# Gen AI Notes.

- Supervised learning                    Unsupervised learning
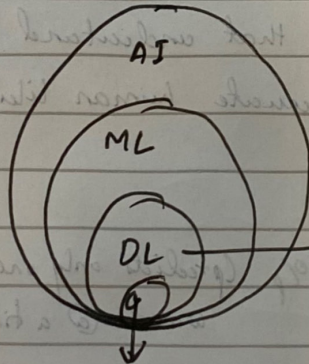  (labeled data)                      (non-labeled data)

Generative model : Generate new data by training on old data

Discriminative model : learn on labeled data and generate/discriminates the data

Machine learning : need lot of training data
                    + computational power
                    + algorithms

Deep Learning

AI
ML
DL → subset of ML that use
                  neural n/ws to solve the problems

Gen AI : subset of deep learning

Conventional AI

i/p data → AI → prediction     (eg) Lots            ✓
                classification     of images → AI
                clustering        of apple          ✗
                NLP
                computer vision

Gen AI

i/p data → AI → Generate     (eg):
                text           lot    → AI → generate
                image      of apple         apple image.
                video
                audio
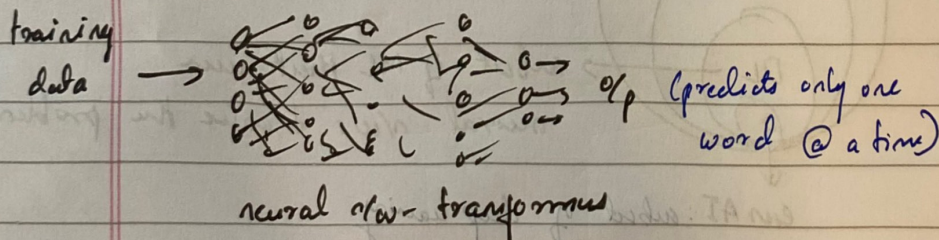                source code

- Machines need to be trained on ① large amount of data
  - ② need high computational power
  - ③ shift to conversational, contextual understanding

## Chat GPT: ① large language model (LLM)
  ② designed for natural language understanding & generation
  ③ can handle conversational context

- A GenAI application
- Developed by Open AI
- trained on billion of documents
- Based on Generative Pre-trained Transformer (GPT) architecture, a type of neural n/w.

## Large Language Models (LLMs): AI models that understand & generate human like text



training
data ⟶ [neural network diagram] o/p (predicts only one word @ a time)

neural n/w - transformer

- pre-training: on large corpus of data
- size & scale: massive neural n/ws (huge # of parameters)
- fine tuning: more 'targeted' training for specific tasks.

## Use cases: content generation
  chatbot & virtual assistant
  language translation
  text summarization
  Q & A

ask unambiguous
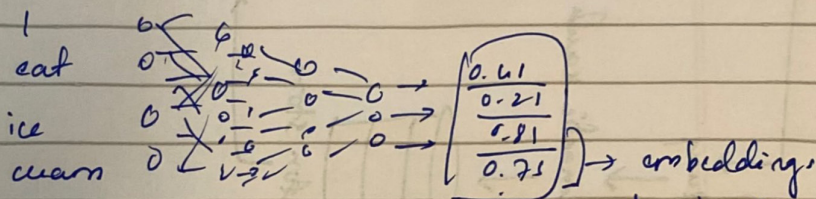clear questions

## Prompt Engineering

Prompts ex: ① Siri → Will it rain today?

Prompt → specific question / command / input that you provide to AI system to request particular response, information / action.

**Embeddings** : numerical representation of text

Eg: I eat icecream
↓
break to small chunks
↓

eat

ice

cream



$$\begin{bmatrix} 0.41 \\ 0.21 \\ .81 \\ 0.75 \end{bmatrix}$$ → embeddings.

for ice & cream au very similar

∴ ice-cream go together.

**Fine tuning**: adapting a pre-trained model for specific tasks. / dataset.

**supervised fine-tuning**

input / output → model

**reinforcement learning**

LLMs → i drink icm  0 ✗
i eat icm  10 ✓

fine tuning is not about ① eliminating data
② not 1-time process

**Gen AI Use Cases Across Industries**  ① S/w development ② Retail ③ Marketing

knowledge
repository/
source

[Pdf] [Pdf]

[Pdf]

chunk

open AI

embeddings

$\left\{\begin{array}{l}0.62\\4525\\9243\end{array}\right\}$
$\left\{\begin{array}{l}0.17\\665\\081\\623\end{array}\right\}$
$\left\{\begin{array}{l}0.2273\\3641\\9645\\2263\end{array}\right\}$

query ↓
semantic search ↓

vector store → ranked result

(open AI) ↓
query ↓

LLM ↑

result ↑ → user