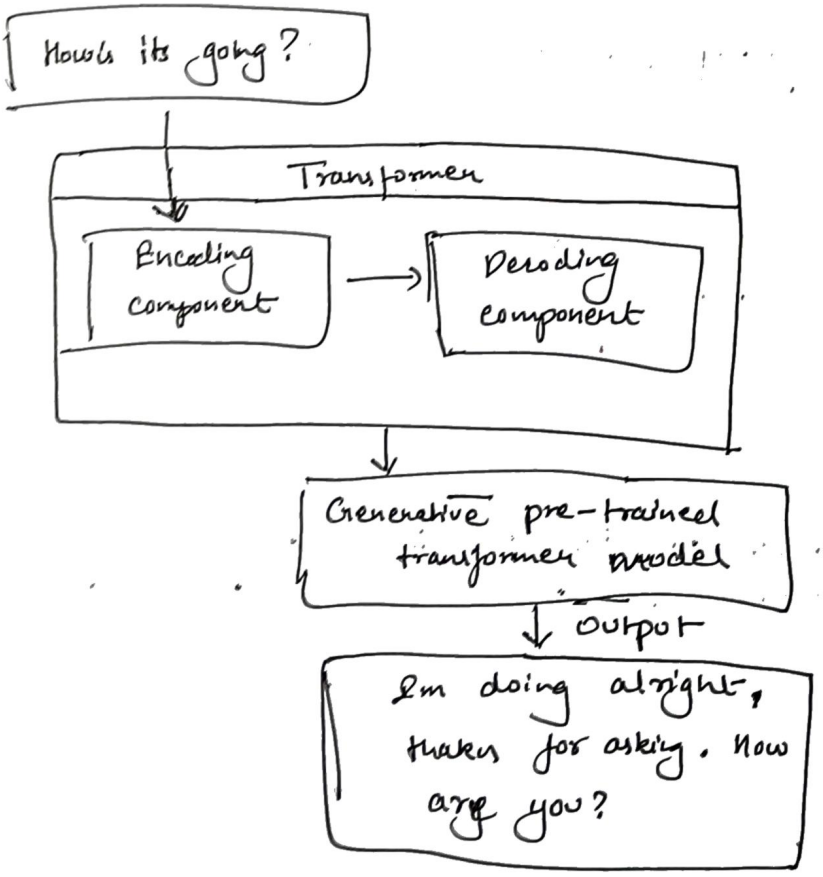
"Gen AI is a type of Artificial Intelligence that creates new content based on what it has learned from existing content!"

Transformers:

The power of GenAI comes from transformers. A transformer is a type of neural network architecture that uses a self-attention mechanism to process and generate sequences of data, like text, by weighing the importance of different words in the input to understand context.

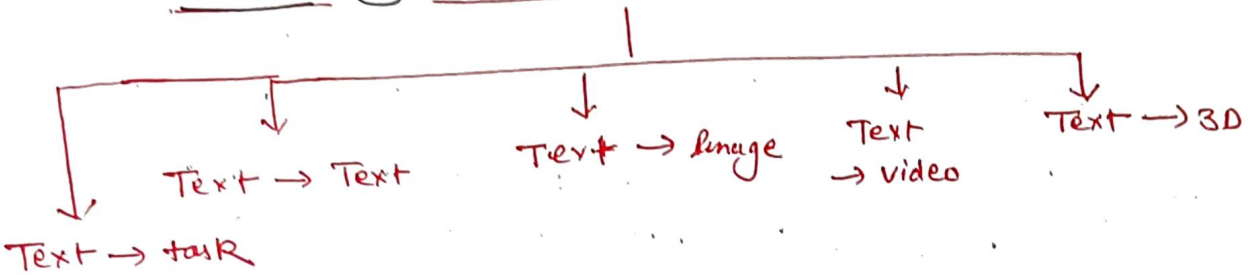
At a high-level, a transformer model consists of an encoder and a decoder

The encoder encodes the input sequence and passes it to the decoder, which learns how to decode the representations for a relevant task.

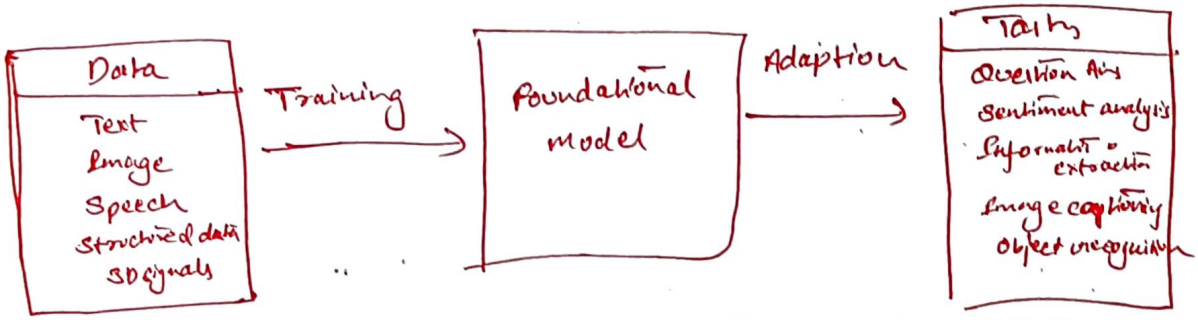


* Prompting is the act of providing a specific instruction or input to an artificial intelligence (AI) system to guide it in generating a desired response.

The model types available to us when text is our input



NOTE: Another model which is ~~mentioned~~ larger than what is mentioned above is a foundational model, which is a large AI model pre-trained on a vast quantity of data designed to be adapted *



Foundational models have the potential to revolutionize many industries, including healthcare, finance, and customer service

Language foundation models : Chat, Text, Code

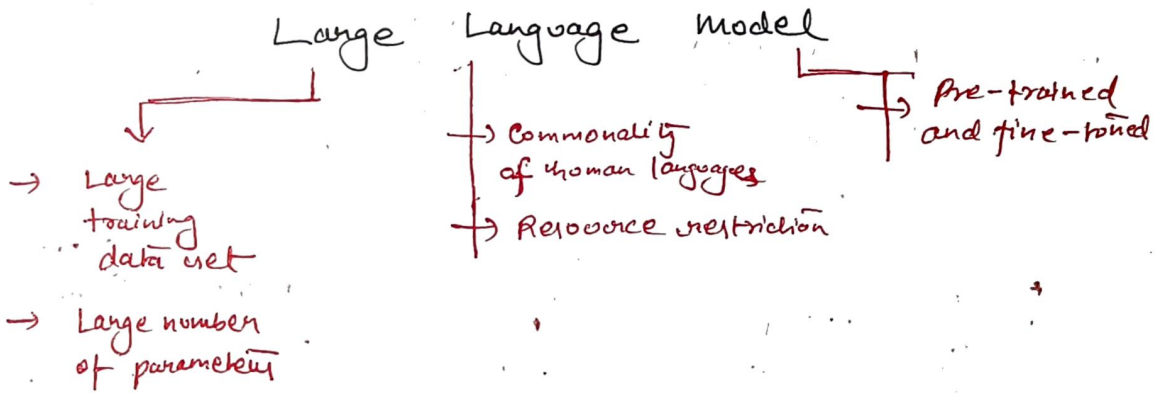
Vision FM : Stable Diffusion v1-5, Embedding extractor

Large Language Models (LLM's)

→ Large Language Model (LLMs) are a subset of Deep learning.

What is it?

→ LLMs refer to the large, general-purpose language models that can be pre-trained and then fine-tuned for specific purposes



LLM use cases: → Language translation
→ Sentence completion
→ Text classification
→ Question Answering

NOTE: What is Question Answering in Natural Language Processing?

Question Answering (QA) is a subfield of NLP that deals with the task of automatically answering questions posed in natural language. QA systems are able to answer a wide range of questions, including factual, definitional, and opinion-based questions.

Prompt design VS Prompt Engineering :

Prompt design is the process of creating a prompt that is tailored to the specific task that the system is being asked to perform

Prompt engineering is the process of creating a prompt that is designed to improve performance

Types of models :

Generic language model : A next word predictor model

Instruction toned Model : Trained on explicit instructions and their corresponding outputs.

Dialog toned models : Trained to have a dialog by predicting the next response.

Dialog tuning is a former specialization of instruction tuning that is expected to be in the context of a longer back-and-forth conversation, and typically works better with natural question-like phrasings.

Efficient methods of tuning :

→ Parameter-efficient tuning methods (PETM)

* Methods for tuning an LLM on your own custom data w/o duplicating the model

→ Prompt tuning

* One of the easiest parameter-efficient tuning methods.