

Language Models

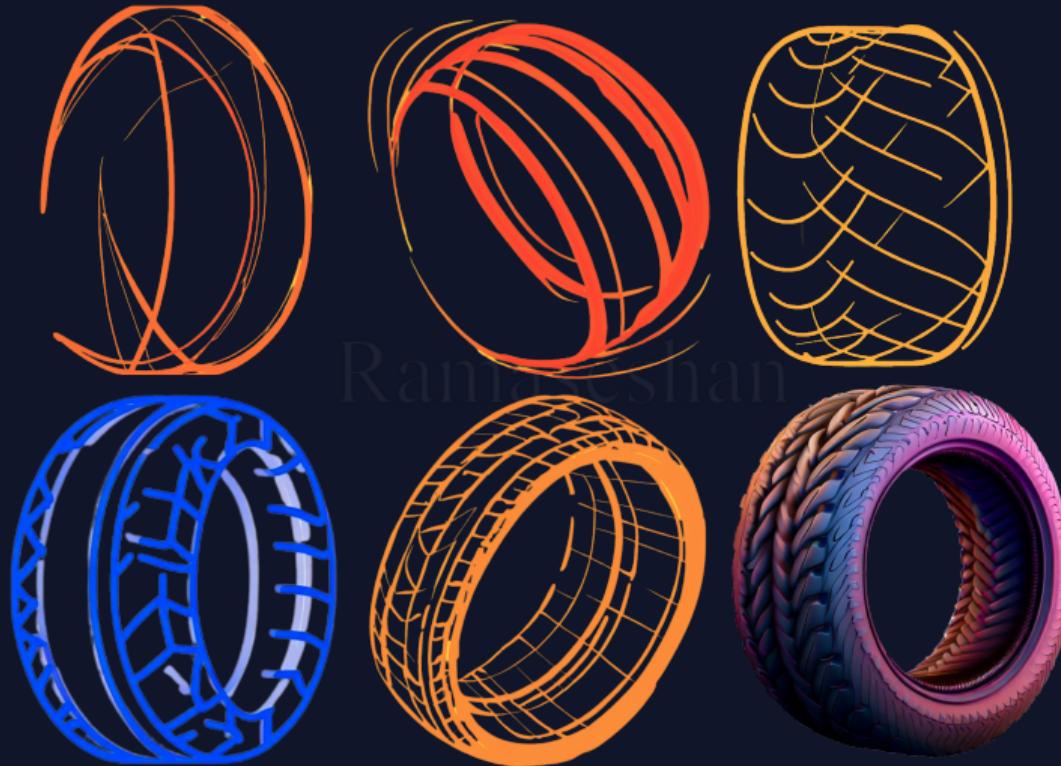
Ramaseshan Ramachandran

Models
Physical Models
Generative AI

Generative AI
Symbolic Reasoning
Probabilistic Reasoning

Ramaseshan

MODELS



WHAT ARE MODELS?

- ▶ A simplified representation of a complex system or phenomenon
- ▶ Designed to understand, analyze, and make predictions about the real world
- ▶ Take different forms, from mathematical equations to physical replicas
- ▶ Consist of several components

Governing equations

Constraints

Variables

Assumptions

Parameters

Ramaseshan

- ▶ Represent encoded information collected from observed data
- ▶ Facilitates prediction of future events
- ▶ Possible next word in a sentence
 - Lexical Knowledge
- ▶ Validates a sentence using learned parameters
- ▶ Representation of words as vectors
- ▶ Understanding word senses
 - Semantic Knowledge
- ▶ Anaphora detection
- ▶ Unsupervised and Supervised
- ▶ No fixed procedural logic
- ▶ Learns input-output relationships

PROPERTIES OF A MODEL

- ▶ Uncovers latent patterns that is difficult to visualize using rigid procedural method
- ▶ Observations of a pattern → **epistemological** understanding of that pattern
 - ▶ Information related correlations
 - ▶ n-gram frequencies, lexical and semantic relationships, word representations as vectors
 - ▶ Incidence matrix of n-grams
- ▶ Does not use any schema and no human annotation
- ▶ Helps in understanding and interpreting information for decision making/classification

Ramaseshan

PROPERTIES OF A MODEL

- ▶ Allows open class queries
- ▶ Makes data driven decisions under uncertainty
- ▶ Presents Inferences based on the domain knowledge - conditional probability
- ▶ Stores vast amount of linguistic and semantic information as parameters knowledge
- ▶ Accesses the knowledge by using the context

Words on the principle of object permanence

Ramaseshan

PORTRAIT OF EDMOND BELAMY



Tribute to Goodfellow - creator of GAN

First artwork created using AI

Featured in a Christie's auction

Signed at the bottom right with generative
Adversarial model loss function

$$\min_G \max_D [E_x[\log(D(x))] + E_y[\log(1 - D(G(y)))]]$$

Trained on a set of 15,000 portraits

- ▶ Learns the patterns of existing content
- ▶ Creates new content such as text, images, or music using learned patterns
- ▶ Generates realistic text, images, and music
- ▶ Hallucinates on unknown areas with known keywords - Generates grammatically correct sentence but factually incorrect sentence like *The cat barks*

- ▶ Customer Support Automation
 - ▶ Idea Generation and Brainstorming
 - ▶ Content Generation
 - ▶ Language Translation
 - ▶ Writing Assistance and Editing
 - ▶ Programming - writing and debugging
 - ▶ Learning and Education
- Ramasesha
- ▶ Data Analysis
 - ▶ Conversational Agents and Virtual Assistants
 - ▶ Mental Health Support
 - ▶ Research and Information Retrieval
 - ▶ Virtual Training
 - ▶ Personal Productivity
 - ▶ Innovation and Creativity

GAI - FAKE INTERVIEW



Ram is a professional surfer. I want the sports reporter Rob to interview him about his professional life as a surfer. Rob is allowed to ask only 3 questions. This is a rapid-fire question and answer session

Rob: "Welcome, Ram! Let's dive right in. Question 1: What inspired you to become a professional surfer?"

Ram: "Growing up by the ocean, the waves always intrigued me, and I wanted to challenge myself to master them."

Rob: "Great! Question 2: Could you share your most memorable surfing moment so far?"

Ram: "Definitely catching a massive wave during a competition in Hawaii and riding it flawlessly to victory."

Rob: "Impressive! Last question, Ram. Question 3: What advice do you have for aspiring surfers looking to make it big?"

Ram: "Stay dedicated, practice consistently, and don't be afraid to take risks. Keep pushing your limits."

Rob: "Thank you, Ram, for the insights into your incredible journey as a professional surfer. We appreciate your time!"

Video
Ramaseshan

PYTHON CODE GENERATION

Could you give me a Python program to add the first 100 Fibonacci numbers?

```
def fibonacci(n):
    fib_sequence = [0, 1] # Initialize with the first two Fibonacci numbers
    while len(fib_sequence) < n:
        next_num = fib_sequence[-1] + fib_sequence[-2]
        fib_sequence.append(next_num)
    return fib_sequence

# Calculate the sum of the first 100 Fibonacci numbers
first_100_fibs = fibonacci(100)
sum_first_100_fibs = sum(first_100_fibs)

print("The first 100 Fibonacci numbers:", first_100_fibs)
print("Sum of the first 100 Fibonacci numbers:", sum_first_100_fibs)
```

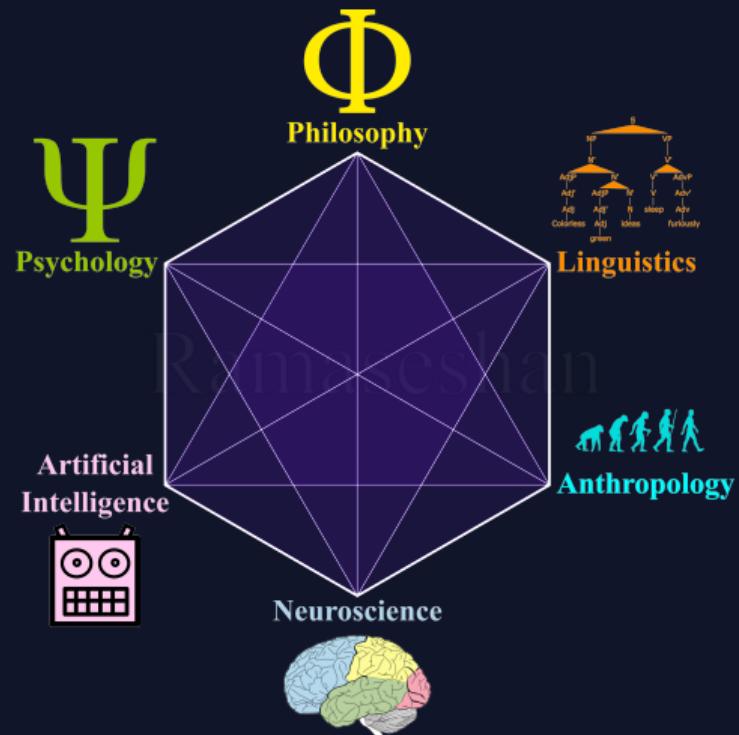
Mimic Human Intelligence - Replicate human-like cognitive functions

Data-Driven Learning - Learn from vast amounts of data, extracting patterns and insights

Adaptive Behavior - Understand context and respond, make decisions

Problem Solving - Read, analyze and solve complex problems, explore potential solutions, and make informed decisions based on available information

- ▶ **Perception** - ability to receive and evaluate information about an environment
- ▶ **Learning** - ability to find common patterns, organize the knowledge
 - Represent knowledge using Propositional and first-order logic - symbolic representation
 - Parameters using statistics and probability - probabilistic representation
- ▶ **Reasoning** - Draw conclusions from the knowledge and learnings such as facts, beliefs, observations/evidence, logical rules
- ▶ Use available/partial information for problem solving
- ▶ Interpretability of reasoning
- ▶ Planning and Execution



- ▶ Knowledge is available as types, relations and their complex relationship - represented in the form of first-order and propositional logic

Who is an all rounder in cricket?

$$\exists x(\text{Bats well}(x) \wedge \text{Fields well}(x) \wedge \text{Bowls well}(x)) \Rightarrow \text{All Rounder}(x)$$

The facts are hard facts - there could be no uncertainty

- ▶ Knowledge is available in the form of parametric relationships - co-occurrences, correlations, incidence matrix, conditional relationships
- ▶ Handles uncertainty well using the observed parameters
- ▶ $\exists w_i \exists w_j (\text{Similar Contexts}(x) \wedge \text{Similar Contexts}(y)) \rightsquigarrow \text{Similar}(x, y)$
- ▶ How are related in a corpus?
- ▶ Context words may differ and not fixed

EXPECTATIONS OF A LANGUAGE MODEL

Find the missing words: Are you _____ in playing cricket?

Rephrase the sentence: You went to a movie yesterday. How was the movie?

Find semantic error in a sentence: - Do you walk a car?

Identify the context: Where is the program?

Generate a sentence: The starting prompt is An adiabatic

...