AI, ML, Deep Learning & Gen AI

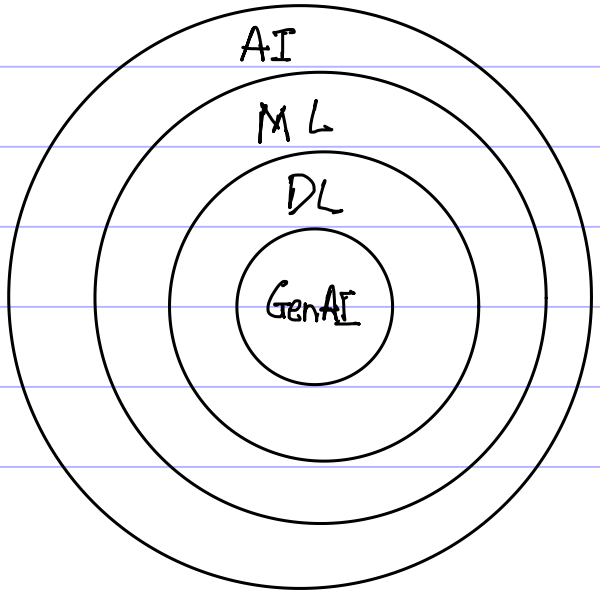
What is AI?

→ AI is a broad field for the development of intelligent systems capable of performing tasks that typically require human intelligence. for example :-

1. Perception
2. Reasoning
3. Learning
4. Problem Solving
5. Decision Making

→ AI, overall is an umbrella term for all the various techniques within the field of AI.

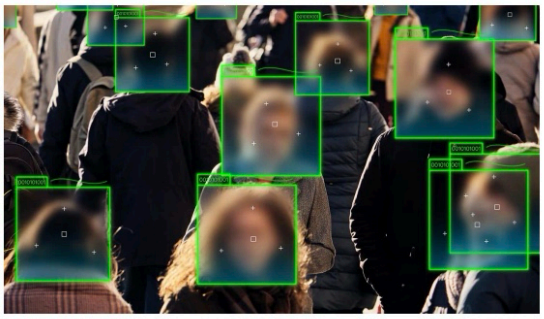
→ AI ⊆ ML ⊆ DL ⊆ Gen AI



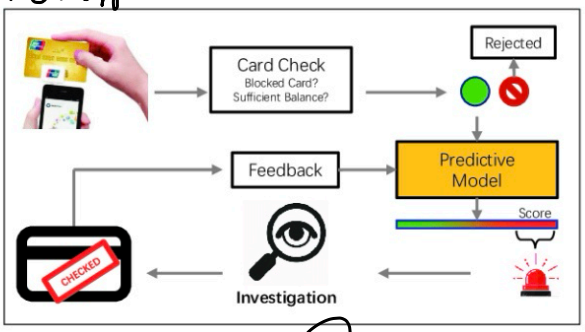
Use Cases of Artificial Intelligence

* Computer Vision for self driving Cars

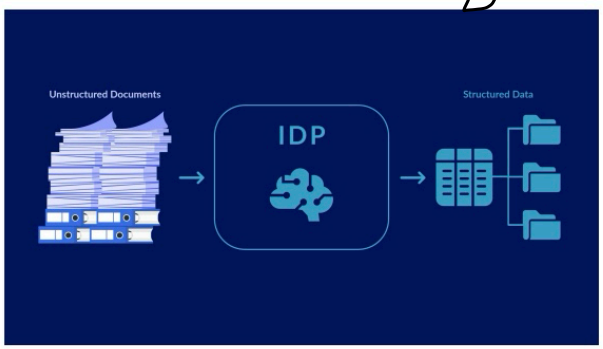
* Facial Recognition



* Fraud detection



* Intelligence document Processing (IDP)



How does AI work? (AI Components)

* We have a data layer, where we are going to collect vast amount of data. This is typically handled by data scientist or data engineers

* → Then we are going to define a machine learning framework or algorithm layer.
→ This is where your data Scientist and your engineers work together to understand the use cases, the requirements, and the frameworks that can help solve your problems.

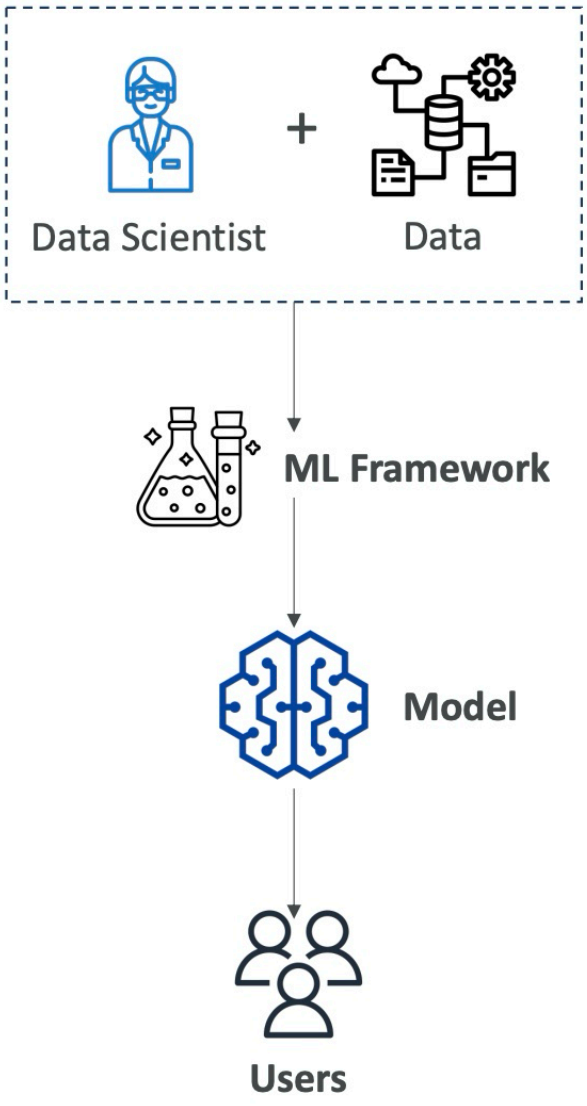
* Model Layer

↳ This is where we implement a model, and we train it.
↳ So we have the structure, we set the correct parameters and function.
↳ We set an optimizer function to create actual model.

* Application Layer

↳ This is the layer where we serve the model to the users.

↳ Then we expose the model's capabilities in a specific way for end-user interaction



Data layer

Model layer

Application layer

What is Machine Learning?

* Machine Learning is a type of AI used to build methods that allow machines to learn from data.

Key Characteristics

* The Data is going to be leveraged because when we have a lot of data, we are going to improve computer performance on a specific set of tasks.

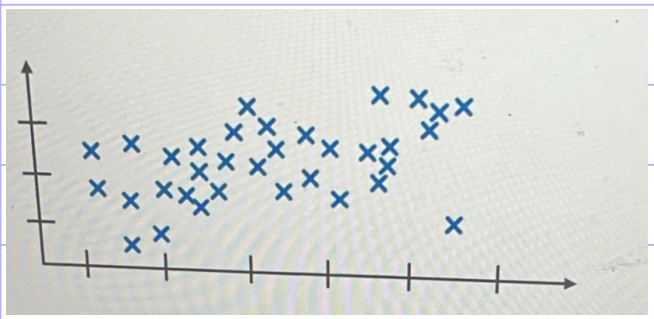
Data Driven improvement

Prediction capability

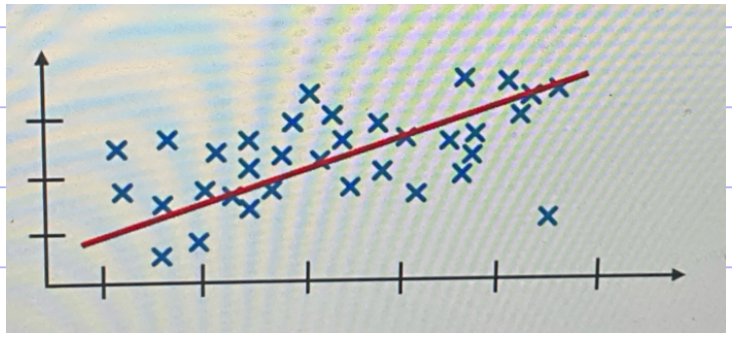
* We can make predictions based on the data used to train the model

Example :- ① Regression

* We have 2 axises, & we have lot of data points.
* We want to sort of predict where the data points are going to be.



* So in the regression, we would have this line right here



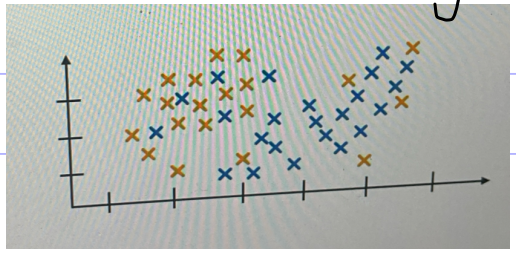
* We would say "Look, it looks like this data set is following this trend of this line"

* Basically in regression, the line is created that best fits the data pattern.

* This will enable predictions based on the identified trend line.

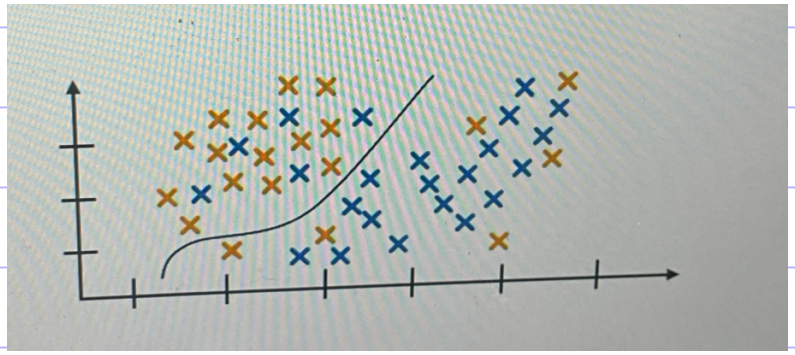
② Classification

* The other one is classification. We say that :- "Here are my all data points"



* Some of them are Blue and some are orange.

* If I draw a specific line, then left of my line is going to be mostly orange points, and the right of my line is going to be mostly blue points



In general, what classification is :-

* Analyzes data points with different categories (eg: blue and orange points)

* Creates decision boundaries to separate categories

* Establishes rules to classify new data automatically.

Note :-

⇒ In ML, we don't explicitly program these rules. We just give data to the algorithm

AI != ML Example

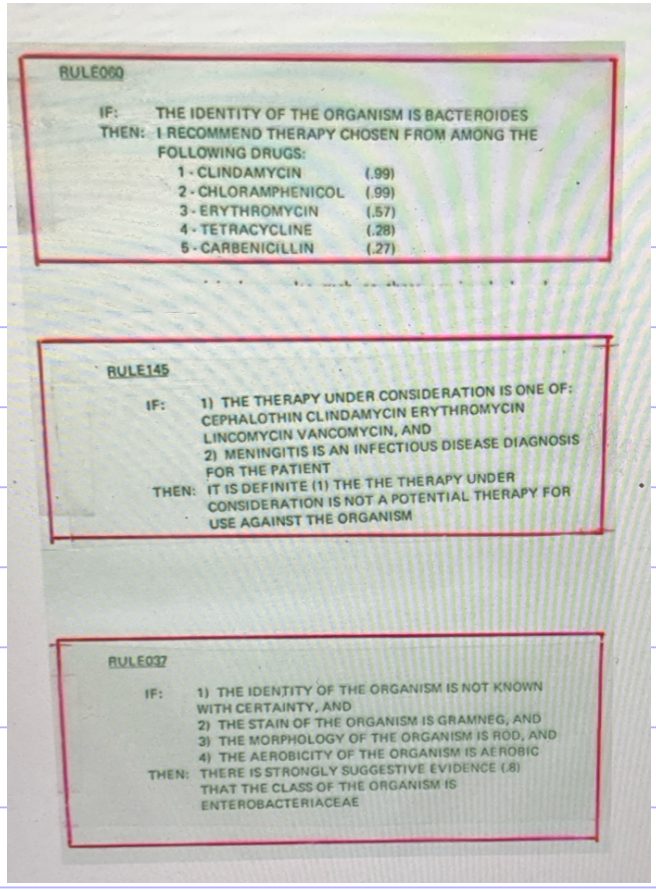
* It is important to remember that AI is not equal to machine learning, although nowadays new AI is machine learning. However, back in 1970s, there are different types of AI systems developed that worked differently from today's approaches.

MYCIN Expert System (1970s)

The MYCIN system was developed in the 1970s to diagnose patients based on their reported symptoms and medical results. This system represents an entirely different approach to AI than what we use today

How MYCIN worked

→ It was created by group of people who programmed over 500 explicit rules.



RULE060

IF: THE IDENTITY OF THE ORGANISM IS BACTEROIDES
THEN: I RECOMMEND THERAPY CHOSEN FROM AMONG THE FOLLOWING DRUGS:
1 - CLINDAMYCIN (.99)
2 - CHLORAMPHENICOL (.99)
3 - ERYTHROMYCIN (.57)
4 - TETRACYCLINE (.28)
5 - CARBENICILLIN (.27)

RULE145

IF: 1) THE THERAPY UNDER CONSIDERATION IS ONE OF: CEPHALOTHIN CLINDAMYCIN ERYTHROMYCIN LINCOMYCIN VANCOMYCIN, AND
2) MENINGITIS IS AN INFECTIOUS DISEASE DIAGNOSIS FOR THE PATIENT
THEN: IT IS DEFINITE (1) THE THE THERAPY UNDER CONSIDERATION IS NOT A POTENTIAL THERAPY FOR USE AGAINST THE ORGANISM

RULE037

IF: 1) THE IDENTITY OF THE ORGANISM IS NOT KNOWN WITH CERTAINTY, AND
2) THE STAIN OF THE ORGANISM IS GRAMNEG, AND
3) THE MORPHOLOGY OF THE ORGANISM IS ROD, AND
4) THE AEROBICITY OF THE ORGANISM IS AEROBIC
THEN: THERE IS STRONGLY SUGGESTIVE EVIDENCE (.8) THAT THE CLASS OF THE ORGANISM IS ENTEROBACTERIACEAE

→ It had very specific rules with Yes/No or textual questions

→ All the rules were explicitly programmed by humans.

System Process

1. Patient would answer a series of questions based on the rule system
2. The system would automatically generate a list of potential bacteria that could cause the disease
3. It provided probability of diagnosis for each potential cause.

4. It also included reasoning behind the diagnosis
5. Then it suggested potential dosage for treatment.

Why MYCIN was never used in Production

* Despite being technically functional, the MYCIN system was never really used in production for practical reasons:-

→ Limited implementation options: There was no easy way to implement such systems in the 1970s

→ Computer Accessibility: Personal computers were not cheap and accessible as they are today.

→ Hardware limitations: The technology infrastructure needed to support such system did not exist

Then: Rule-Based Programming

- Humans explicitly programmed every rule
- Required extensive manual coding of logic
- Based on expert knowledge translated into IF-THEN statements

Now: Machine Learning Approach

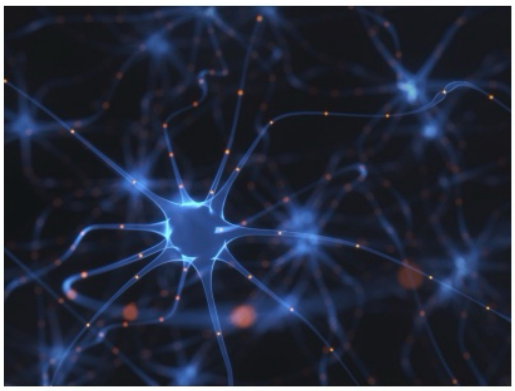
- We don't program explicit rules anymore
- Instead, we throw large amounts of data at machine learning algorithms
- The algorithms automatically generate models and discover patterns
- The system learns relationships and rules from data rather than being explicitly programmed

This represents a fundamental shift from human-programmed logic to data-driven pattern recognition in AI development.

What is Deep Learning?

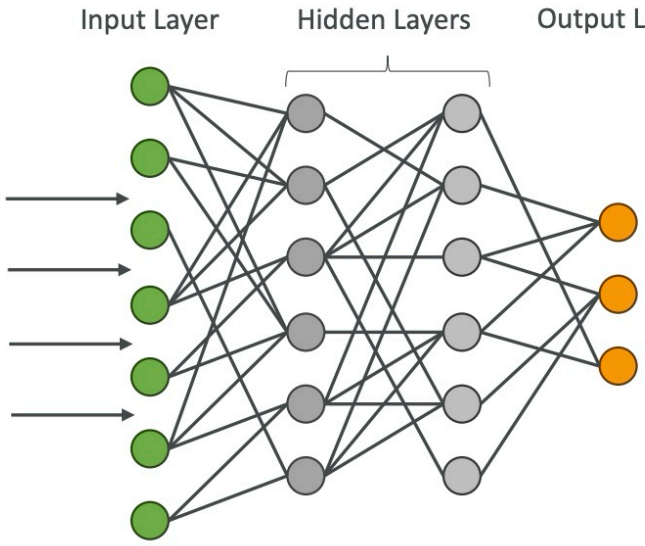
* Deep Learning is a subset of Machine Learning. DL uses the concept of neurons and synapses like our brain to train a model.

* We were really inspired by how our brain functions to develop these models.



* With Deep learning :-
  a. You are able to process more complex patterns in the data than with the traditional machine learning techniques like regression and classification

* So deep learning is following what a brain looks like. We have :-
  1. Input layer :- we give a lot of data here
  2. Hidden layer :- more than one hidden layer are there
  3. Output layer :- this is where we get the answer that we are looking for



* It is called deep learning because there is more than one layer of learning.

* Examples :-

① Computer Vision such as
- Image Classification
- Object detection
- Image Segmentation

② Natural Language Processing (NLP) such as
- Text Classification
- Sentiment Analysis
- Machine translation
- Language Generation

* With DL, to have a very good model, you need to have a very large amount of input data.

* On top of it, it is computationally heavy.

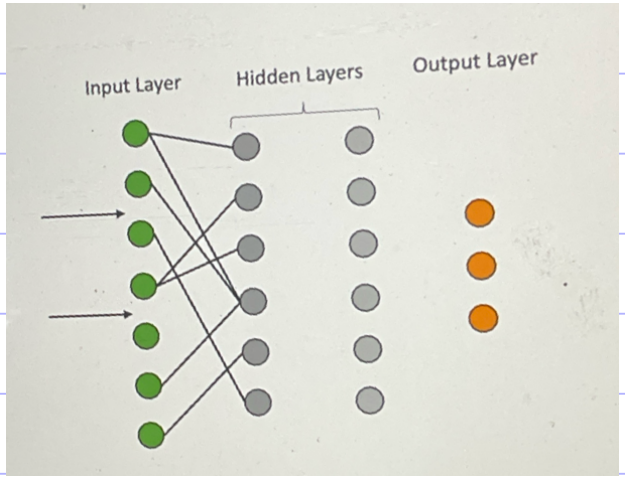
meaning ↳ it requires GPUs (Graphical Processing Units)

- a gpu is something which your computer has to actually display
- It is also used in parallel computations
  ↳ DL use case.

Neural Networks : How do they Work?

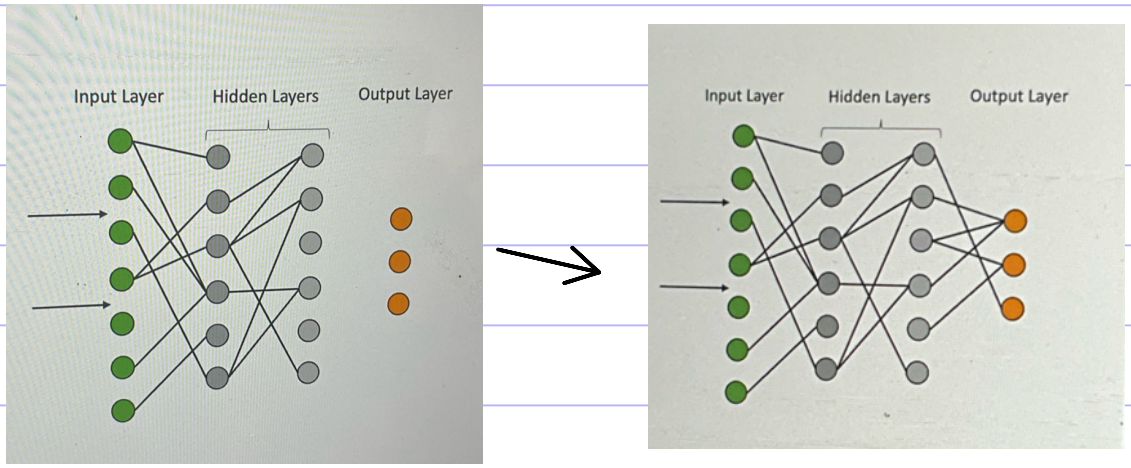
(will not be asked in detail for the exam)

* We have the input data, which will be put into our network :-



* By putting the data, it is going to create connections between our different layers. (see the pic above)

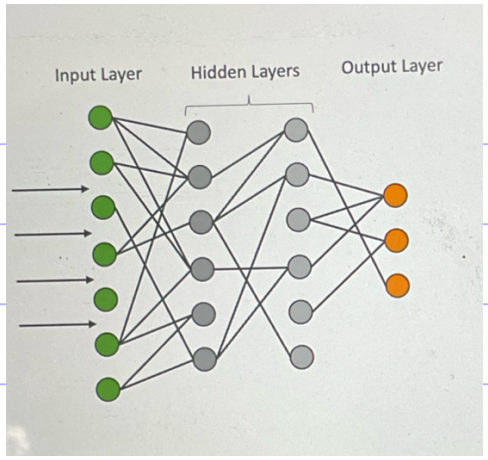
* Over time, new connections will be created, all the way down to the output layer



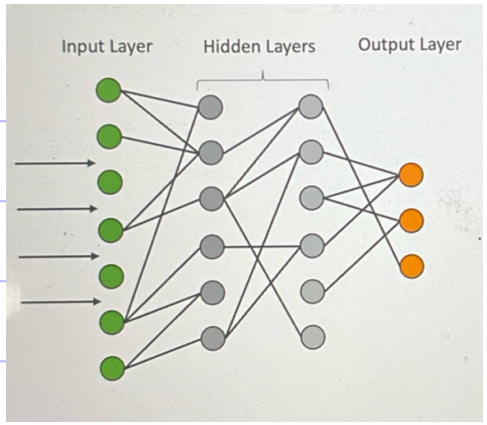
* These nodes (tiny units) are connected together and organized into layers
  - Input layer (1)
  - Hidden layer (many)
  - Output layer (1)

* When the neural network sees a lot of data, it identifies patterns and changes the connections between the nodes.

* Nodes are "talking" to each other, by passing on (or not) data to the next layer (creating new connections or removed)

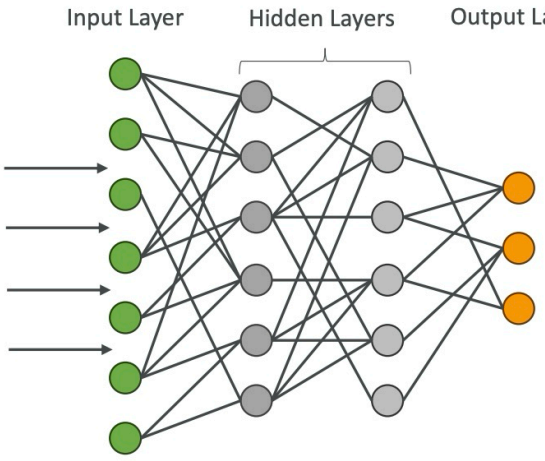


OR



Creating new connections

Removing old connections



Removing old & adding more connections (It totally depends upon the data added)

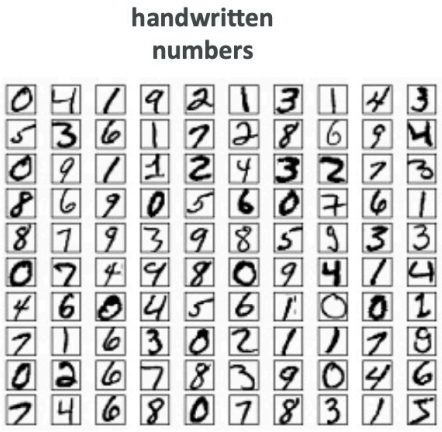
Note!-

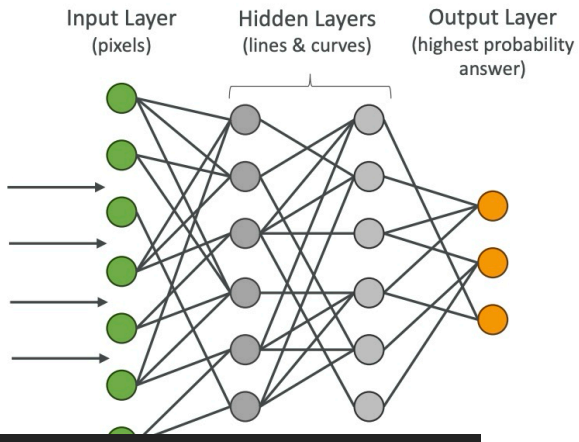
* The math and all the parameters behind tuning a neural network is way beyond the level of this course.

* In practice, neural networks have billions of nodes and many, many different layers. That's why it's called deep learning.

Deep Learning Example: Recognizing hand-written Digits

* Here is a simple example of how deep learning works with handwritten number recognition:





How the System Works:

Input Layer (Pixels):

- The input layer represents the pixels of handwritten numbers

Hidden Layers (Lines & Curves):

- Hidden layers represent what the model identifies as lines or curves based on the data it has seen
- **Example:** Vertical lines detected for numbers 1, 4, and 7 (which all contain vertical lines graphically)
- **Example:** Curved bottoms detected for numbers 6, 8, and 0 (which all have curved bottom features)

Output Layer (Highest Probability Answer):

- When you combine all these layers that detect vertical lines, curved bottoms, and other features, the system automatically determines the correct number

Key Insight:

Intuitively, each layer learns about a "pattern" in the data:

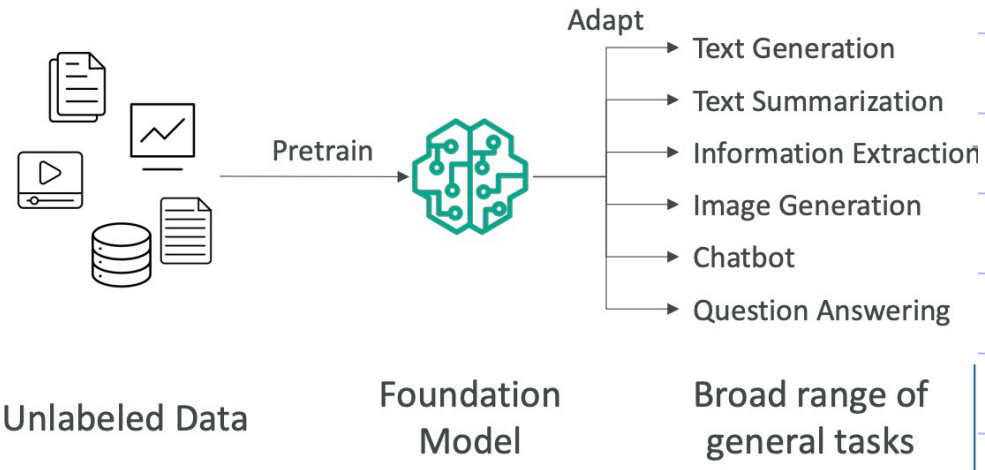
- Some layers detect vertical lines
- Some layers detect curved features
- Each layer builds upon the previous layer's findings

Important: All of this pattern recognition is not manually programmed by humans. It is automatically learned by the neural network through exposure to training data.

This represents the fundamental power of deep learning - the ability to automatically discover and learn complex patterns without explicit programming.

Generative AI

The GenAI space works by giving a lot of data to pre-train a foundational model, and this foundation model is very versatile. It can adapt to a broad range of general tasks.



What is Generative AI?

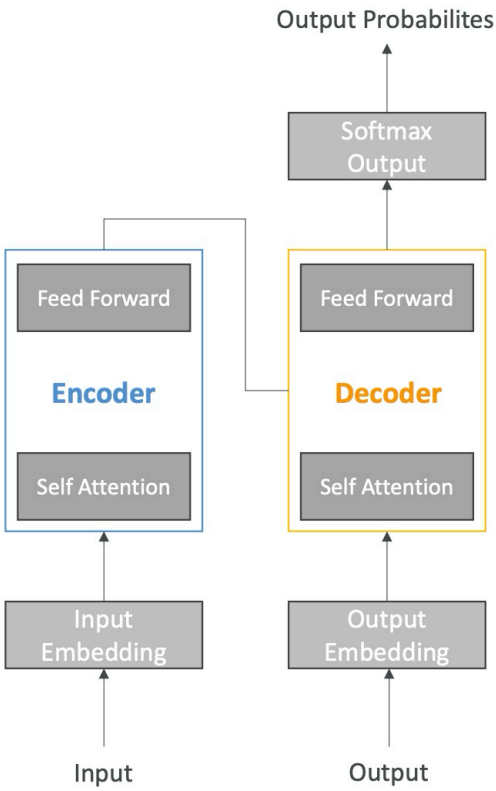
Generative AI is a subset of deep learning. Here we have a multipurpose foundation model that is actually backed by multiple neural networks. These models can also be fine-tuned as we want with our data to better fit our use cases.

(Text Only)

Transformer Models

* These GenAI models leverages transformer model.
* The transformer model is just a name (no need to remember exactly how it works).
* The idea is that it is an optimization that allows a model to process a sentence as a whole instead of word by word, which gives us :-
  → Faster text processing
  → More efficient text processing
  → less training time

* This is what the transformer model looks like :-



This is just visual of how transformer architecture looks like (won't be asked in the exams)

* Based on the architecture, you need to remember that the transformer can process sentences very efficiently and give relative importance to specific words in a sentence.

Transformer-Based Large Language Models (LLMs)

* We have transformer-based LLMs, so they are models that can understand and generate human like text. They are trained on a lot of data from :-
  - → The Internet
  - → Books
  - → Other resources

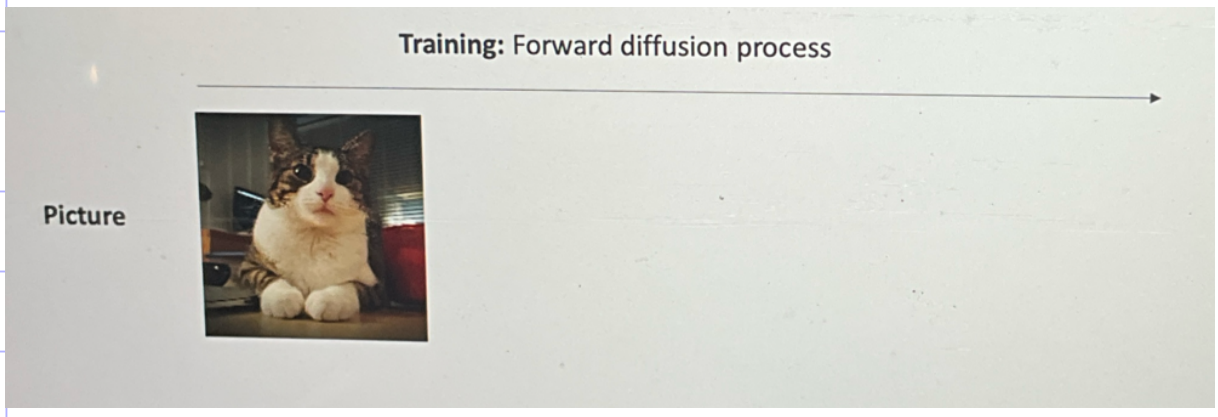
* They learn patterns very efficiently between words and phrases.

* Why transformer matter : Google Bert and OpenAI ChatGPT are based on it. ChatGPT means "Chat generative pre-trained transformer". This is why the transformer architecture is important - It is very commonly used architecture nowadays, and this is what ChatGPT and a lot of other foundation models are based on.

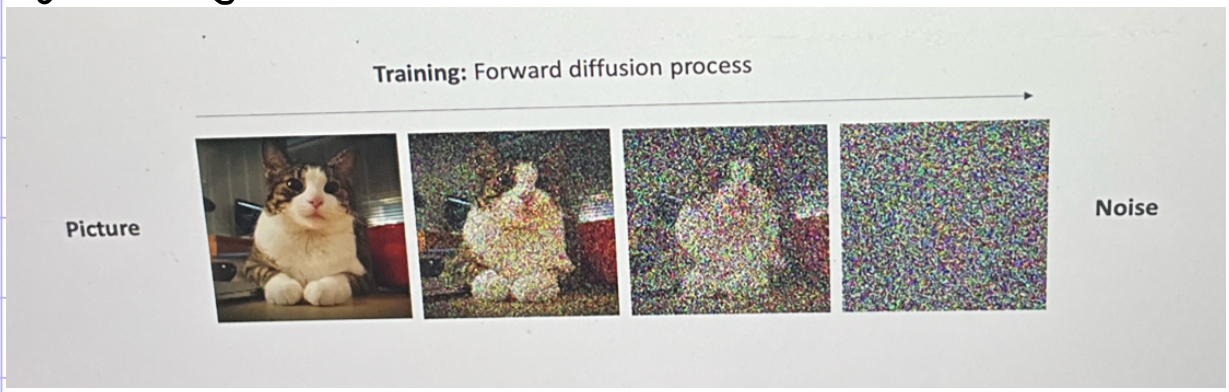
Diffusion Models (Image Only)

Note :- In Amazon Bedrock, diffusion models was explained in detail.

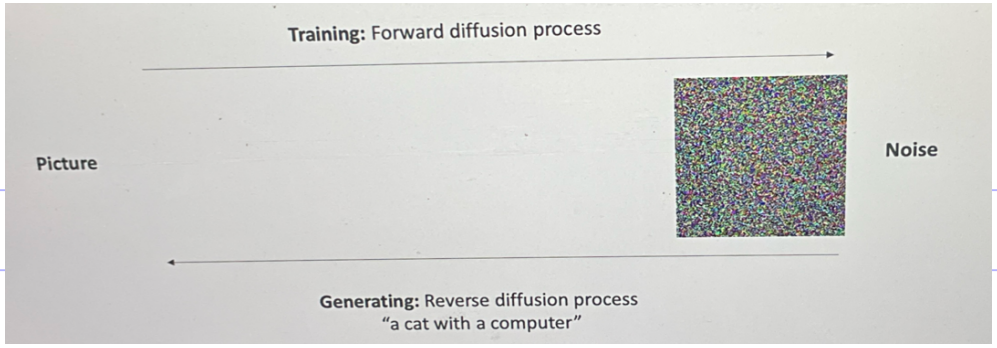
* In diffusion model, we have a picture and we do a forward diffusion process



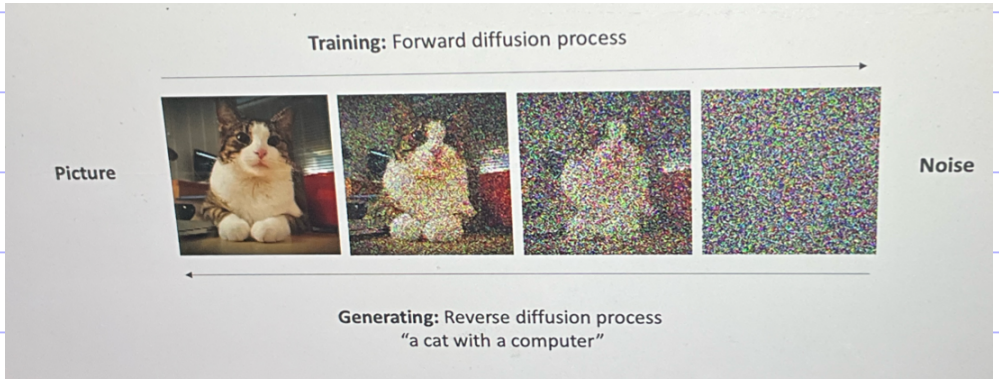
by adding noise over time.



* Then to generate images, we do the opposite. We generate from noise, back to Cat image



by providing a prompt



Multimodal Models (ex: GPT-4o)

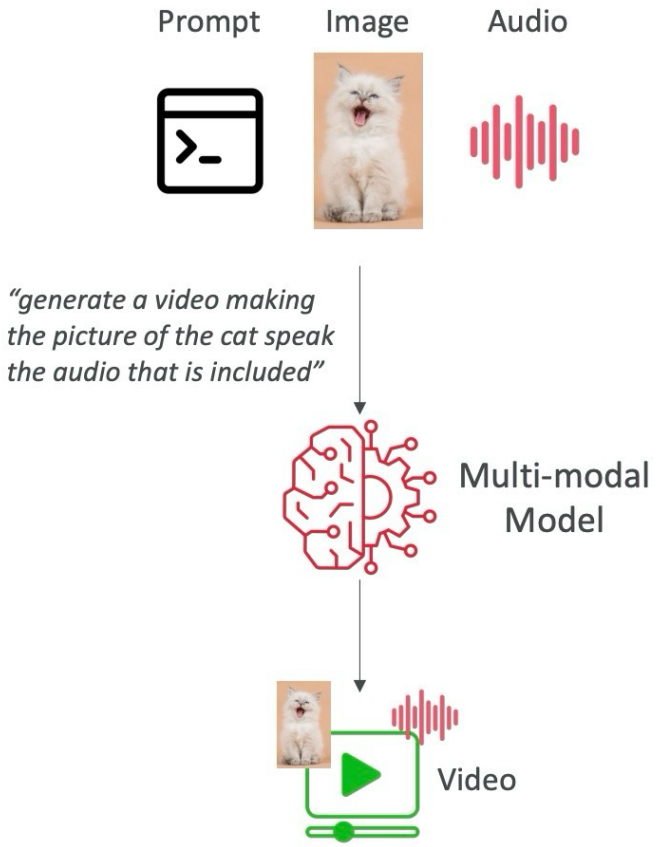
* Multi-modal models are models where we have multiple types of inputs and multiple types of outputs in terms of formats.

* For example :- ① a multimodal model can take a mix of audio, image, and text. Output of mix of video & text would be generated.

② We give an image of a cat, and we give an audio file, and as a prompt

we tell the model, "Please generate a video making the cat in the picture speak: what is included in the audio."

↳ The model will understand all these things, combine them together, and create a video of a cat with the specific audio.



Imp points to remember :-

① Multimodal does NOT :-

↳ rely on a single type of input (text, or images, or audio only)

↳ create a single type of output

The Four Levels of AI (How Humans Are a Mix of AI)

To help you understand exactly what works, humans are a mix of AI. Here are the different levels:

1. Artificial Intelligence

When we say, "Well, if this happens, then do that."

- Example: "If there's a fire in my home, put some water on it to extinguish it."

2. Machine Learning

We don't have just the if/then rules. We talk about what we've seen before, and we classify them.

- Example: We've seen a lot of dogs before, and we see a new dog, and we're very confident that this is a dog.

3. Deep Learning

We haven't seen something before, but we learned from other similar concepts what it was, and therefore, we can make a decision.

- Example: Say you have seen a lot of animals in your life - dogs, cats, elephants, giraffes, and so on. All of a sudden, you see a tiger. This is the first time you've seen a tiger, and you've never learned about the tiger, but it looks like it has legs, a mouth, it is moving, it's alive. So you say, "Well, this looks like an animal to me."

4. Generative AI

We are able to learn things from similar concepts, but now we can actually generate content. We can be creative, even though we've never seen it.

- Example: Maybe we haven't seen a specific type of poem, but we invent one because we are being very creative, and this is what humans are able to create.

This shows how in our own personal reasoning process, we use different level types of AI between artificial intelligence, machine learning, deep learning, and generative AI.