

# Evolution of AI: From Rule-based to Generative Models

**Jiaul Paik**

**Department of AI  
IIT Kharagpur**

*jiaul@ai.iitkgp.ac.in*

# Outline of the Talk

- **Background**
- **A Brief History of AI**
- **Stages of AI Evolution**
  - Expert Systems
  - Machine Learning
  - Deep Learning
  - Generative Models
- **The Future and Ethical Issues**



AI4ICPS



IIT Kharagpur

# Background

# What is AI? (Very Informal)

- AI is the capability of a computer system to perform tasks that typically require *human intelligence*
  - Learning
  - Reasoning
  - Problem-solving



AI4ICPS

IIT Kharagpur

# Basic Architecture of Intelligence





AI4ICPS



# Basic Architecture of Intelligence

30<sup>th</sup> Over? Defend

45<sup>th</sup> over? Go for 6!

An in-swinger

Reasoning

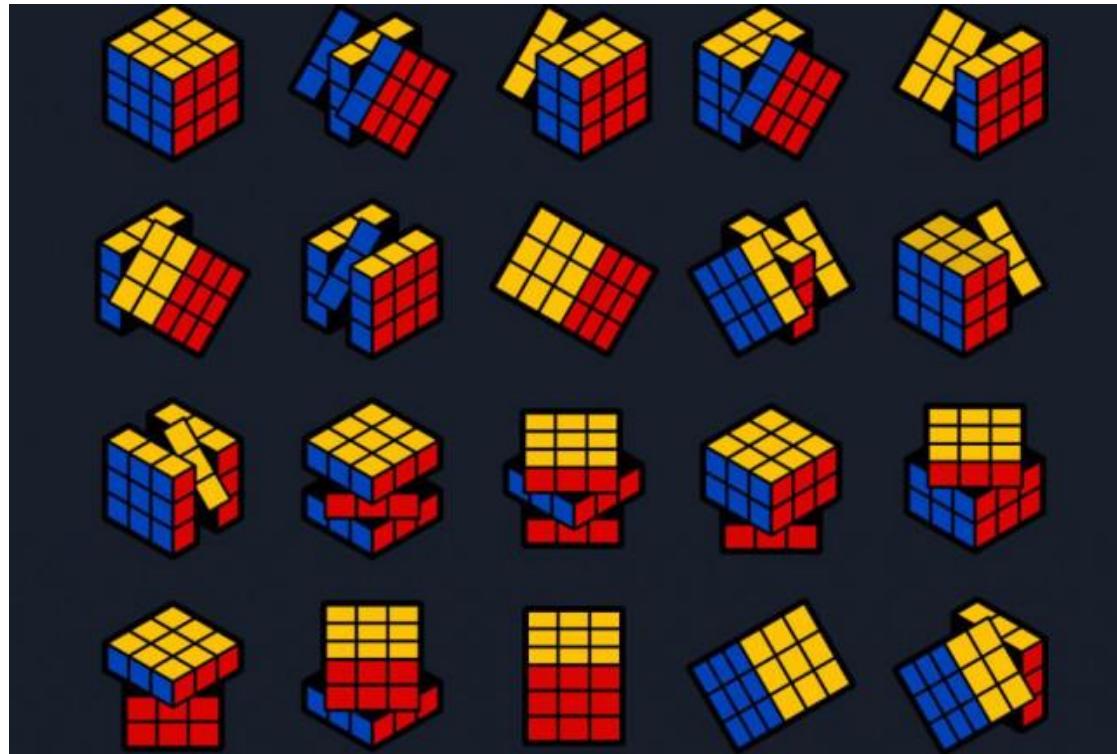
In similar situation, I hit down  
the ground in earlier match

Classification

Reinforcement



# Decision Making: Generalized Problem Solving



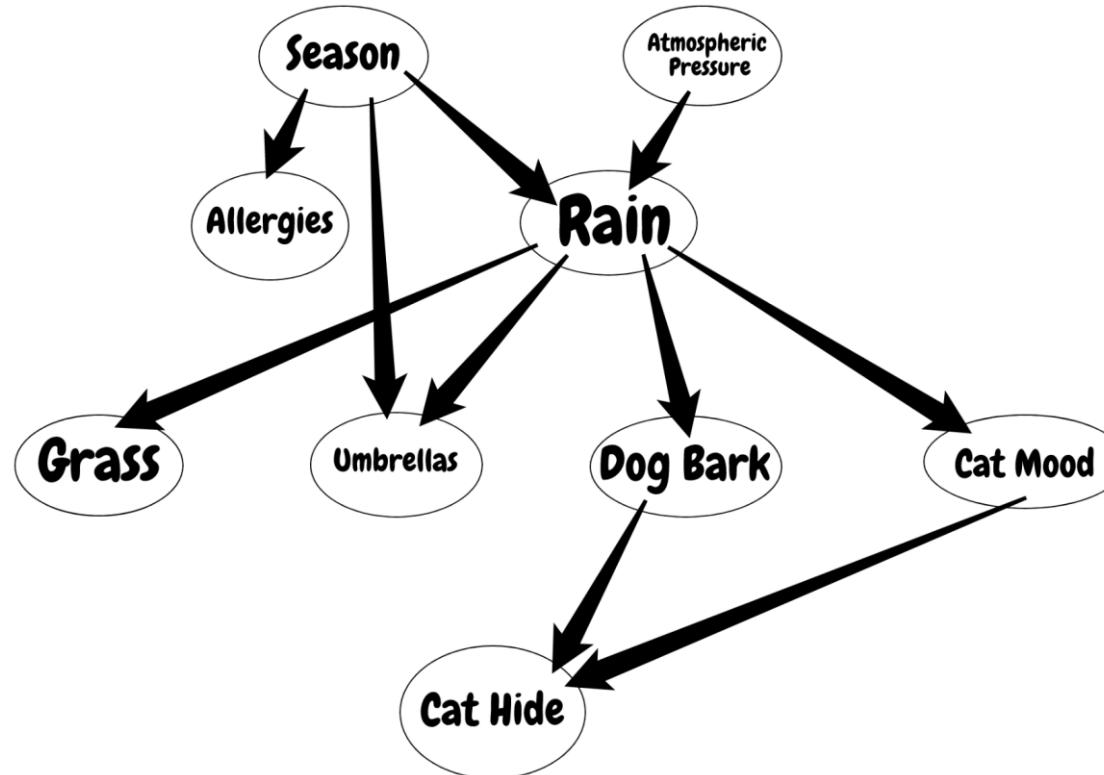
Planning Problem



AI4ICPS



# Decision Making: Reasoning under Uncertainty





# Reasoning + Decision Making



IBM Deep Blue wins over Garry Kasparov

# Types of AI

- **Artificial Narrow Intelligence (ANI)**
- **Artificial general Intelligence (AGI)**
- **Artificial Super Intelligence (ASI)**



AI4ICPS



# Types of AI

ANI

- Performs specialized tasks

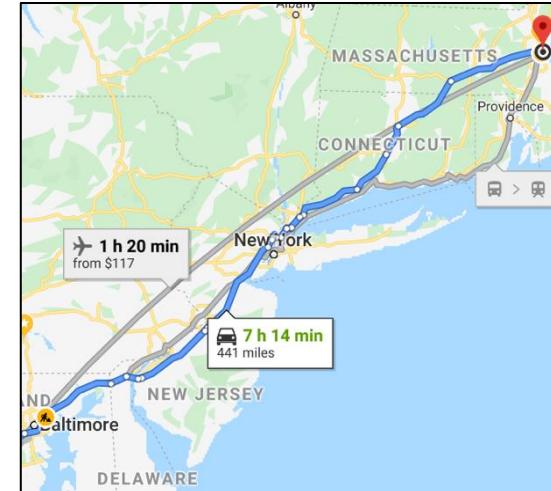
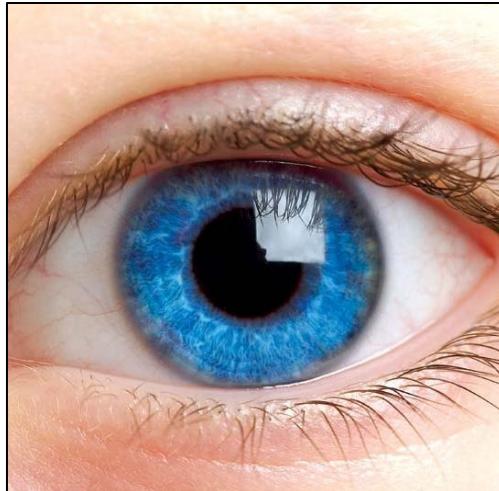
AGI

- Capable of human level cognitive skills
- But, not yet exists (some hits, many misses)

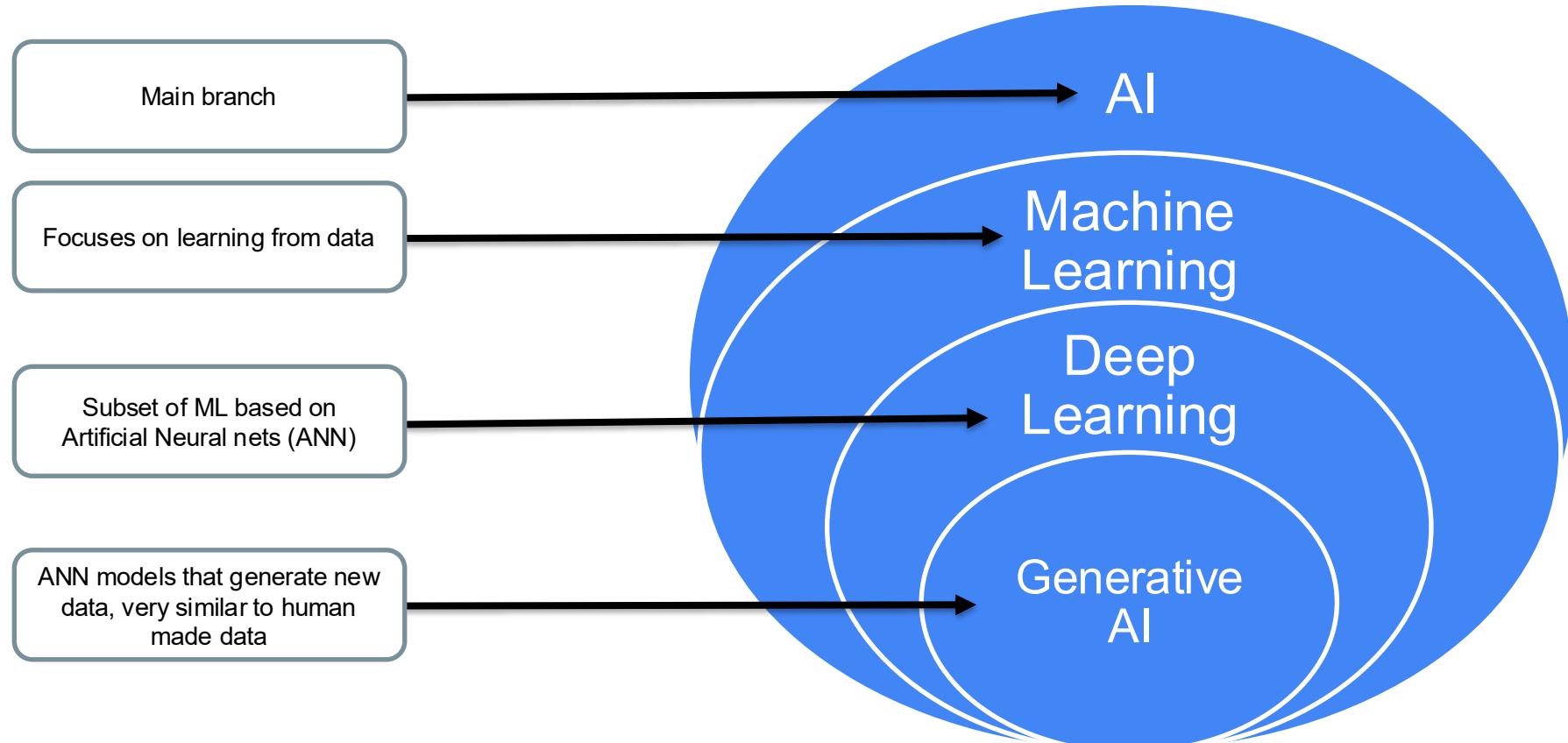
ASI

- More than a human can do
- Too speculative, too far away

# Typical “narrow” tasks: vision, NLP, and Planning.



# AI & Machine Learning & Deep Learning





AI4ICPS



IIT Kharagpur

# A Brief History of AI



AI4ICPS

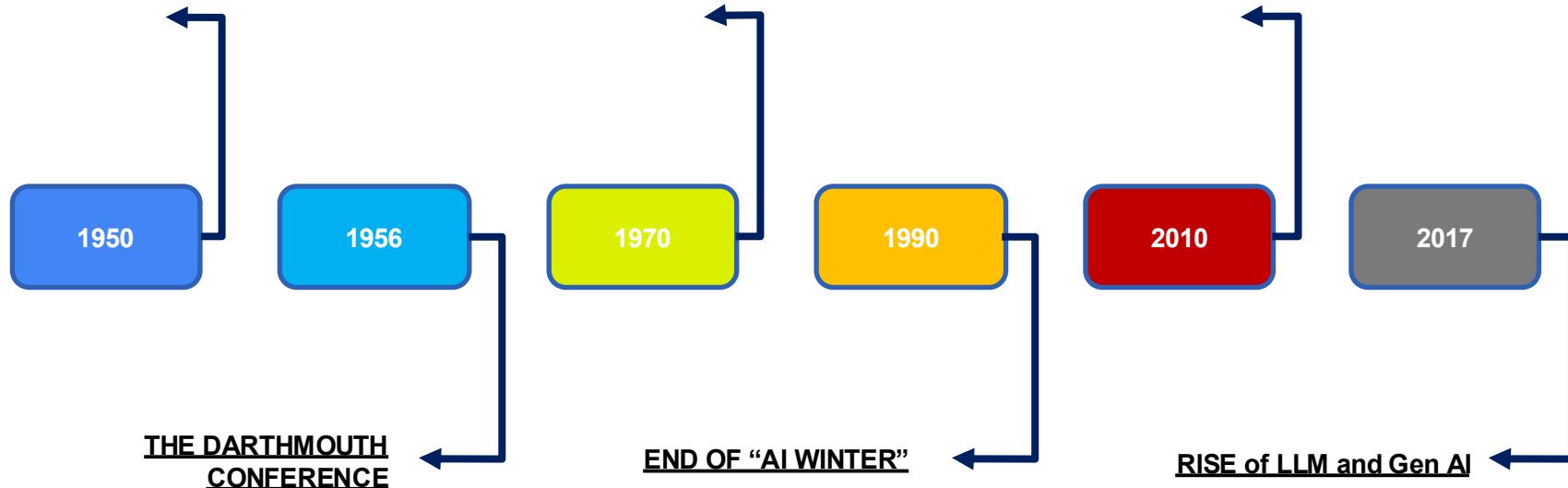
IIT Kharagpur

# A Brief History of AI

CAN MACHINES THINK?

“AI WINTER”  
ABRUPT HALT OF AI  
DEVELOPMENT

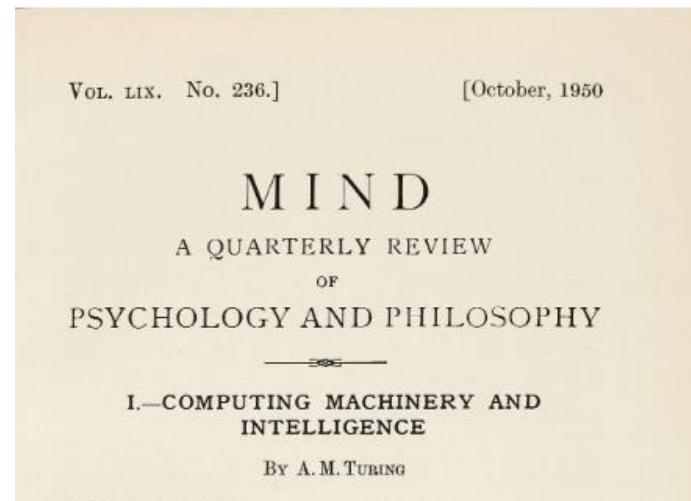
“AI SPRING”  
AI RENAISSANCE



# History: 1950

- **Alan Turing asked:**

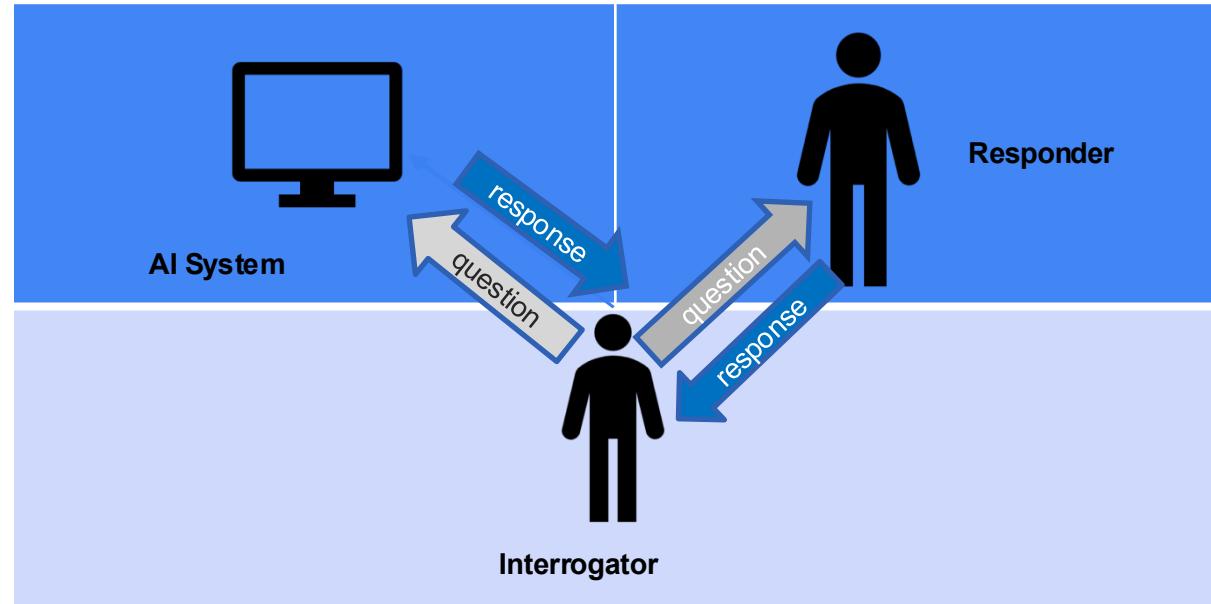
- Can machines think?
- How could you tell?





# Turing Test

- Interrogator asks a question
- Gets answers from both human and machine
- Can't identify who is what



# 1956: Dartmouth Conference

- Convened the leading thinkers of the day (Shannon, Minsky among them)
- Goal was to make a "significant advance" in 2 months
- Coined the term "artificial intelligence"
  
- **Overwhelming Optimism from the great minds ....**
  - **Herbert Simon:** "Machines will be capable, within twenty years, of doing any work a man can do."
  - **Marvin Minsky:** "Within 10 years the problems of artificial intelligence will be substantially solved."
  - **Claude Shannon:** "I visualize a time when we will be to robots what dogs are to humans, and I'm rooting for the machines."

# History: 1966

- **ALPAC (Automatic Language Processing Advisory Committee)** report cut off government funding for machine translation
- **But why this happened?**
  - **Limited computation:** search space grew exponentially, outpacing hardware
  - **Limited information:** complexity of AI problems (number of words, objects, concepts in the world)

# History: 1970 – 80s

- **Knowledge-based systems**
  - Encode specific domain knowledge from experts in form of rules
- **Examples:**
  - DENDRAL: infer molecular structure from mass spectrometry
  - MYCIN: diagnose blood infections, recommend antibiotics
  - XCON: convert customer orders into parts specification

# History: 1943 onwards, Intro to Neural Net

- **1943:** artificial neural networks, relate neural circuitry
- **1949:** "cells that fire together wire together" learning rule (Hebb)
- **1958:** Perceptron algorithm for linear classifiers (Rosenblatt)
- **1969:** Perceptrons book showed that linear models could not solve XOR, killed neural nets research

# History: 1985 onwards

- **Statistical machine learning (Learning from data)**

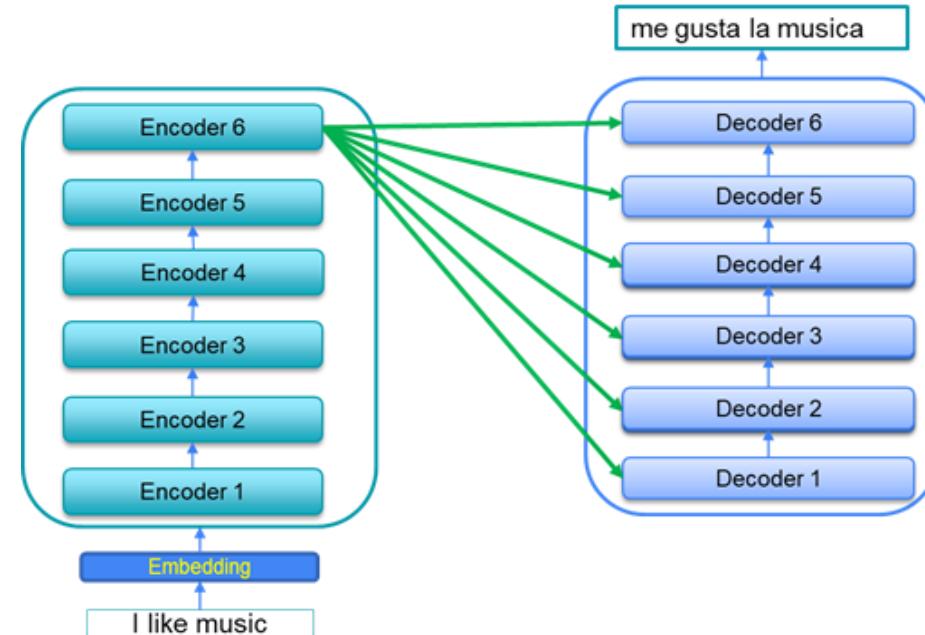
- 1985: Bayesian networks
- 1995: support vector machines
- 1999: variational inference
- 2001: conditional random fields
- 2003: topic modeling

# History: 2006 onwards, Deep Learning Era

- **2006:** unsupervised layerwise pre-training of deep networks
- **2012:** AlexNet obtains huge gains in object recognition; transformed computer vision community
- **2014:** sequence-to-sequence modeling (for machine translation)
- **2015:** Attention mechanism (for machine translation)
- **2016:** AlphaGo uses deep reinforcement learning, defeat world champion

# History: 2017 onwards, Era of LLM and Gen AI

- **2017: Transformer Model**





AI4ICPS

IIT Kharagpur

# Rule-based Expert Systems

# What is an Expert System?

- **A computer program with**
  - Contains expert knowledge about a particular problem,
  - Often in the form of a set of if-then rules,
  - Able to solve problems at a level equivalent to human experts

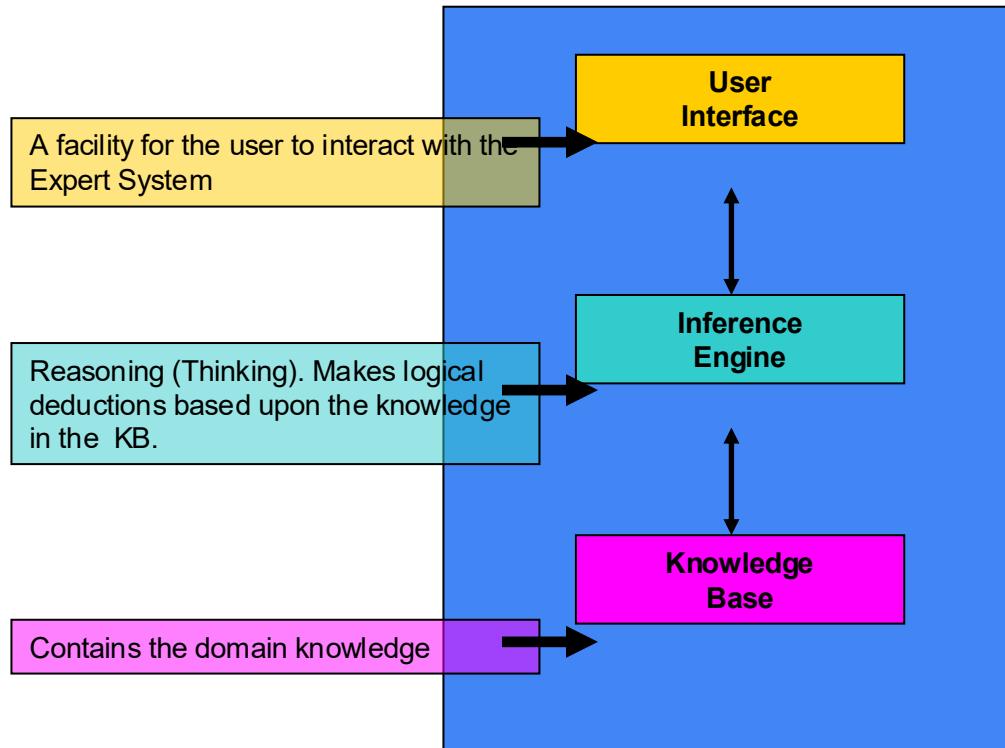
# Building Expert Systems

- **Objective of an expert system**
  - To transfer expertise from human experts to a computer system and
  - Then on to non-experts
- **Activities**
  - Knowledge acquisition
  - Knowledge representation
  - Knowledge inferencing
  - Knowledge transfer to the user

# Expert Systems: Some Use-cases

- **Diagnose human illnesses**
- **Make financial forecasts**
- **Schedule routes for delivery vehicles**
- **Decide the strategy in computer games**

# Important Expert System Components



# Production Rules

- **Condition-Action Pairs:**
  - A RULE consists of an IF part and a THEN part (also called a *condition* and an *action*). if the IF part of the rule is satisfied; consequently, the THEN part can be concluded, or its problem-solving action taken.
- **Rules represent a *model* of actual human behavior**

# Expert Systems Benefits

- Helps preserve knowledge
- Helps if expertise is scarce, expensive, or unavailable
- Helps if under time and pressure constraints
- Helps in training new employees

# Limitations of Expert Systems

- 1) Knowledge is not always readily available
- 2) Expertise can be hard to extract from humans
- 3) Expert Systems work well only in a *narrow domain* of knowledge
- 4) Knowledge engineers are rare and expensive

# Examples of Some Expert System

- **MYCIN (1972-80)**
  - An interactive program that diagnoses certain infectious diseases, prescribes antimicrobial therapy, and can explain its reasoning in detail
  
- **DENDRAL (1965-83)**
  - Rule-based expert systems that analyzes molecular structure. DENDRAL proposes plausible candidate structures for new or unknown chemical compounds.



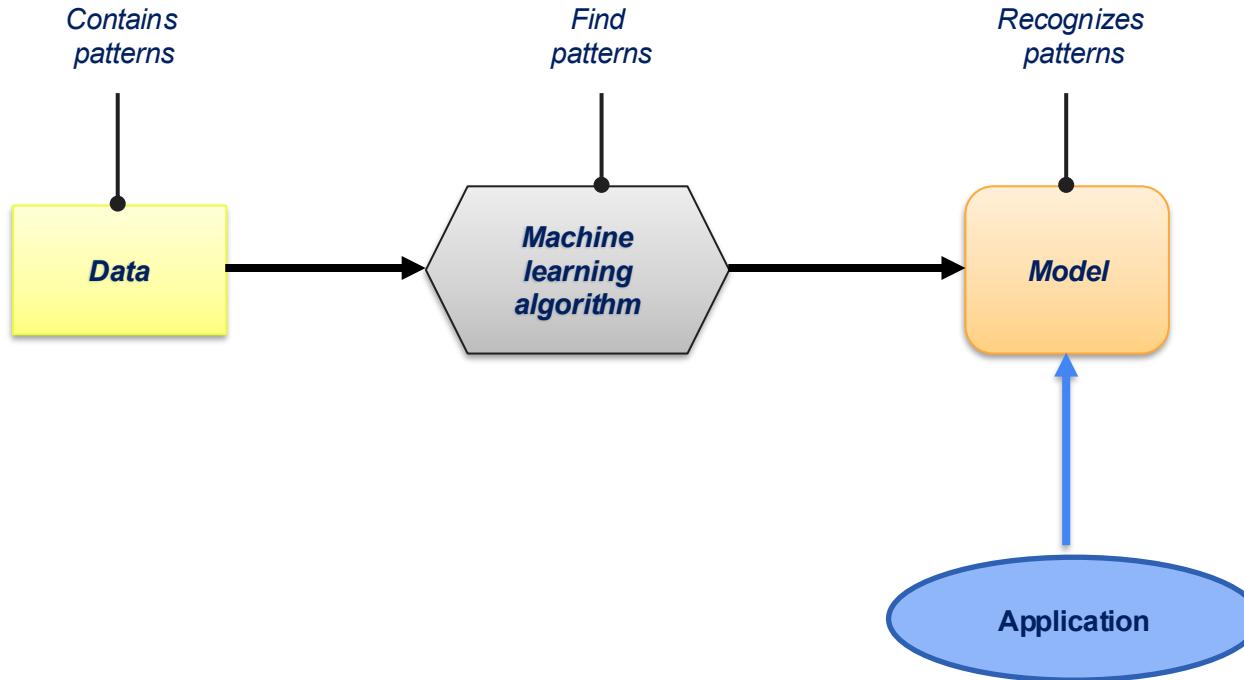
AI4ICPS

IIT Kharagpur

# Machine Learning

(Learning from Data)  
(Often with human generated features)

# Machine Learning in Nutshell





AI4ICPS

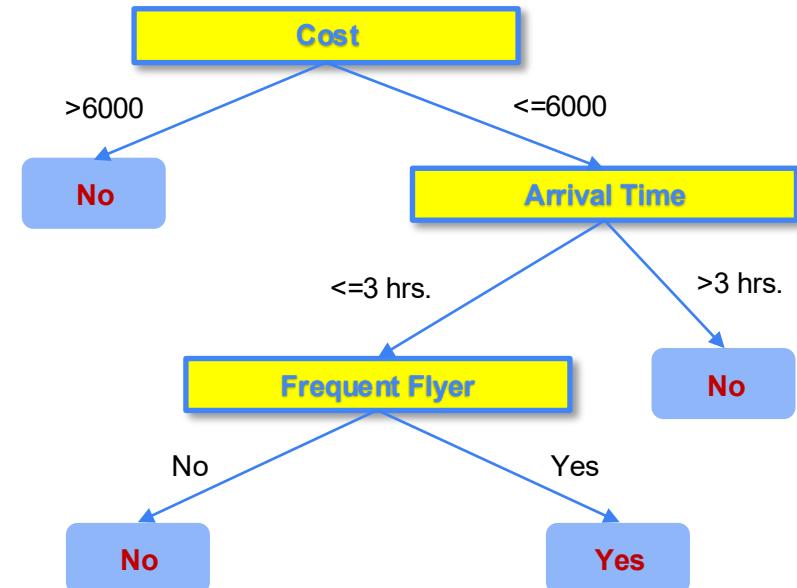


# How does a human mind decide?

Flight	Time	Cost	Frequent flyer?	Buy ticket?
AB234	5	10000	N	N
CD567	3	6000	N	N
JE819	2	5000	Y	Y
BG567	7	8000	Y	N

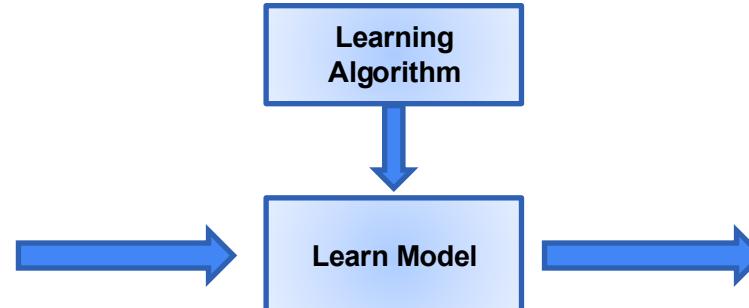
I choose JE819

Why did I choose it?



# How does Machine Learning decide?

Flight	Time	Cost	Frequent flyer?	Buy ticket?
AB234	5	10000	N	N
CD567	3	6000	N	N
JE819	2	5000	Y	Y
BG567	7	8000	Y	N



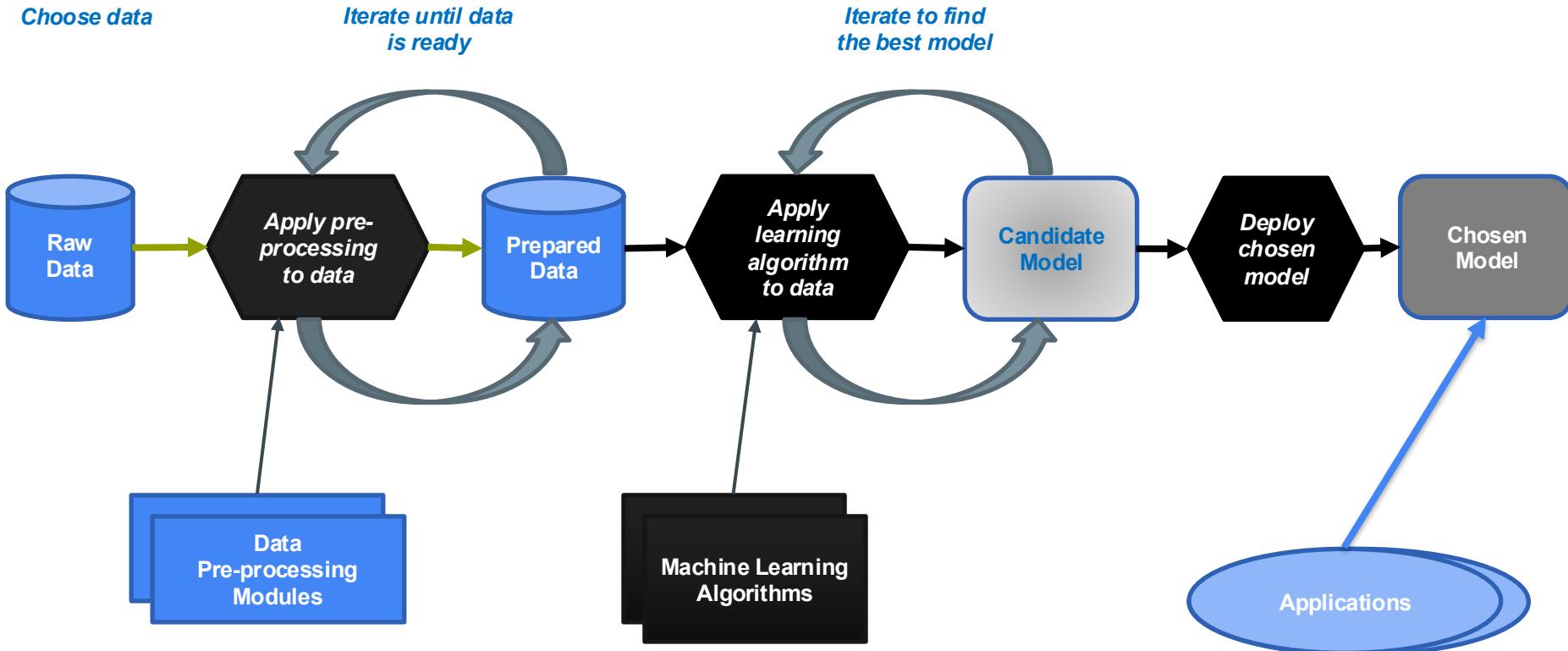
## Test Set

Flight	Time	Cost	Frequent flyer?	Buy ticket?
ZZ234	4	9000	N	?
PD567	5	7000	Y	?

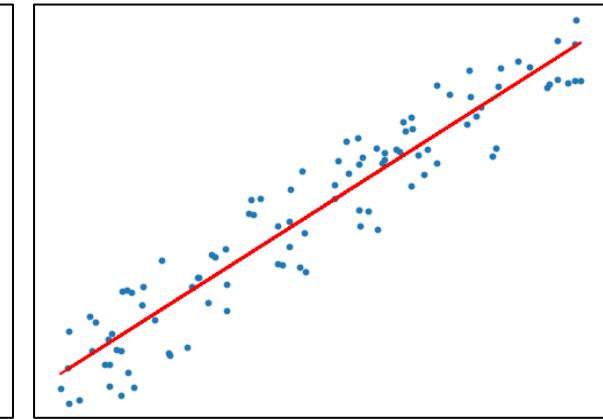
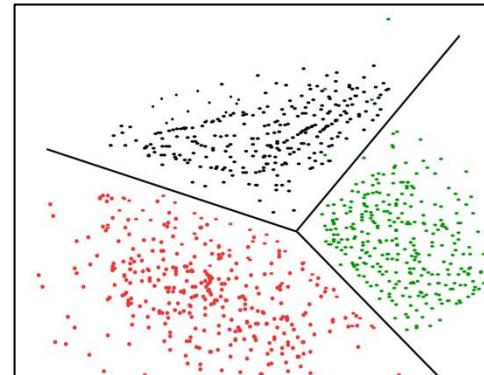




# Machine Learning Pipeline



Machine learning can perform many tasks:  
classification, clustering, and regression.

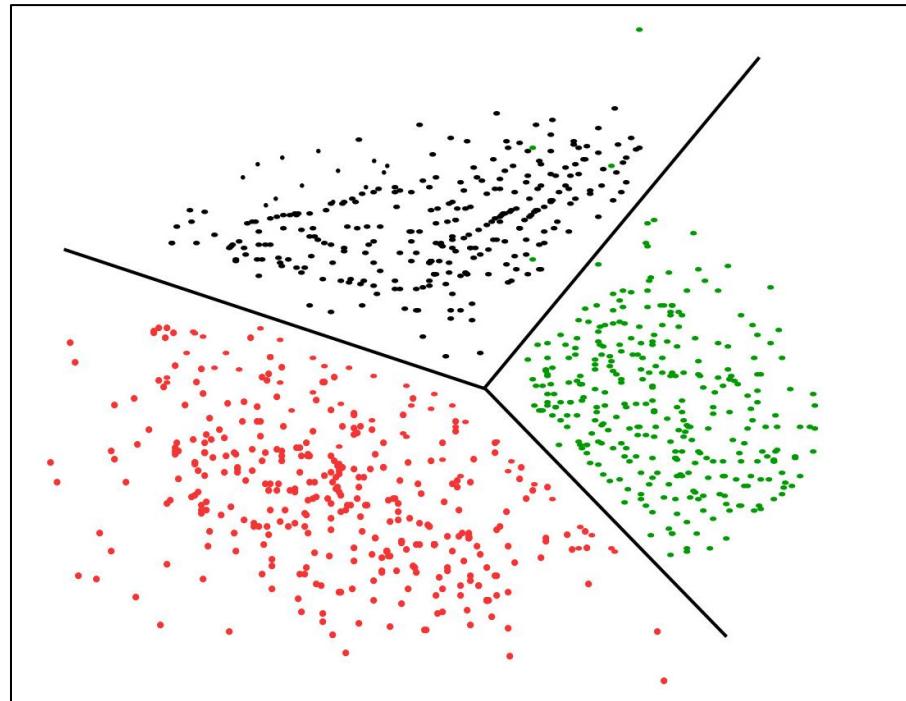


Machine learning can perform many tasks,  
i.e. **classification**, clustering, and regression.

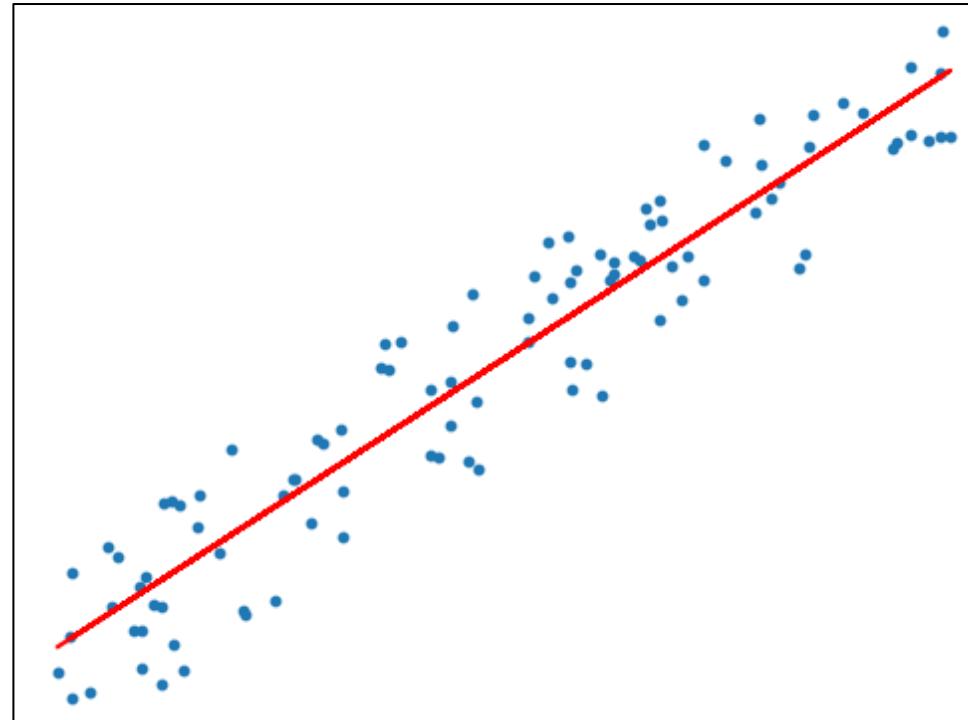


CAT

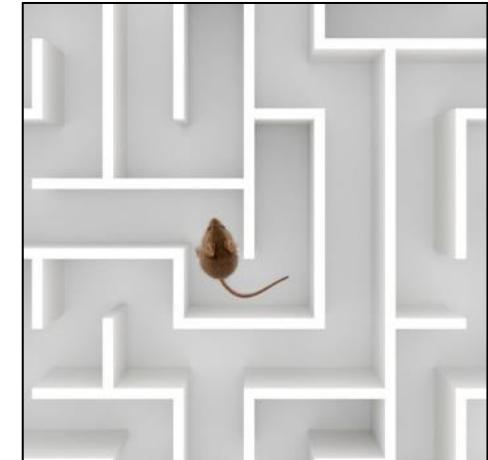
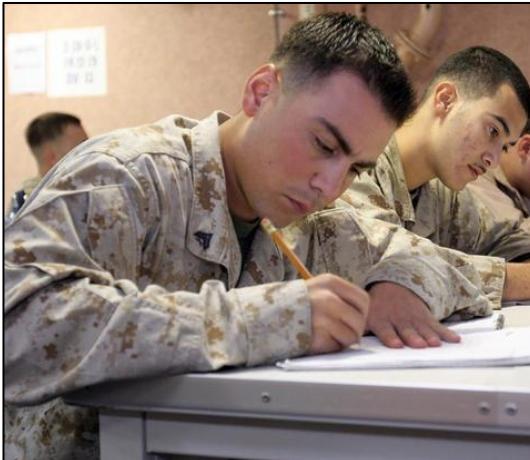
Machine learning can perform many tasks,  
i.e. classification, **clustering**, and regression.



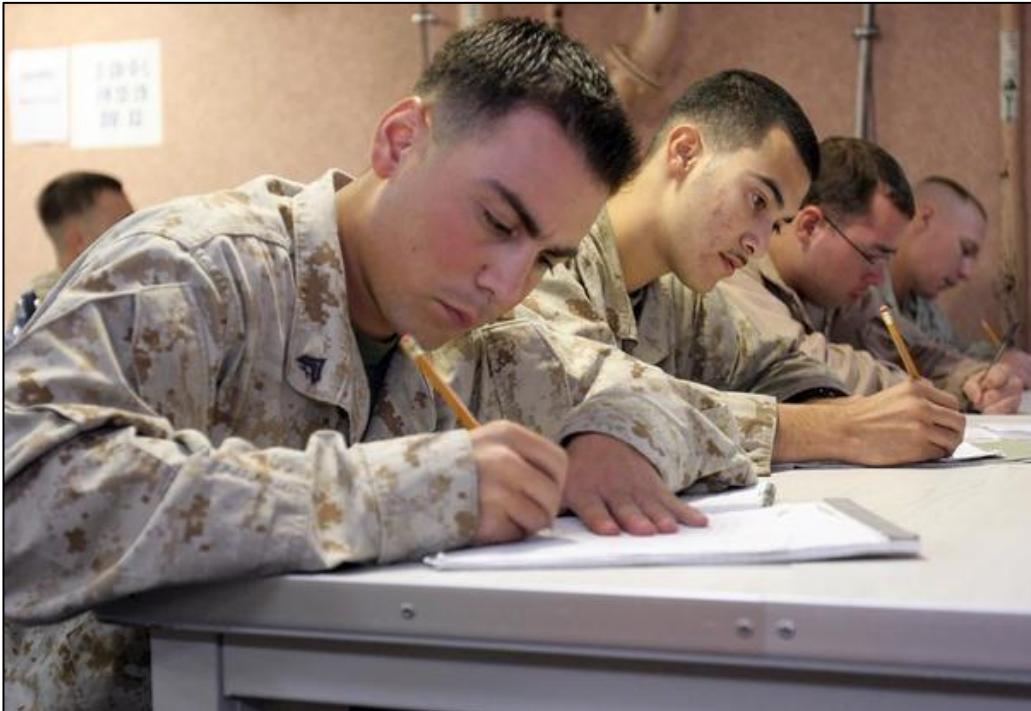
Machine learning can perform many tasks,  
i.e. classification, clustering, and **regression**.



# Three types of learning: Supervised, unsupervised, and reinforcement learning.

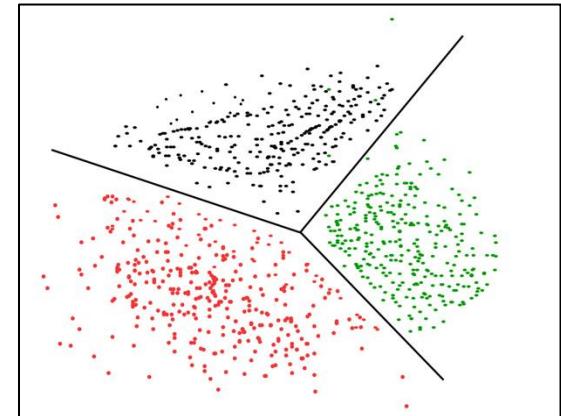


There are three types of learning: **supervised**, unsupervised, and reinforcement learning.



CAT

There are three types of learning: supervised, **unsupervised**, and reinforcement learning.



There are three types of learning: supervised, unsupervised, and **reinforcement learning**.

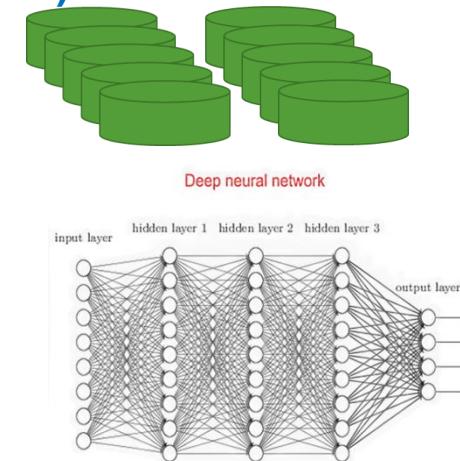


# Deep Learning

(Automatic Feature Generation)  
(Representation Learning)

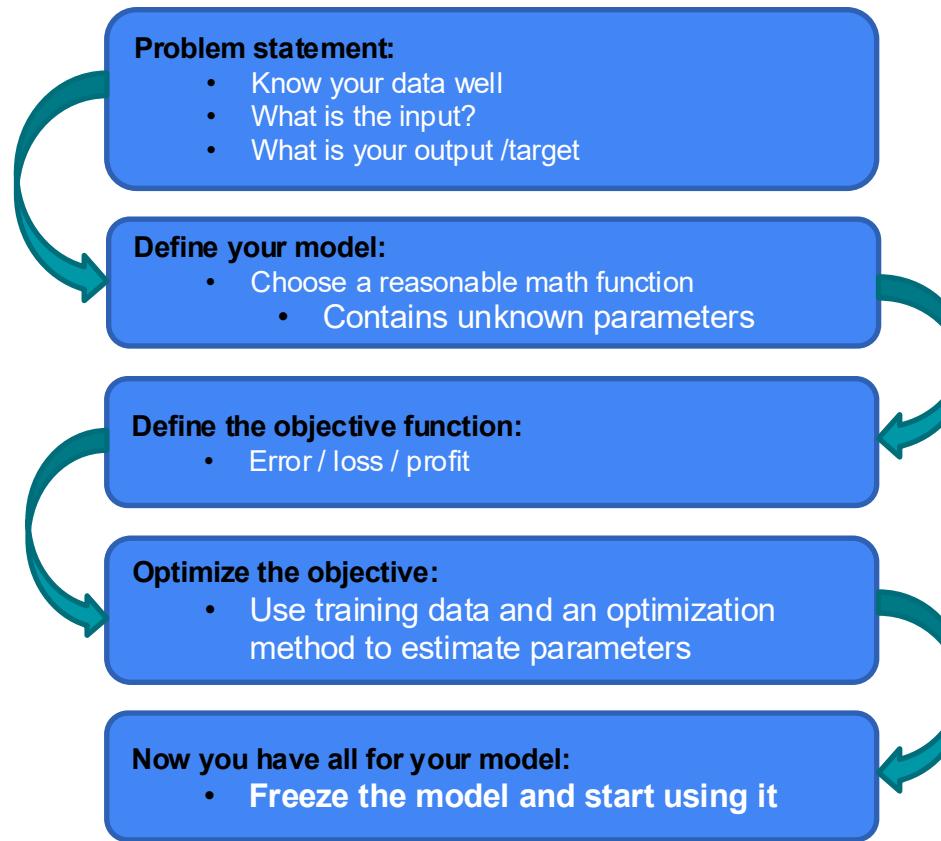
# Three Pillars of Deep Learning/Modern AI

- **Data (as the repository of knowledge and wisdom)**
- **Large Models (as the learning engine)**
- **GPU (as the computing powerhouse)**



Source: nvidia.com

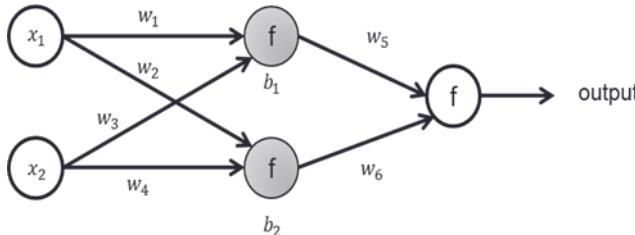
# Learning Pipeline



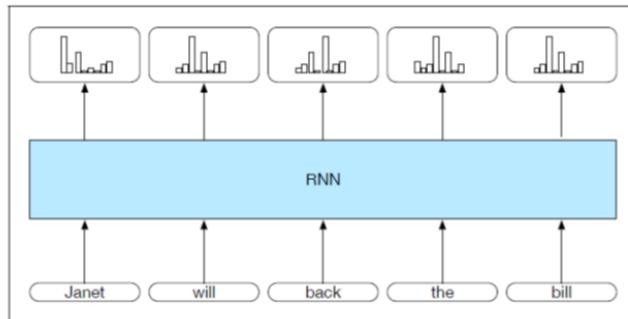
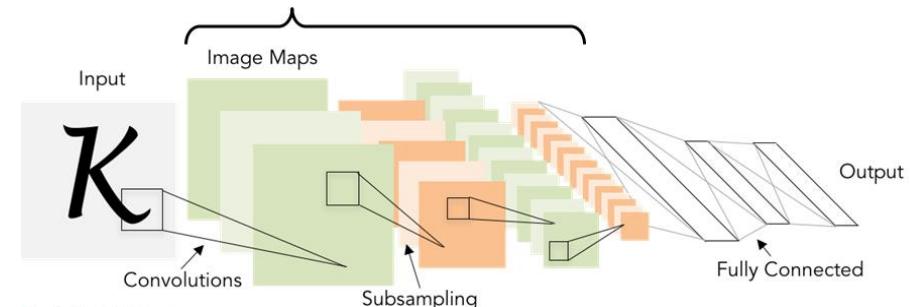


# Deep Neural Nets: Big Four of Modern AI

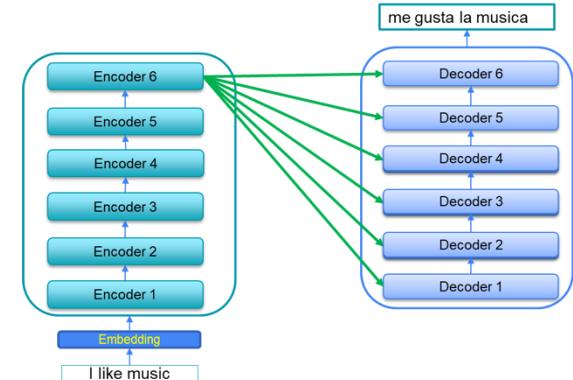
## Feedforward Neural net



## Convolutional Neural net: For Image/videos



RNN: For Sequence data



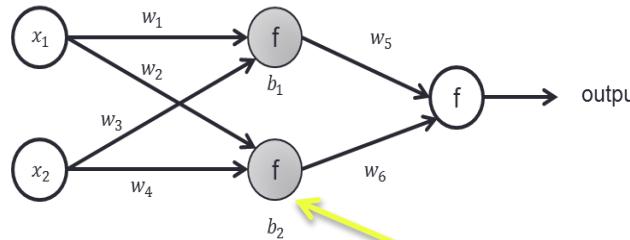
Transformer: For Languages/sequences



# Feedforward Neural Network

Movie Name	Mary's Rating	John's rating	I like?
Lord of the rings	1	3	No
Roman Holiday	4	5	Yes
Life is Beautiful	4	4	Yes
LA Confidential	2	3	?

Each neuron takes inputs from nodes from previous layer,  
linearly combines them,  
and then passes through the activation function



$$\text{Input: } w_2 * x_1 + w_4 * x_2 + b_2$$

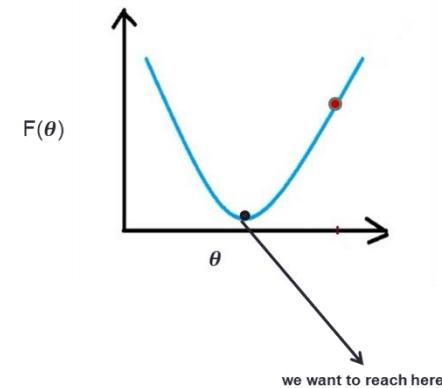
$$\text{Output: } \mathbf{f}(w_2 * x_1 + w_4 * x_2 + b_2)$$

**f** is the activation function (e.g., sigmoid function)

# Optimization: Gradient Descent Algorithm

1. Randomly set the values of parameters (thetas)
2. Repeat until convergence

- $\theta_j^{t+1} = \theta_j^t - r * \frac{\partial E}{\partial \theta_j}$  for all j

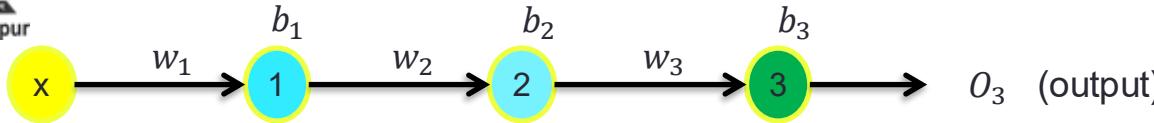




AI4ICPS



# Training Neural Network: Backpropagation



$$I_1 = w_1 * x + b_1 \quad I_2 = w_2 * f(I_1) + b_2 \quad I_3 = w_3 * f(I_2) + b_3$$

$$O_1 = f(I_1)$$

$$O_2 = f(I_2)$$

$$O_3 = f(I_3)$$

- Error/performance function:  $E = \frac{1}{2}(O_3 - y)^2$

## Gradients:

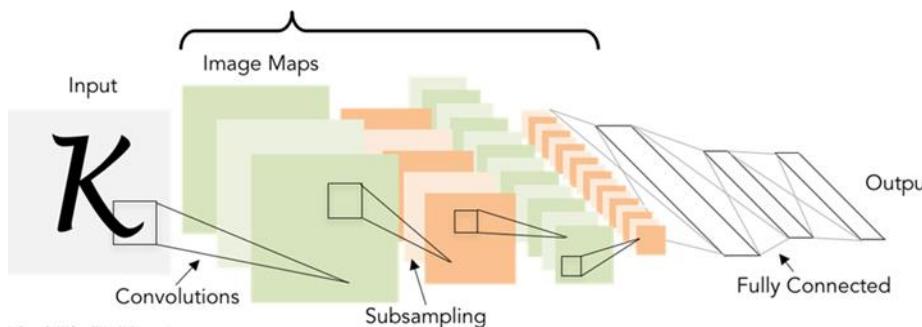
$$\frac{\partial E}{\partial w_3} = \boxed{\frac{\partial E}{\partial O_3} \times \frac{\partial O_3}{\partial I_3}} \times \frac{\partial I_3}{\partial w_3}$$

$$\frac{\partial E}{\partial w_2} = \boxed{\frac{\partial E}{\partial O_3} \times \frac{\partial O_3}{\partial I_3}} \times \boxed{\frac{\partial I_3}{\partial O_2} \times \frac{\partial O_2}{\partial I_2}} \times \frac{\partial I_2}{\partial w_2}$$

$$\frac{\partial E}{\partial w_1} = \boxed{\frac{\partial E}{\partial O_3} \times \frac{\partial O_3}{\partial I_3}} \times \boxed{\frac{\partial I_3}{\partial O_2} \times \frac{\partial O_2}{\partial I_2}} \times \boxed{\frac{\partial I_2}{\partial O_1} \times \frac{\partial O_1}{\partial I_1}} \times \frac{\partial I_1}{\partial w_1}$$

# Convolutional Neural Network (CNN)

For images and computer vision tasks

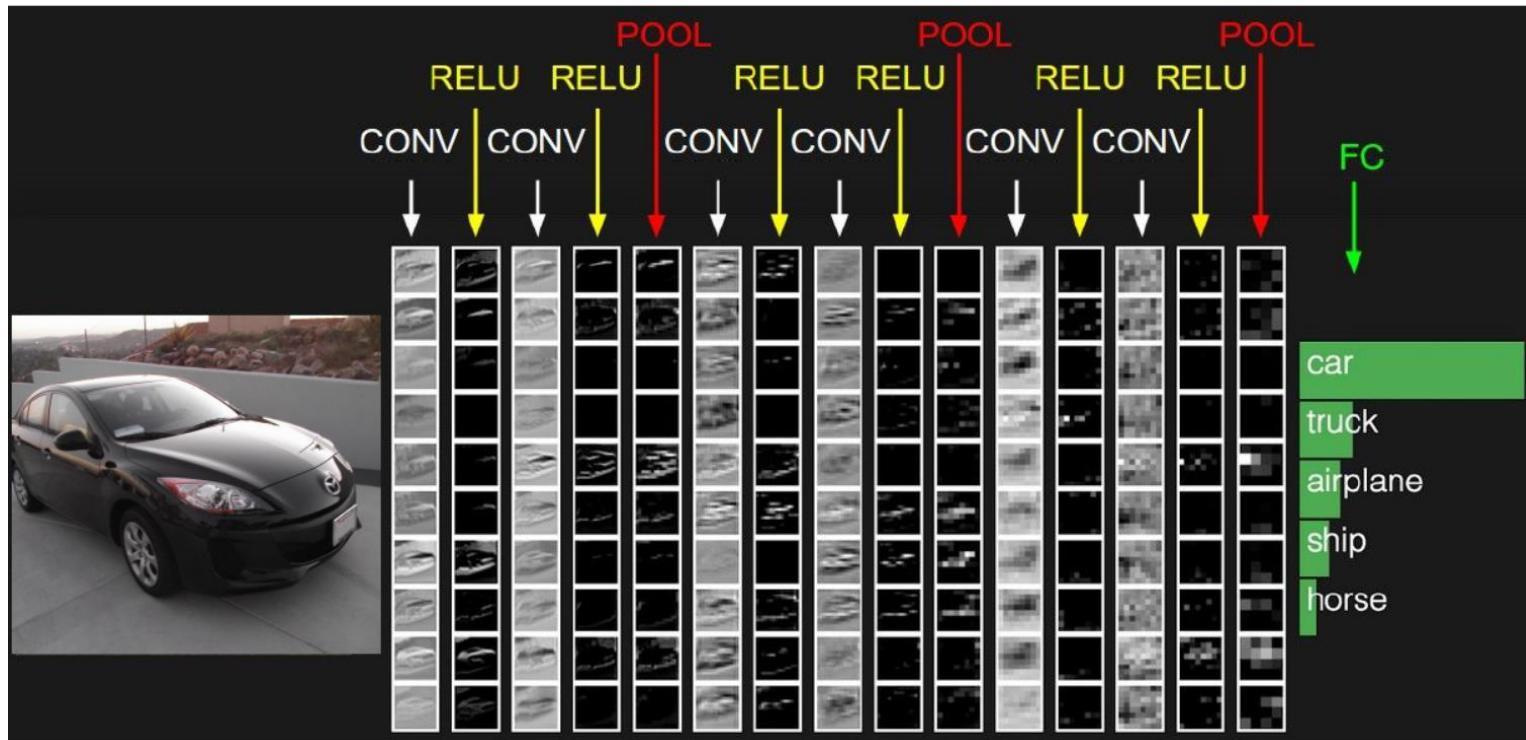




AI4ICPS



# CNN: An Example

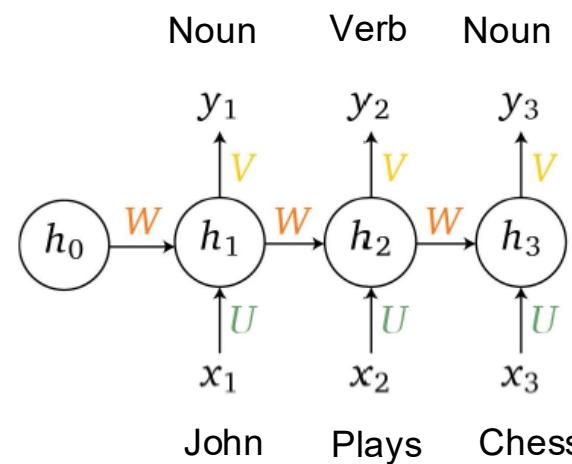


# RNN: For Sequence Data

**Problem:** Parts of speech tagging.

**Input:** John plays chess.

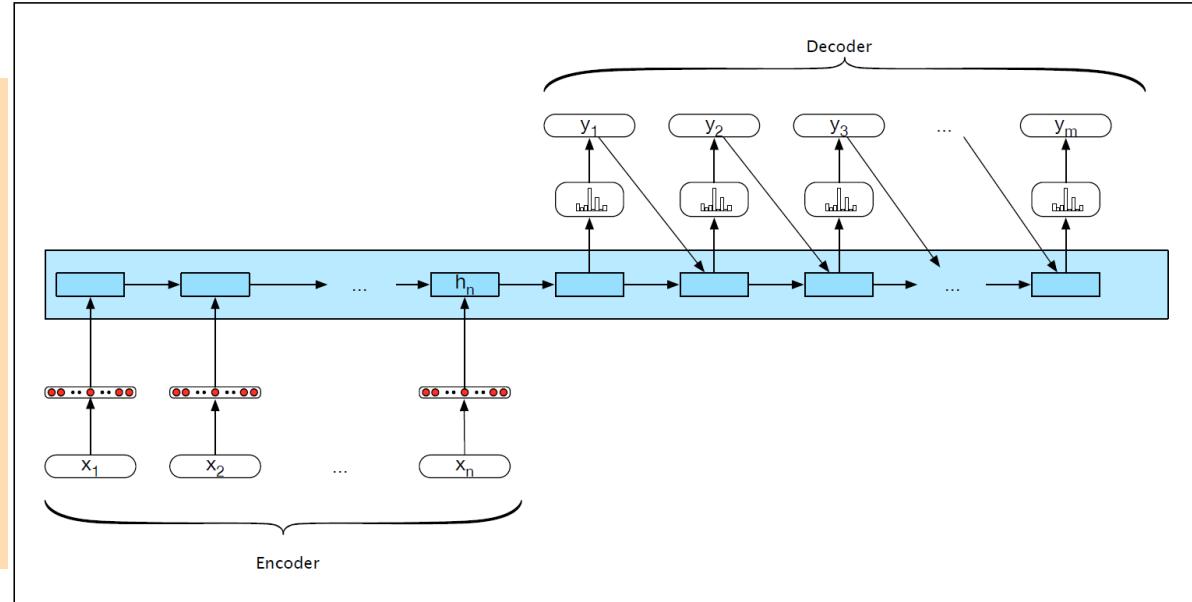
**Output:** Noun Verb Noun.



# Simple Encoder Decoder Networks

## Limiting design choices

- **E** and **D** have the same internal structure
- Final state of the **E** is the only context available to **D**
- This context is only available to **D** as its initial hidden state.



- Encoder generates a contextualized representation of the input (last state).
- Decoder takes that state and autoregressively generates a sequence of outputs

# AI Application Domain



Precision Medicine



Precision Agriculture



Online Retail



Social Media



Planetary Missions



Assistive Technology



AI4ICPS



# AI on Languages

## Information Retrieval



Doc 1  
Doc 2  
Doc 3

## Sentiment Analysis



## Information Extraction



## Machine Translation



# Natural Language Processing

## Question Answering



Human: When was Apollo sent to space?

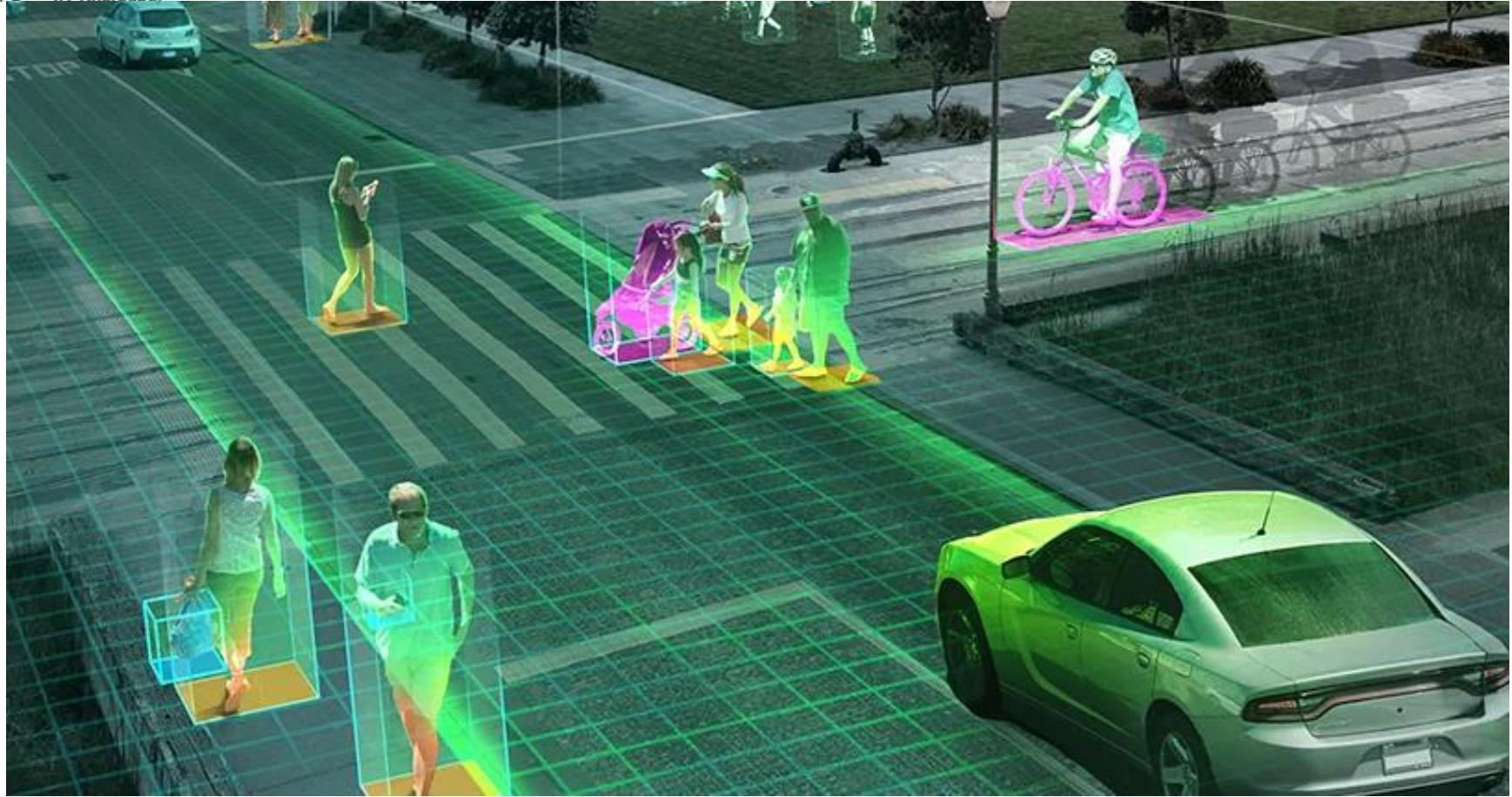
Machine: First flight - AS-201, February 25, 1966



AI4ICPS

IIT Kharagpur

# Computer Vision



# CNN Applications: Segmentation



Source: Farabet et al., 2012

# CNN Applications: Self Driving Car



Source: *Fei Fei Li*

# CNN Applications: Image Captioning



*A white teddy bear sitting in the grass*



*A man in a baseball uniform throwing a ball*



*A woman is holding a cat in her hand*



*A man riding a wave on top of a surfboard*



*A cat sitting on a suitcase on the floor*



*A woman standing on a beach holding a surfboard*



AI4ICPS

IIT Kharagpur

# Game Playing: Alpha Go





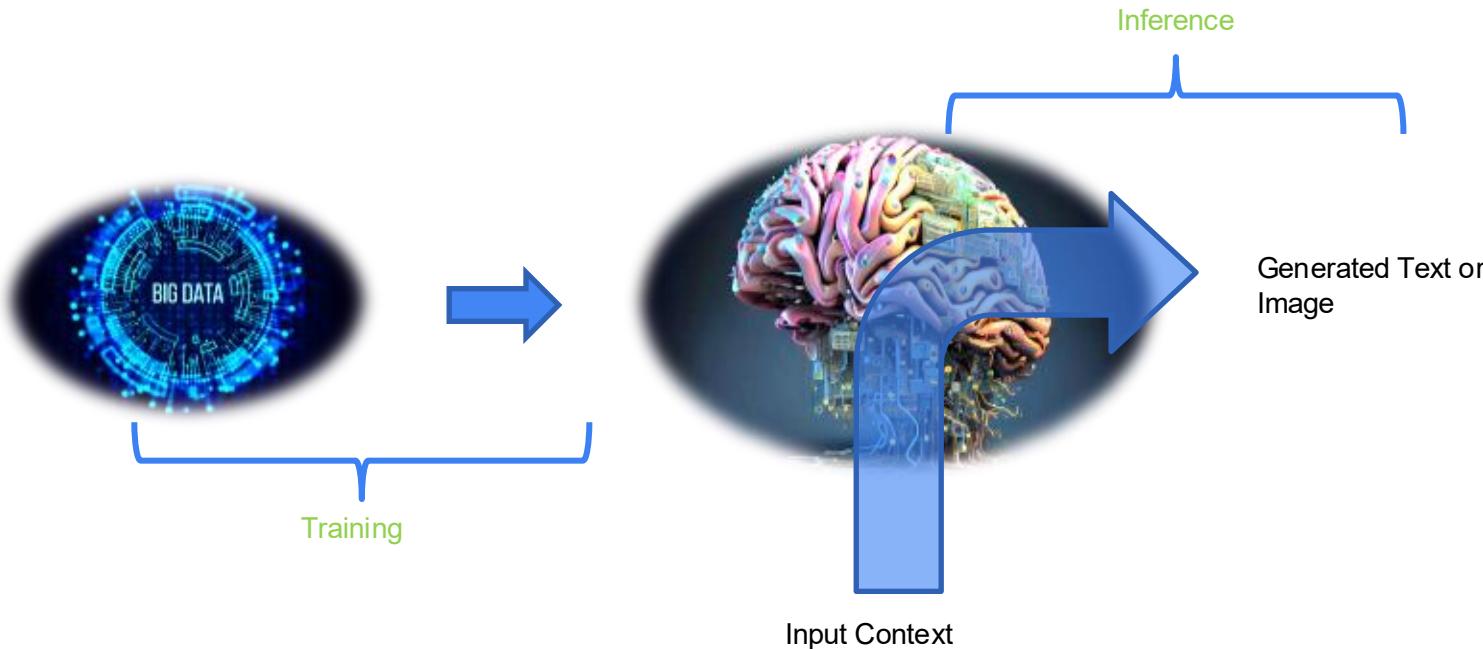
AI4ICPS



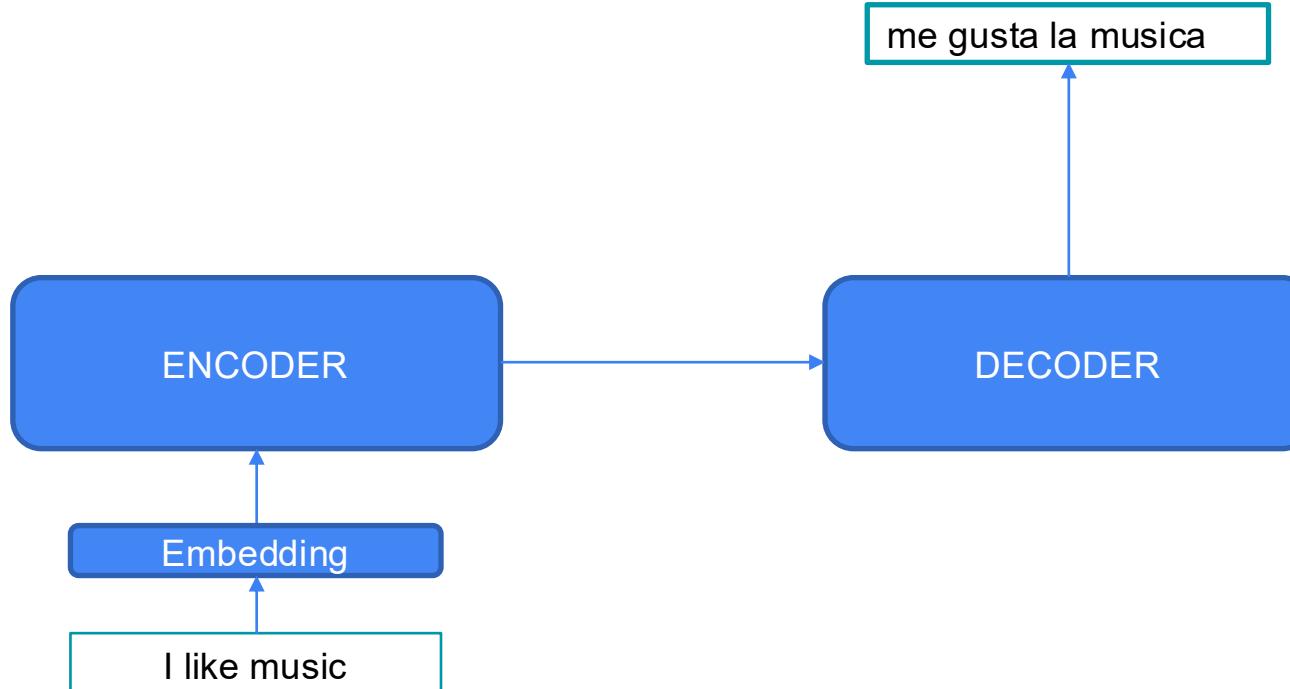
IIT Kharagpur

# Generative AI

# How Does a Generative AI System Work?

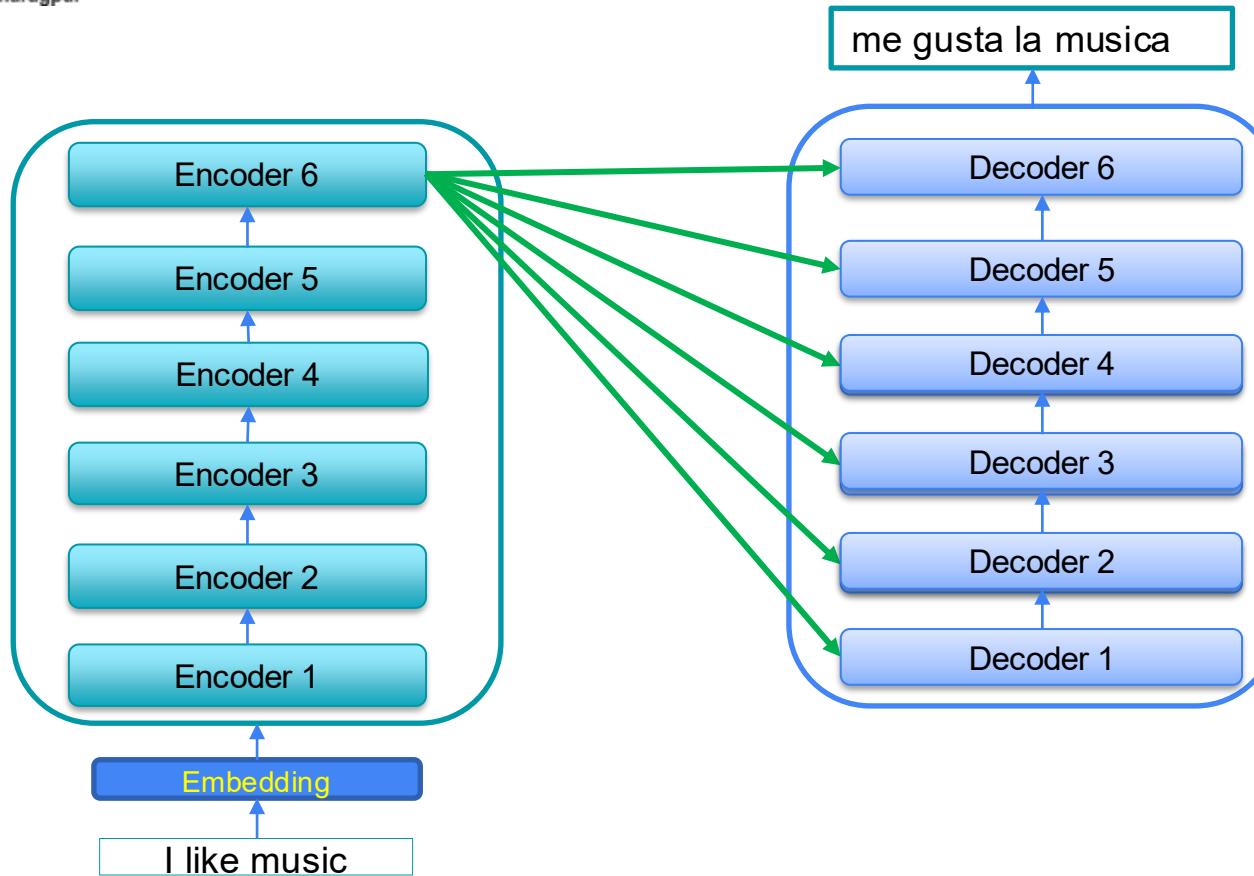


# Transformer Architecture





# Transformer Architecture



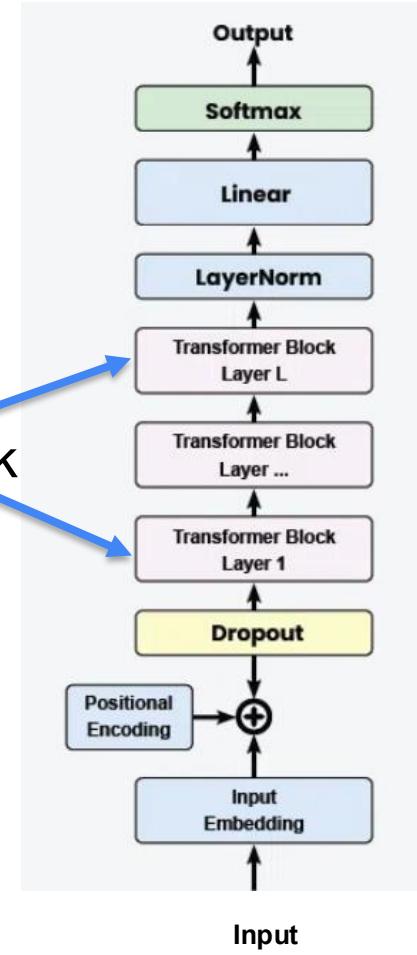


AI4ICPS

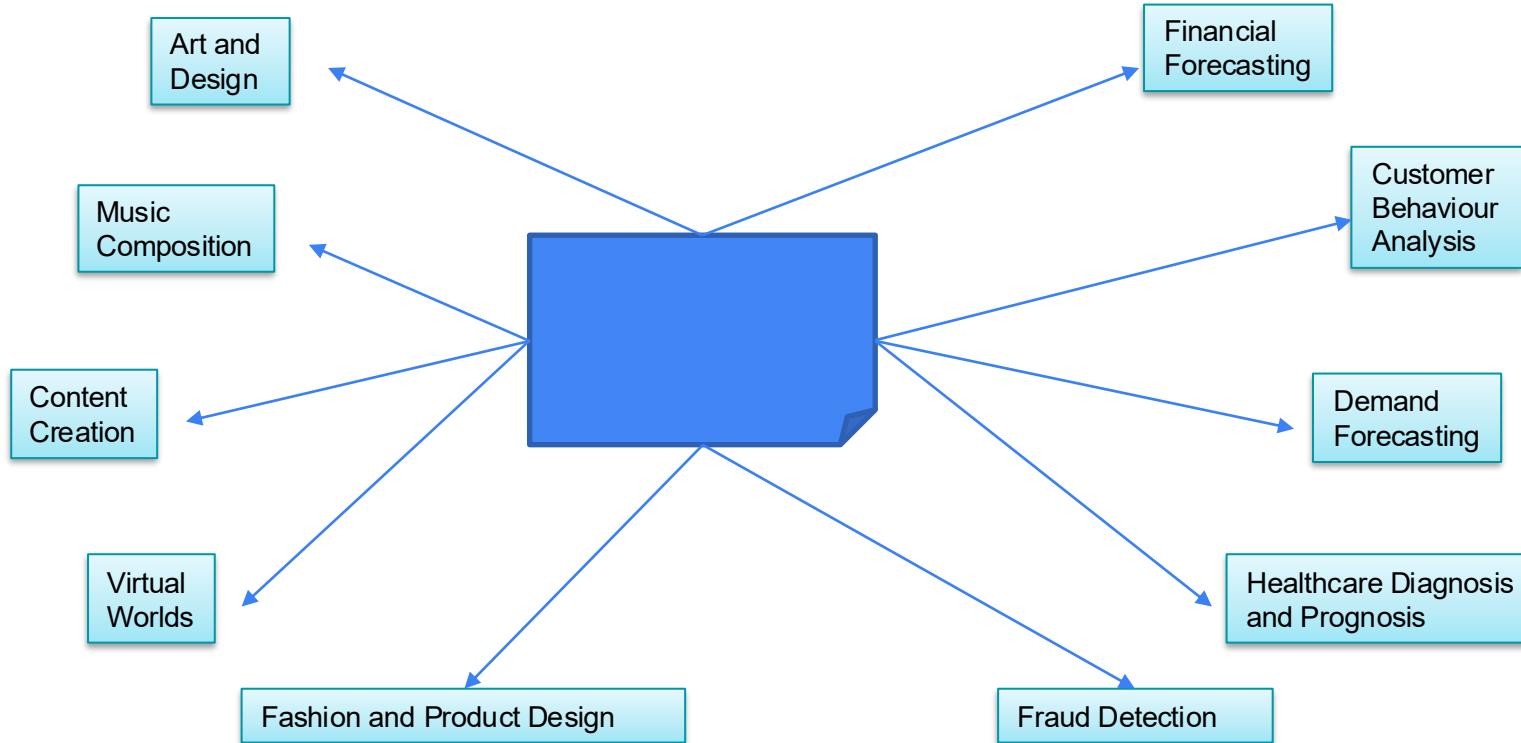


# GPT Architecture

Several Transformer Block



# Generative AI: Applications





AI4ICPS



# Generative AI Systems Show “Emergent” Behavior

$2 \Rightarrow 4$

$3 \Rightarrow 9$

$4 \Rightarrow 16$

$25 \Rightarrow$

In context learning

## Standard Prompting

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The answer is 27.

## Chain-of-Thought Prompting

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls.  $5 + 6 = 11$ . The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had  $23 - 20 = 3$ . They bought 6 more apples, so they have  $3 + 6 = 9$ . The answer is 9.

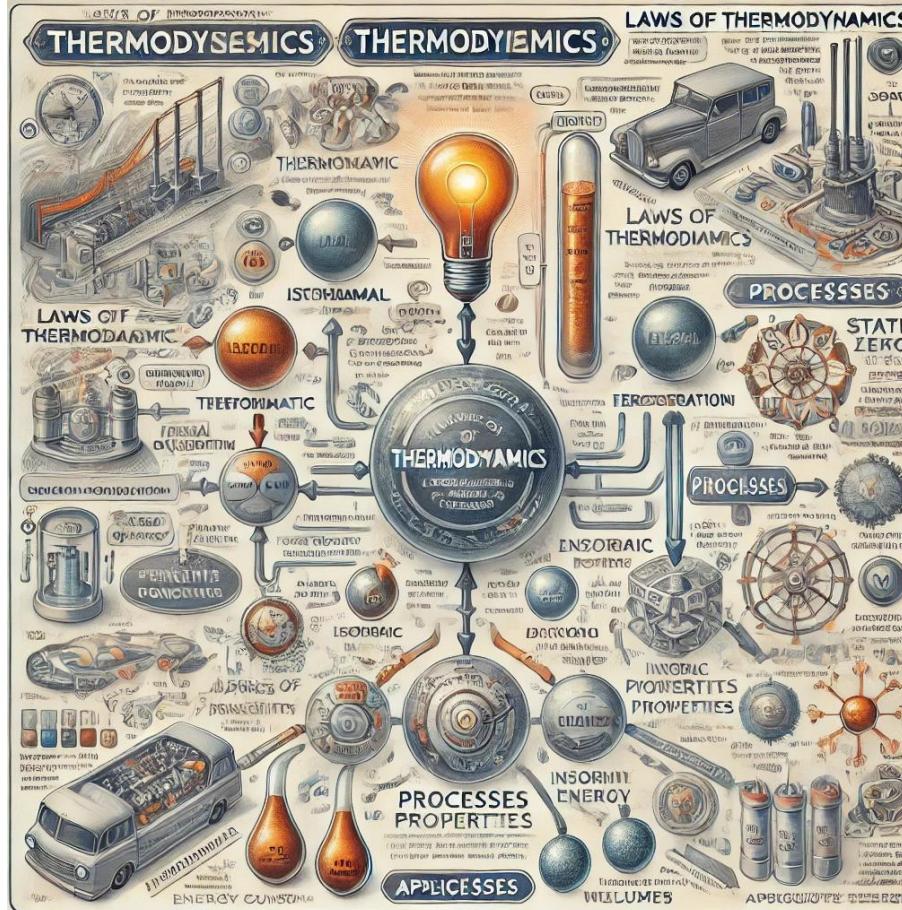
Prompting



AI4ICPS



# Lesson Planning



## Create a concept map of Thermodynamics



AI4ICPS

IIT Kharagpur

# Varied Example for Concept Reinforcement

Imagine that you are a teacher planning to explain the concept pseudo force to a 9th grade student. Give five example scenarios for pseudo force in five different paragraphs that will explain different aspects of pseudo force

## 1. Sudden Braking in a Car

Imagine sitting in a car when it suddenly comes to a halt. Although the car stops, you feel like are thrown forward. From the perspective of someone inside the car (a non-inertial reference frame), there seems to be a "force" pushing you forward. This force is the **pseudo force**. However, from the perspective of someone standing outside the car (an inertial frame), it's just inertia keeping your body in motion as the car stops. This explains how pseudo force appears due to acceleration in non-inertial frames.

## 2. Rotating Merry-Go-Round

On a spinning merry-go-round, you feel as though you're being flung outward. To you, a force seems to pull you away from the center. This "force" is called the **centrifugal pseudo force**. In reality, what's happening is that your body tries to maintain a straight-line motion (inertia), while the merry-go-round pushes you toward the center (centripetal force). The pseudo force arises in your reference frame of reference to balance this motion.

## 3. Elevator Acceleration

When an elevator starts moving upward, you feel heavier as if a force is pushing you down. Conversely, when it accelerates downward, you feel lighter. In both cases, these are pseudo forces experienced due to the elevator's acceleration. For someone outside the elevator (in an inertial frame), these sensations can be explained by changes in normal force and gravity.

## 4. Tilting Train

Imagine standing in a train as it makes a sharp turn. You feel as though you're being pushed to one side. This sideways "force" is a pseudo force because, in reality, your body is trying to continue in a straight-line motion due to inertia. The curved motion of the train creates the sensation of a force acting in the opposite direction.

## 5. Earth's Rotation

On the Earth's surface, we experience a pseudo force called the **Coriolis force** due to the Earth's rotation. For instance, winds and ocean currents do not move straight but are deflected (to the right in the Northern Hemisphere and to the left in the Southern Hemisphere). This deflection can only be explained using the concept of pseudo force in the non-inertial frame of reference of the rotating Earth.

# Provide Multiple Levels of Explanation

I am going to introduce cell division concept to a 8th grader. They will encounter the concept for the first time. Keep the description simple, engaging and try to use analogy from daily experience.

## Introducing Cell Division: "The Magic of Making Copies"

Imagine you have a favorite book, and you want to share it with a friend. But instead of giving your friend the original, you make an exact copy of the book so you both can have one. Now, imagine this happens over and over again until everyone who wants the book has their own copy. This process of copying is similar to what happens during **cell division** in your body.

## What is Cell Division?

Cells are like tiny building blocks that make up everything in your body—your skin, muscles, bones, and even your blood. To help your body grow, repair itself, or replace old and damaged cells, these building blocks need to multiply. **Cell division** is the process where one cell splits into two, creating two "daughter" cells that are just like the original.

## Analogy: A Factory Making New Workers

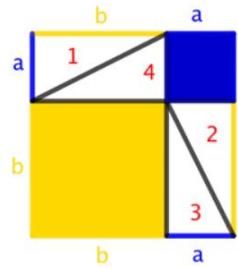
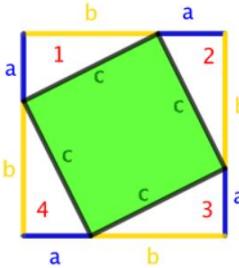
Think of a cell as a worker in a factory. When the factory needs more workers, one worker divides into two. Both new workers know exactly how to do the job because they carry the same "instructions" (DNA) as the original. This ensures the factory (your body) keeps running smoothly.



AI4ICPS



# Provide Multiple Ways to Solve a Problem



## Proof 1: Algebraic (Using Area Decomposition)

1. Construct a square with side length  $a + b$ .
2. Inside the square, arrange four identical right triangles such that their hypotenuses form a smaller square in the center.
3. The area of the large square is  $(a + b)^2$ , and the area of the smaller square is  $c^2$  (the square formed by the hypotenuses).
4. The total area can also be expressed as the sum of the areas of the four triangles:

$$(a + b)^2 =$$

1. Construct two identical large squares, each with side  $a + b$ .

5. Simplify to get:

$$a^2 + 2ab + b^2 = 2ab$$

2. In one square, place four identical right triangles in one arrangement, leaving a smaller square with side  $c$  in the center.

3. In the second square, rearrange the triangles to leave two rectangles with areas  $a^2$  and  $b^2$ .

4. Since the overall area is preserved, equate:

$$c^2 = a^2 + b^2$$

I will like to show the students that there can be multiple proofs for a geometry problem. Provide at least three alternative proofs for the Pythagoras theorem. Also discuss the merits and demerits of the generated proofs.

## Proof 2: Geometric (Using Similar Triangles)

1. In a right triangle, draw an altitude from the right angle to the hypotenuse, dividing it into two segments.
2. The two smaller triangles are similar to the original triangle and to each other (by AA similarity).
3. Let the segments of the hypotenuse be  $d$  and  $e$ , with the hypotenuse  $c = d + e$ . The altitude is  $h$ .

$$h^2 = ab$$

$$h^2 = bd$$

$$c^2$$

## Proof 3: Rearrangement (Using Visual Puzzle)



AI4ICPS



IIT Kharagpur

# Agentic AI

# Agentic AI





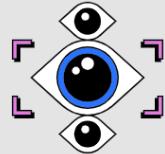
AI4ICPS



IIT Kharagpur

Many hits  
Yet, Many Misses  
Still miles to go .....

# What we need to build true Intelligent system?



Visual perception



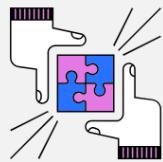
Audio perception



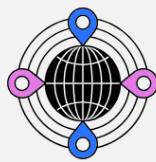
Fine motor skills



Natural language processing



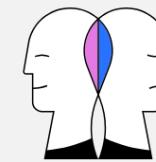
Problem-solving



Navigation



Creativity



Social and emotional engagement

ALL IN ONE,  
SINGLE DEVICE

# Safety & ML Algorithm Bias

AI4ICPS

IIT Kharagpur

Tesla hit parked police car 'while using Autopilot'

© 30 May 2018

f    Share



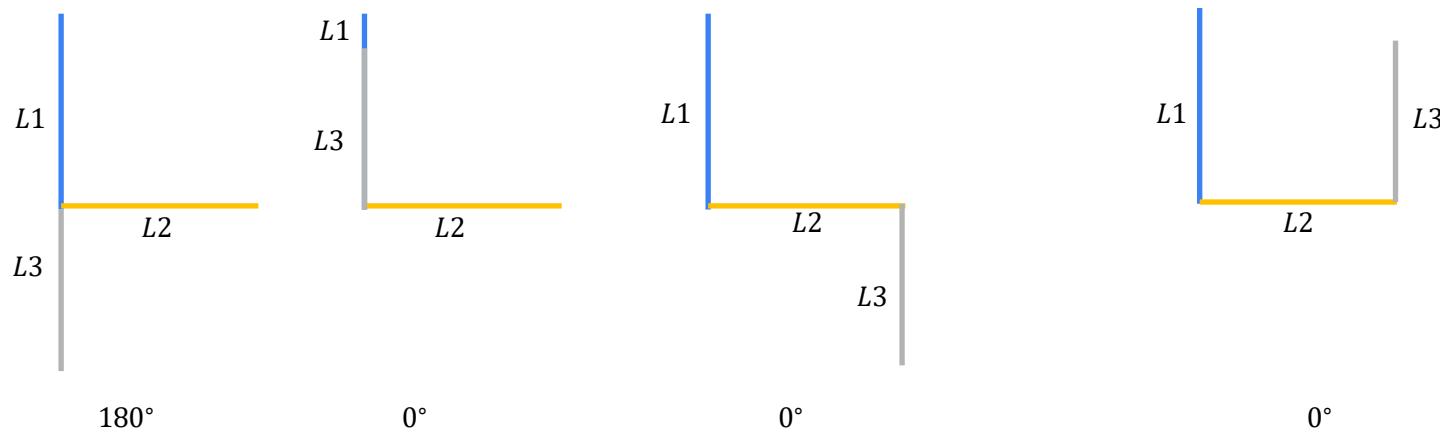
Bias in Algorithms: Accuracy of predictions

**98.7%** **68.6%** **100%** **92.9%**



# Assessing Students' Learning

**Response to question:** In a two-dimensional plane, two lines L1 and L2 are perpendicular to each other. The lines L2 and L3 are mutually perpendicular. Find the angle between L1 and L3, the answer of the student is 180 degrees. The reasoning the student has provided is that if the line L2 is perpendicular to L1 and L3 is perpendicular to L2, then L1 and L3 are co-linear. Hence, L1 and L3 would form an angle of 180 degree. Check whether the answer of the student is correct. Otherwise, specify why the student's answer is incorrect.





AI4CPSS



IIT Kharagpur

# Assessing Answer

The student's answer is **incorrect**, as their reasoning is based on a misunderstanding of the geometric relationships between perpendicular lines in a two-dimensional plane. Here's why:

## Why the Student's Reasoning is Incorrect

### 1. Understanding the Relationship Between the Lines:

- In a two-dimensional plane, if  $L_2$  is perpendicular to  $L_1$ , then  $L_2$  lies at a  $90^\circ$  angle to  $L_1$ . Similarly, if  $L_3$  is perpendicular to  $L_2$ , then  $L_3$  lies at a  $90^\circ$  angle to  $L_2$ .
- This setup forms an orthogonal relationship:  $L_3$  is perpendicular to  $L_2$ , and  $L_2$  is perpendicular to  $L_1$ . Thus,  $L_1$  and  $L_3$  will also be **perpendicular to each other**, forming a  $90^\circ$  angle.

### 2. Why $L_1$ and $L_3$ Are Not Co-linear:

- The claim that  $L_1$  and  $L_3$  are co-linear is incorrect. Two lines are co-linear only if they lie along the same straight path or can be described by the same line equation. In this case,  $L_1$  and  $L_3$  are at  $90^\circ$  to each other and do not lie along the same line.
- The student confuses the property of "being perpendicular to a common line" with "co-linearity." Perpendicularity does not imply co-linearity.

## Correct Answer

- The angle between  $L_1$  and  $L_3$  is **90 degrees**, not 180 degrees.

AI Systems Hallucinate

# AI Systems Hallucinate

Prove that the equation  $1 - x^n - (1 - x^n) = 0$  for  $n \geq 0$  has exactly 10 solutions and list the solutions.

## Conclusion

The equation  $x^n + (1 - x)^n = 1$  has exactly  $n$  solutions in the interval  $[0, 1]$ , and for  $n = 10$ , there are precisely 10 solutions, symmetrically distributed around  $x = 0.5$ .

Prove that the equation  $1 - x^n - (1 - x^n) = 0$  for  $n \geq 0$  does not have more than 2 solutions

## Step 4: Conclusion

For  $n \geq 2$ , the equation  $1 - x^n - (1 - x)^n = 0$  has at most two solutions in  $[0, 1]$ . These solutions are symmetric about  $x = 0.5$ .

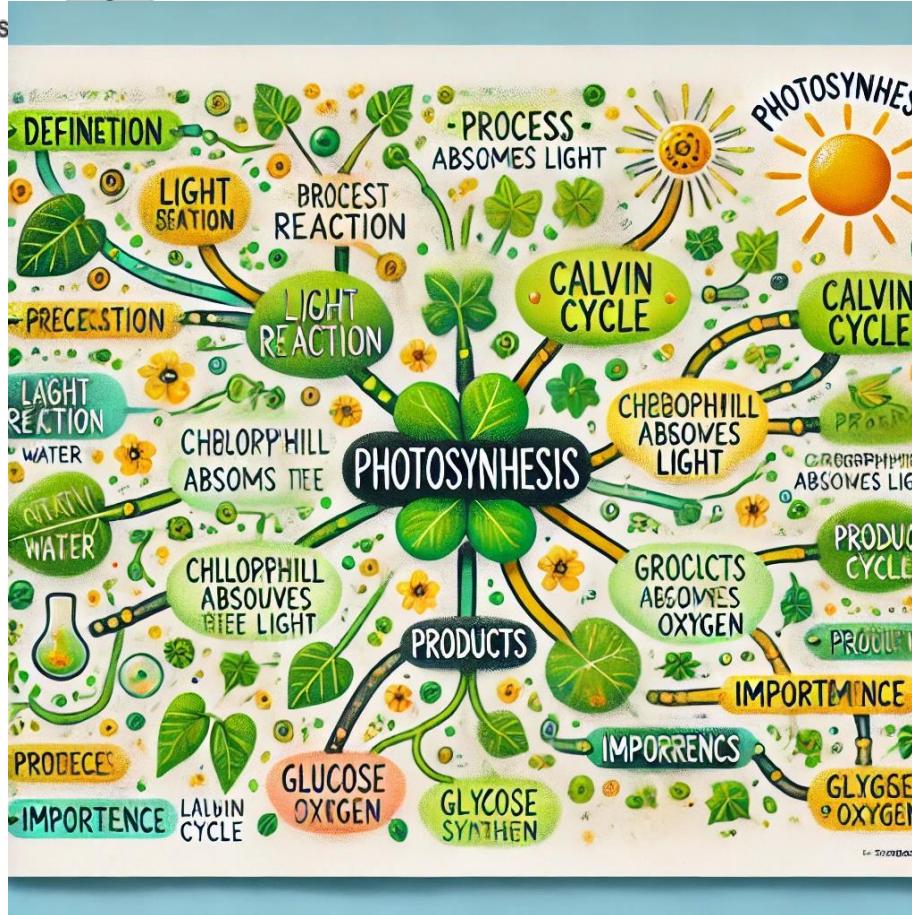
Prove that the equation  $1 - x^n - (1 - x^n) = 0$  for  $n \geq 0$  has exactly 2 solutions

## Step 5: Conclusion

The equation  $x^n + (1 - x)^n = 1$  has exactly **two solutions** in the interval  $[0, 1]$  for  $n \geq 0$ . These solutions are symmetric about  $x = 0.5$ .



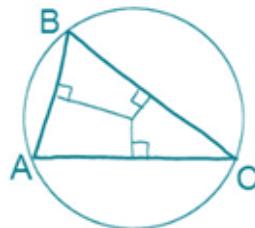
# Lesson Planning



Very attractive but gibberish



AI4ICPS



A circle passes through the points A( -1,5), B(7, 1), and C(5, -3).  
Find the equation of the circle

## Human Solution

- 1) Find midpoint of AB
- 2) Find slope of AB
- 3) Slope of perpendicular bisector of AB
- 4) Equation of perpendicular bisector
- 5) Midpoint of BC
- 6) Slope of BC
- 7) Slope of perpendicular bisector of BC
- 8) Equation of perpendicular bisector
- 9) Circumcenter is where the two bisectors meet
- 10) Find center of circle
- 11) Find radius
- 12) Equation:  $(x - 2)^2 + (y - 1)^2 = 25$

## AI Solution

- 1) General equation of a circle  
$$x^2 + y^2 + Dx + Ey + F = 0$$
- 2) Generate three equations corresponding to A( -1,5), B(7, 1), and C(5, -3)
- 3) Solve the equation to find out D, E and F
- 4) Equation:  $x^2 + y^2 - 4x - 2y - 20 = 0$

Your focus is more to develop conceptual understanding than devising a solution

Which method will you choose?



AI4ICPS



IIT Kharagpur

# Thank you!