

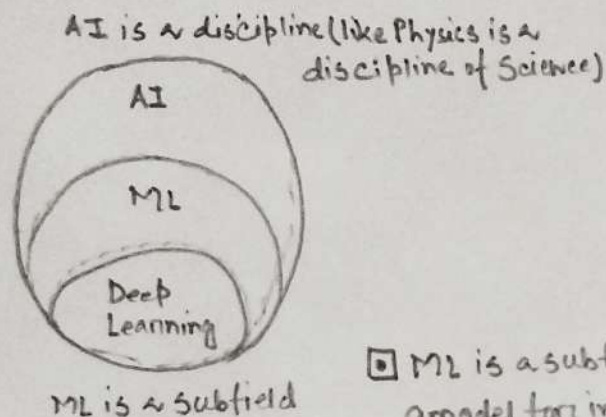
The course covers 4 things:

- Define generative AI
- Explain how generative AI works
- Describe generative AI model types
- Describe generative AI applications

What is generative AI?

- Generative AI is a kind of artificial intelligence technology that produces different kind of content including text, imagery, audio and synthetic data

Before going to the details of generative AI, let's find out the difference b/w 'Artificial Intelligence' and 'Machine Learning.'



- Artificial Intelligence is a branch of Computer Science that deals with the creation of intelligent agents, and the system that can reason, learn and act autonomously.
- AI has to do with theory and methods to build machines that think and act like humans.
- ML is a subfield of AI. ML is a program or system that trains a model for input data.

- The trained model can do some useful predictions on new/never before seen data drawn from the same one used to train the model.
- ML gives the computer the ability to learn without explicit programming.
- ML models can primarily be classified into 2: Supervised (labeled i/p data) and Unsupervised (unlabeled i/p data)

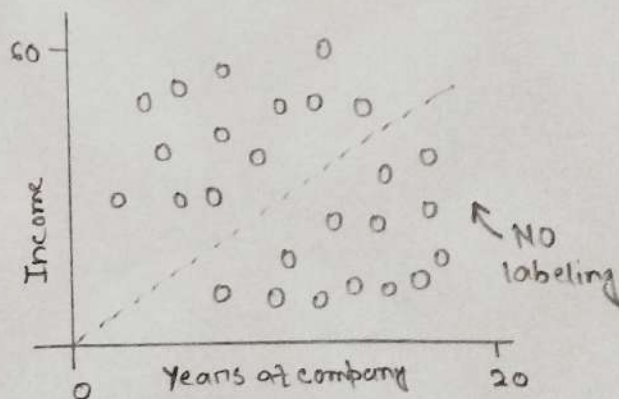
Supervised Model:
Restaurant tips by Order type



○ → pickup
X → delivery

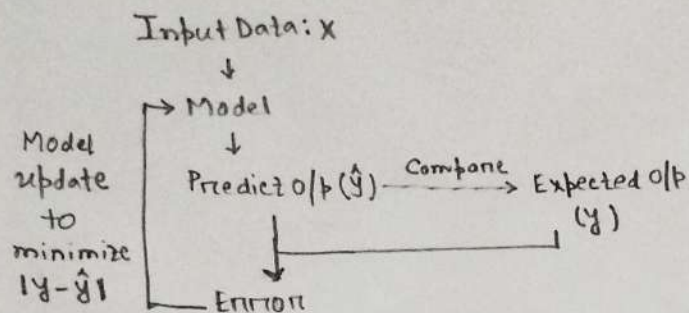
Here, supervised model can predict the tip amount for an unknown bill amount for order types b/w pickup and delivery

Unsupervised Model: Income vs. job tenure



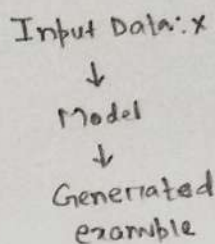
In unsupervised models, we need to look at the raw data and try to find out naturally the data can be categorized into groups.

Supervised Learning



Classic example of optimization problem

Unsupervised Learning

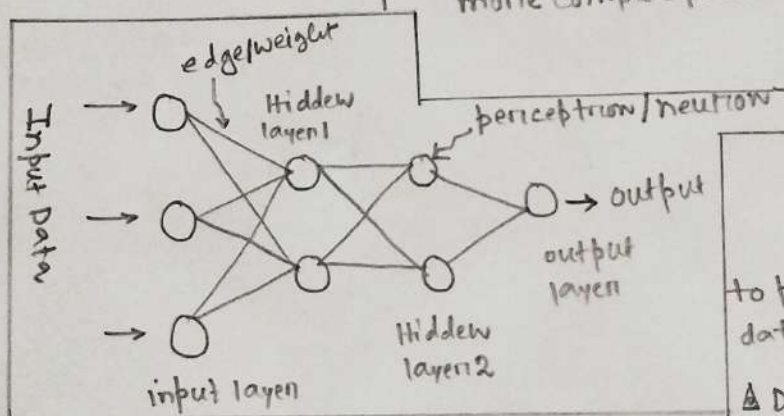


Machine Learning

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning
- Deep Learning

Post-discussion a bit about a subfield of Machine Learning, i.e., Deep Learning, we will start with the discussion of Generative AI.

△ Deep Learning is a type of Machine Learning that uses artificial neural networks, allowing to process more complex patterns in Machine Learning.



△ Artificial Neural Network (ANN) is like human brain. Like our brain, ANN is made up of many interconnected nodes or neurons that learn to perform tasks by processing data and making predictions.

△ Deep Learning models typically have many layers of neurons which allows them to learn more complex patterns as compare to traditional Machine Learning models.

△ Neural Network (NN) can take both labeled and unlabeled data; this is called semi-supervised learning. Typically, in semi-supervised learning, the neural network is trained with small # of labeled data and large # of unlabeled data.

△ Labeled data helps NN to learn basic concepts of tasks, whereas, unlabeled data helps NN to generalize to new examples.

Generative AI (GenAI)

GenAI is a subfield of Deep Learning. GenAI uses artificial neural network (ANN), which can process both labeled and unlabeled data using Supervised, Unsupervised and Semi-supervised methods. Large Language Models (LLMs) are also subset of Deep Learning models.

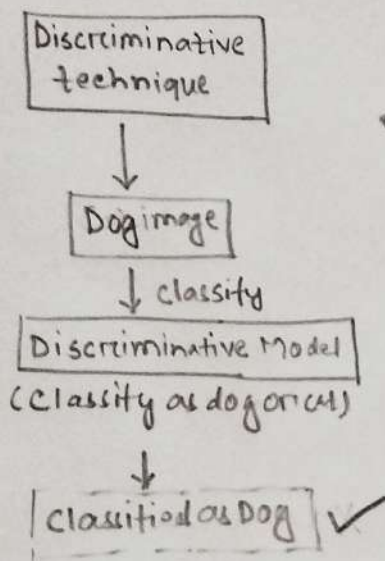
Deep Learning models or Machine Learning models in general can broadly be divided into 2 types: 1. Generative model; 2. Discriminative model

Discriminative Models

- Type of models that use classification to predict labels for data points
- Typically trained on labeled dataset and learn the relationships between features of data points and labels
- Once the model is trained, it can be used to predict the labels of new data points.

Generative Models

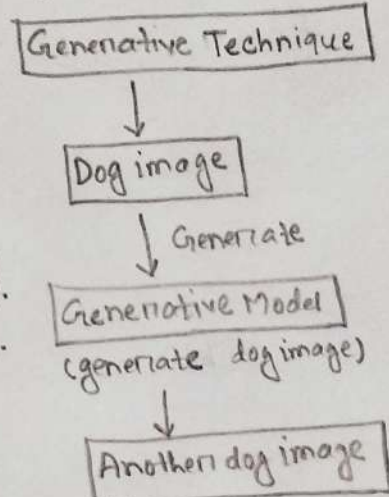
- generate new contents
- generate new data instances based on learned probability distribution of existing data.



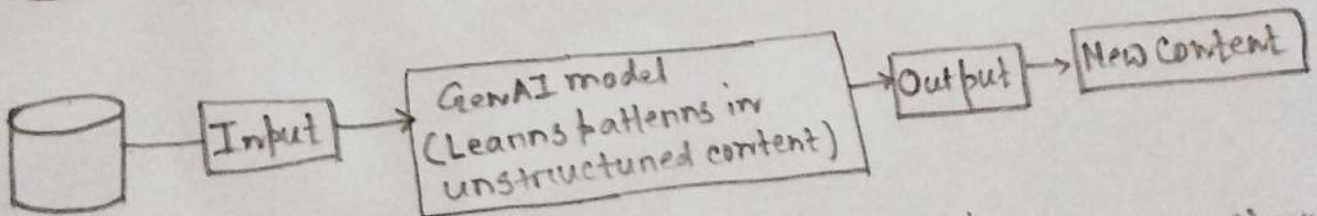
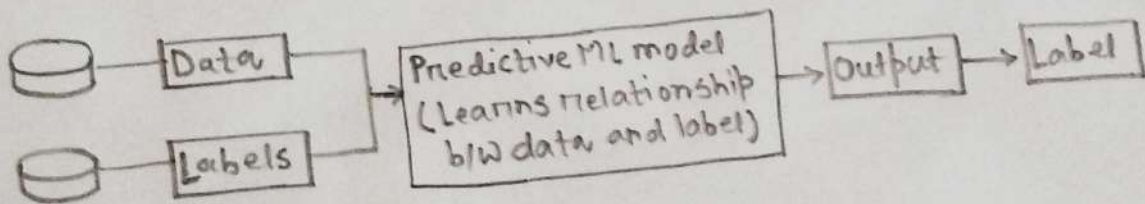
Discriminative model learns the conditional probability distribution or probability of $Y(o/p)$ given $X(i/p)$. Here, the discriminative model classified a dog image as dog.

Summarization

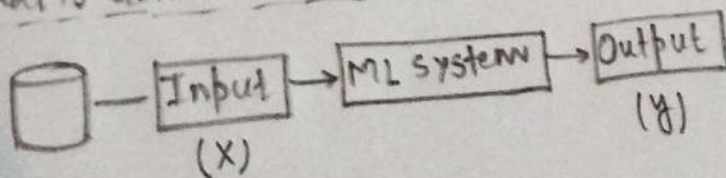
- Generative models generate new data instances.
- Discriminative models discriminate b/w diff. kinds of data instances.



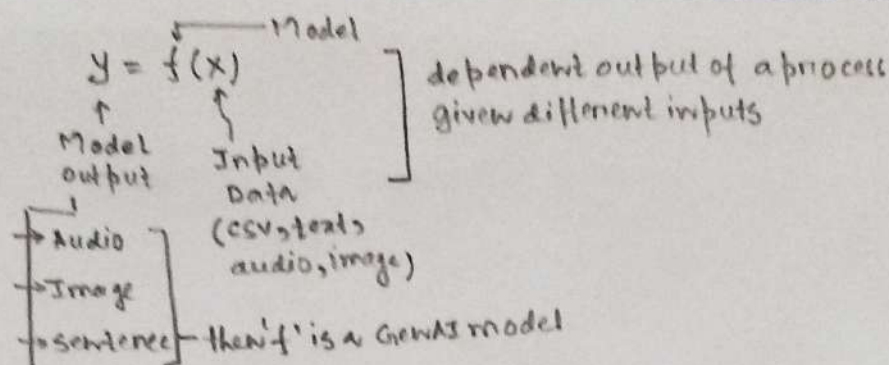
Generative model learns the joint probability distribution, i.e., $p(X, Y)$, and predicts the conditional probability that it is a dog, and can generate a picture of a dog.



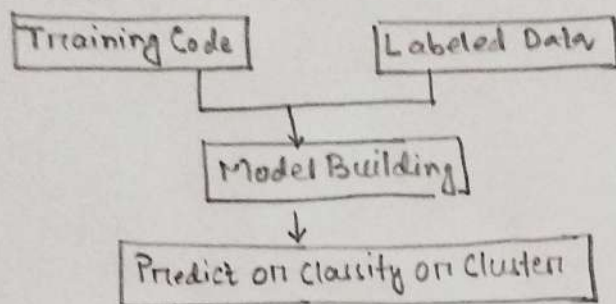
Unstructured content
What is GenAI and what is NOT?



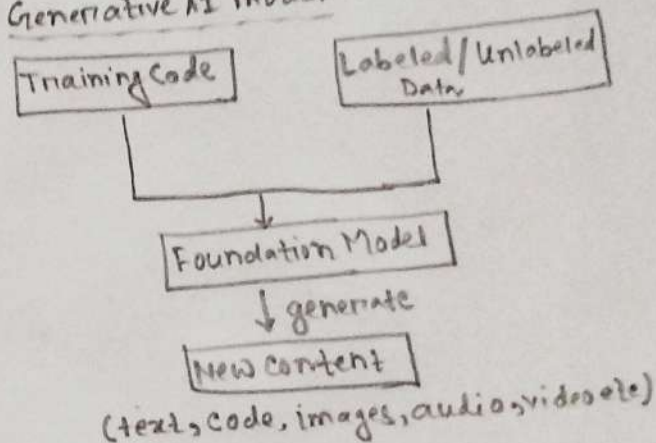
- | | |
|-----------------------|----------------------|
| NOT GenAI, when y is: | IS GenAI, when y is: |
| + Number | + Natural language |
| + Discrete | + Image |
| + class | + Audio |
| + Probability | |



Traditional ML model



Generative AI model



Traditional ML Model → Neural Network
GenAI ←

Generative Language Models: LAMBDA, PaLM, GPT, etc.

Language Model for Dialogue Applications

Pathways Language Model

Generative Pre-trained Transformers

Generative language models ingest very very large data from multiple sources across internet and build Foundation Language Model, which gives output by just typing in prompt or asking in prompt.

→ If we write what is cat? - It gives output about all the details that it learned about cat from multiple sources.

Formal definition of GenAI

GenAI is a type of Artificial Intelligence that creates new content based on what it has learned from existing content.

→ Process of learning from existing content is called Training and it results in the creation of Statistical Model.

→ When given a prompt, GenAI uses the statistical model to predict what an expected response might be and this generates new content

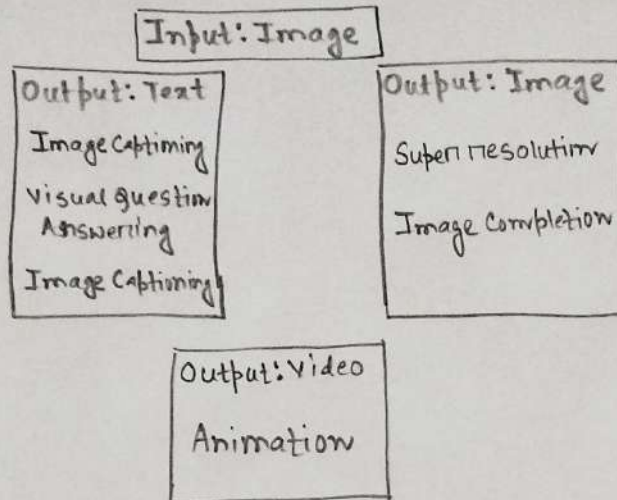
→ GenAI learns the underlying structure of the data and can then generate new samples that are similar to the data it was trained on.

Generative Language Models

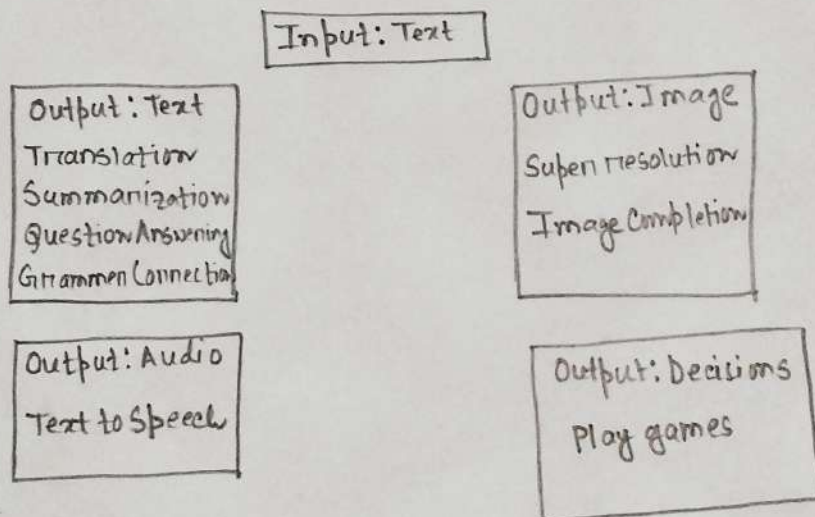
Generative language models learn about patterns in language through training data. Then given some text, it predicts what comes next.

Generative Image Models

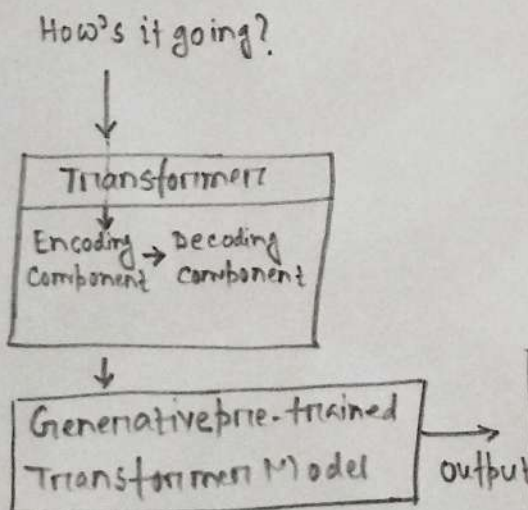
Generative image models produce new images using techniques like diffusion. Then given a prompt on related imagery, they transform random noise into images or generate images from photos. Here, the o/p can be another image, text or video.



Generative Language Models



Generative AI and Transformers



Sometimes, Transformers run into issues through hallucinations (words or phrases that are generated by the model that are often non-sensical or grammatically incorrect)

Hallucinations can be caused by no. of factors like

- Model is NOT trained on enough data
- Model is trained on noisy or dirty data
- Doesn't give enough contexts
- Doesn't give enough constraint

I'm doing alright, thanks for asking. How are you?

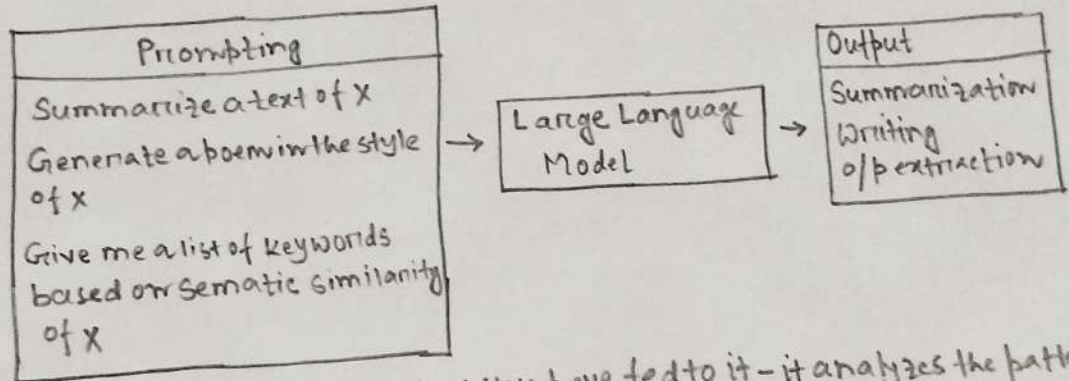
Hallucinations in Transformers

- Can make the output text difficult to understand
- Can make the model more likely to generate incorrect or misleading information
- Simply, hallucinations are bad for Transformers

Prompt

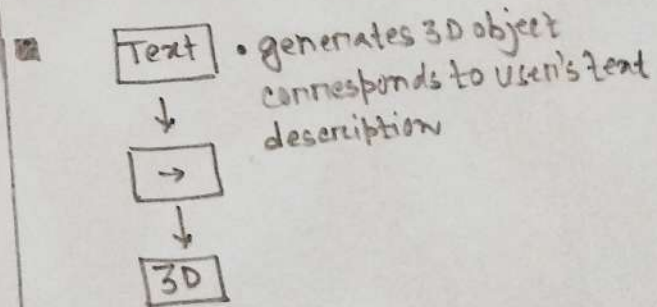
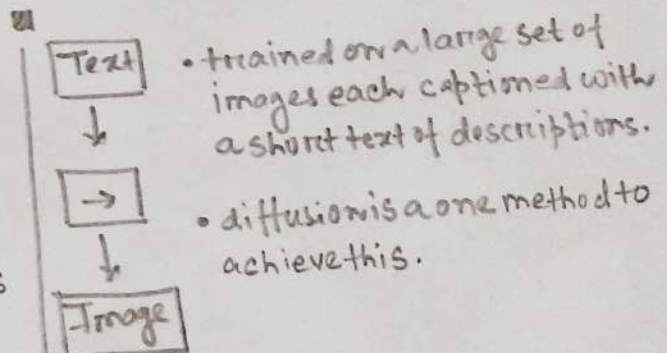
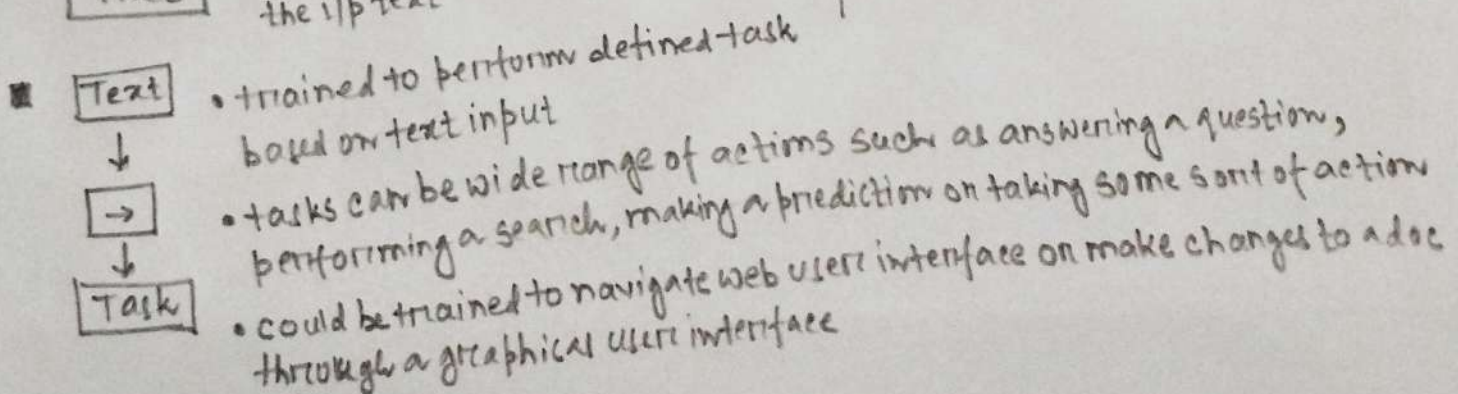
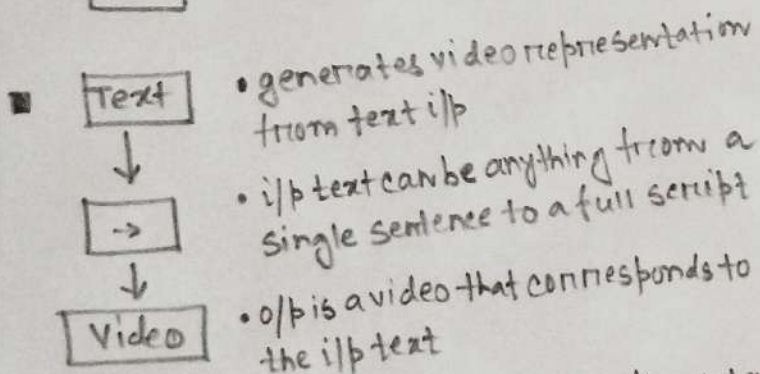
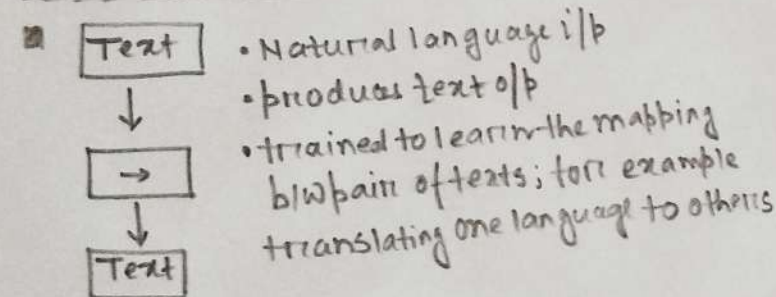
Prompt is a short piece of text that is given to a large language models or LLMs as input and it can be used to control the output of the model in a variety of ways.

↳ Prompt design is a process of creating a prompt that will generate a defined output from an LLM.

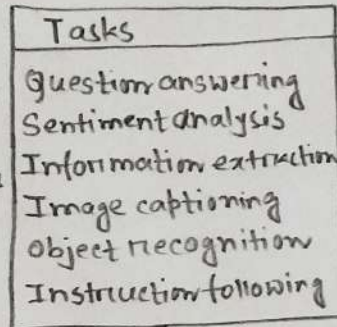
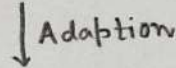
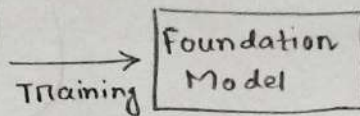
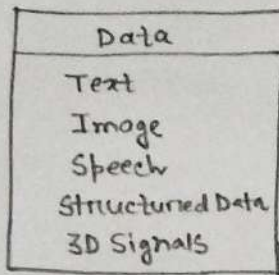


GenAI depends a lot on the training data that you have fed to it - it analyzes the patterns and structures of i/p data thus learns. But with access to a browser based prompts, the users can generate their own content.

GenAI Model Types

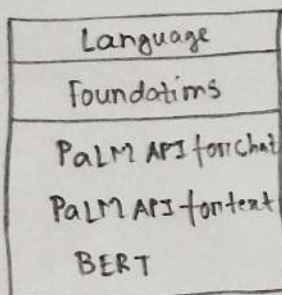


Foundation Model



- Vertex AI offers a model garden that includes Foundation models

- Language Foundation Model includes



BERT → Bidirectional Encoder Representations from Transformers

GenAI Use Cases

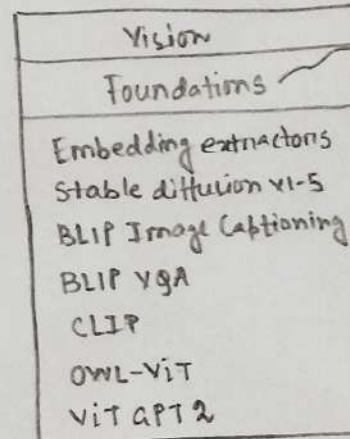
- Need to gather sentiments about how your customers feel about your product or service
- Occupancy analytics

- Large AI model pre-trained on a vast quantity of data designed to be adapted or fine-tuned to a wide range of Downstream tasks

- Foundation model has the potential to revolutionize many industries including healthcare, finance, customer service

- can be used for fraud detection, personalized customer support

- Vision Foundation Model includes



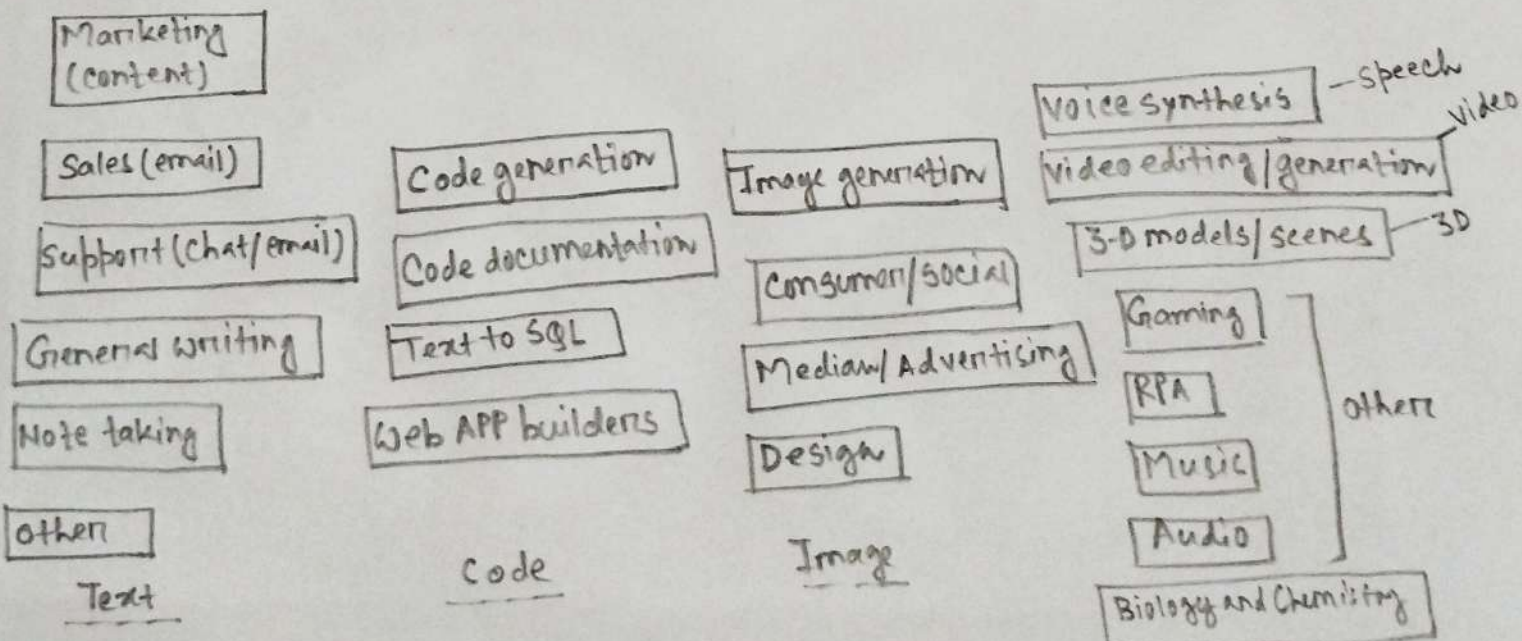
generating high quality images from text descriptions

BLIP → Bootstrapping Language-Image Pre-training
VQA → Visual Question and Answering

CLIP → Contrastive

Language-Image Pretraining
ViT → Visual Transformers

Generative AI Applications Landscape



Gemini Code Generation Example

Prompt: I have a Pandas DataFrame with two columns - one with the filename and one with the hour in which it is generated:

File	Hour
F1	1
F2	2
F3	1
F4	1

I am trying to convert to a JSON file with the following format:

```
{ "File": "F1", "Hour": 1 }
{ "File": "F2", "Hour": 2 }
{ "File": "F3", "Hour": 1 }
{ "File": "F4", "Hour": 1 }
```

How can I do this?

Output: Sure, here are the steps on how to convert a Pandas DataFrame to a JSON file with desired format:

1. Import the Pandas library
2. Create the Pandas DataFrame with following data:

File	Hour
F1	1
F2	2
F3	1
F4	1

```
[
  {
    "File": "F1",
    "Hour": 1
  },
  {
    "File": "F2",
    "Hour": 2
  },
  ...
]
```

Gemini code generation can help

- code debugging
- code explanation line-by-line
- crafting SQL queries for database
- Translating code from one language to another
- Generation of documentation and tutorials for source code

Applications of Generative AI on Google Cloud

- quickly explore and customize GenAI models
- Developers create and deploy GenAI models
- Library of pre-trained models
- Tool for fine-tuning models
- Tools for deploying models to production
- community forum for developers to share ideas and collaborate

Vertex AI

Vertex AI would also be helpful for those who don't have much of coding experiences:

- can build generative AI search and conversations for customers and employees with Vertex AI Search and conversation formerly GenAI APP builder
- creations of chatbots, digital assistants, custom search engines, knowledge bases, training applications, and more

Vertex AI has PALM API:

- leveraging test and experiment with Google's Large Language Models and GenAI tools to make prototyping quick and more accessible
- developers can integrate PALM API with maker suite and use it to access API using GUI
- Maker suite contains different tools such as model training tools, model deployment tool and model monitoring tool
 - Model Training Tool: Helps developers to train ML models on their data using different algorithms
 - Model Deployment Tool: Helps developers to deploy ML models to production with a number of different deployment options
 - Model Monitoring Tool: Helps developers to monitor the performance of their ML models in production using a dashboard and no. of different metrics

In summary, Gemini is a multi-modal AI model unlike the traditional large language models, it's NOT limited to understanding the text alone, it can analyze images, understand the nuances of audio and even interpret programming code.

↳ This allows Gemini to perform complex tasks that were previously impossible for AI. Due to Gemini's advanced architecture, it is incredibly scalable and adaptable, making it suitable for diverse applications.

↳ Model Garden is continuously updated to include new models.