Fast-Tracking course of AI

From "What is AI?" to "How ChatGPT works”.

What does it mean to learn?

What does it mean to learn?

Changing yourself based on experiences, basically adjustments

Learning is a loop

Try something -> See what happens(feedback) -> Adjust -> Repeat

What is knowledge?

What does it mean to know?

What is knowledge?

What does it mean to know?

Two types of knowledge:

- Explicit Knowledge
- Implicit Knowledge

What is Intelligence?

What is Intelligence?

Intelligence is the ability to achieve goals in a wide range of situations.

What is AI?

What comes to mind when you hear "Artificial Intelligence"?

AI = The Goal

Making machines do things that would require intelligence if a human did them.

Examples

- Recognizing faces
- Understanding language
- Playing chess
- Driving a car

Attempt #1: Write the Rules

> "If the email contains 'free money' → spam"

> "If temperature > 100°F → fever"

> "If chess piece can move to square AND square has enemy piece → capture"

This is called "Expert Systems" or "Rule-Based AI"

The Problem with Rules

How would you write rules for tasks that humans do intuitively?

Recognizing a Face

How do you define the exact distance between eyes or the curve of a nose for every possible angle and lighting?

Understanding Sarcasm

"Oh, great." Does it mean something is good, or is it a complaint? Rules can't easily capture tone and context.

Decoding Idioms

"It's raining cats and dogs." A rule-based system might look for falling animals instead of understanding the metaphor.

Attempt #2: Let Machines Learn

Instead of writing rules…

Show the machine thousands of examples and let it figure out the patterns.

This is called Machine Learning (ML)

How Machine Learning Works (Intuitively)

01 Show the machine many examples

02 Machine makes a guess

03 Tell it if it was right or wrong

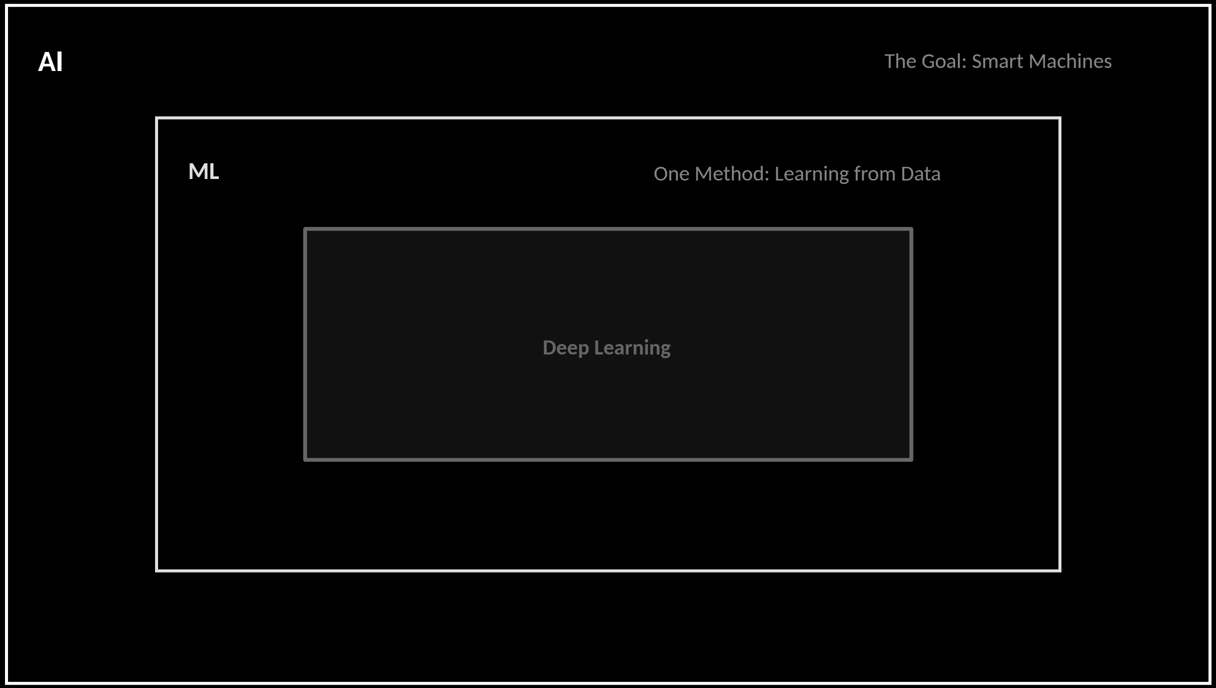
04 Machine adjusts itself slightly

05 Repeat millions of times

The Analogy

"Like a child learning by trial and error"

AI vs ML


AI
The Goal: Smart Machines
ML
One Method: Learning from Data
Deep Learning

Early ML: Modest Successes

What Worked

- Spam filters got better
- Recommendation systems (Netflix, Amazon)
- Basic image recognition

The Bottleneck

- Couldn't have a conversation
- Couldn't understand a paragraph
- Couldn't recognize objects as well as a toddler

Why Early ML Was Limited

Bottleneck 01

Not enough data

Machine Learning requires massive amounts of examples to learn patterns effectively.

Where do you get millions of labeled examples?

The internet was in its infancy.

Bottleneck 02

Not enough compute

Learning from data requires intense mathematical calculations.

Computers were too slow for deep training.

Training on millions of examples would take years.

AI Winters: When Hope Died (Twice)

1970s Winter

The First Collapse

Researchers promised human-level AI in 20 years. When they failed to deliver, funding dried up and interest vanished.

1980s–90s Winter

The Expert System Bust

Expert systems were supposed to revolutionize business, but they were too brittle for the real world. Funding vanished again.

"The field became a joke. Saying you worked on 'AI' was career suicide."

Why AI Exploded (After 2012)

01. Data

The Internet

Massive datasets of text, images, and video became available for the first time.

02. Compute

GPUs

Graphics cards, built for gaming, turned out to be perfect for the math ML needs.

03. Algorithms

Deep Learning

Researchers finally cracked how to train much deeper and more complex neural networks.

> All three ingredients converged around 2012.

Now Let's Go Deep on One Problem

We'll use **language understanding** as our lens.

Why Language is Hard for Computers

Why Language is Hard for Computers

> ***"I saw the man with the telescope."***
>
> Who had the telescope? Me or the man?

Computers need precision, but human language is **messy, ambiguous, and context-dependent**. A single sentence can have multiple valid interpretations, making it one of the hardest problems in AI.

Why Language is Hard for Computers

“Fast-tracking the AI course” and “Fast-tracking the course of AI”

NLP Attempt #1: The Dictionary Approach

Apple

/ˈap(ə)l/

1. The round fruit of a tree of the rose family, which typically has thin red or green skin and crisp flesh.

The Problem

A dictionary definition isn't enough for real-world language.

- "Apple" is also a trillion-dollar company.
- "Apple" is a famous record label.
- Context changes everything.

NLP Attempt #2: Statistical Patterns

> "If 'New' is followed by 'York' → City"

This approach powered Google Translate for nearly a decade.

The Problem

Pattern Matching ≠ Understanding

The machine could predict the next word based on frequency, but it had no concept of what the words actually meant.

The Breakthrough: Words as Numbers

"Apple" → [0.92, -0.14, 0.05, ...]

Computers don't understand words. They understand **numbers**. The challenge is: how do we turn a word into a list of numbers that captures its **meaning**?

Word Embeddings (The "Secret Sauce")

How do we turn a word like "Apple" into a number that captures its meaning?

Instead of one number, we give each word a **list of numbers** (a vector).

The Vector Representation

Word: "Apple"

[0.9 - color(red) , -0.1 shape(square), 0.05, ...]

Each number represents a specific "dimension" of meaning.

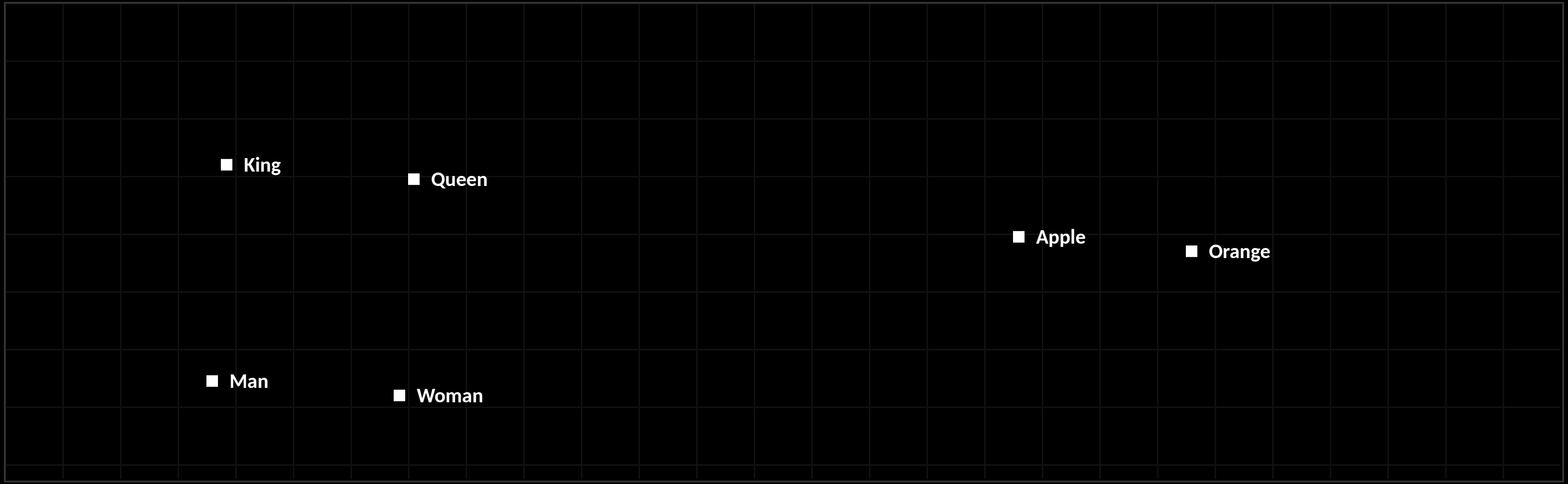
Dimensions of Meaning

Word	Royalty	Gender (M)	Edibility
King	0.98	0.95	0.01
Queen	0.97	0.05	0.02
Apple	0.02	0.00	0.94

"King" and "Queen" have **similar numbers for royalty** but vastly different values for gender.

This is how machines represent concepts as data.

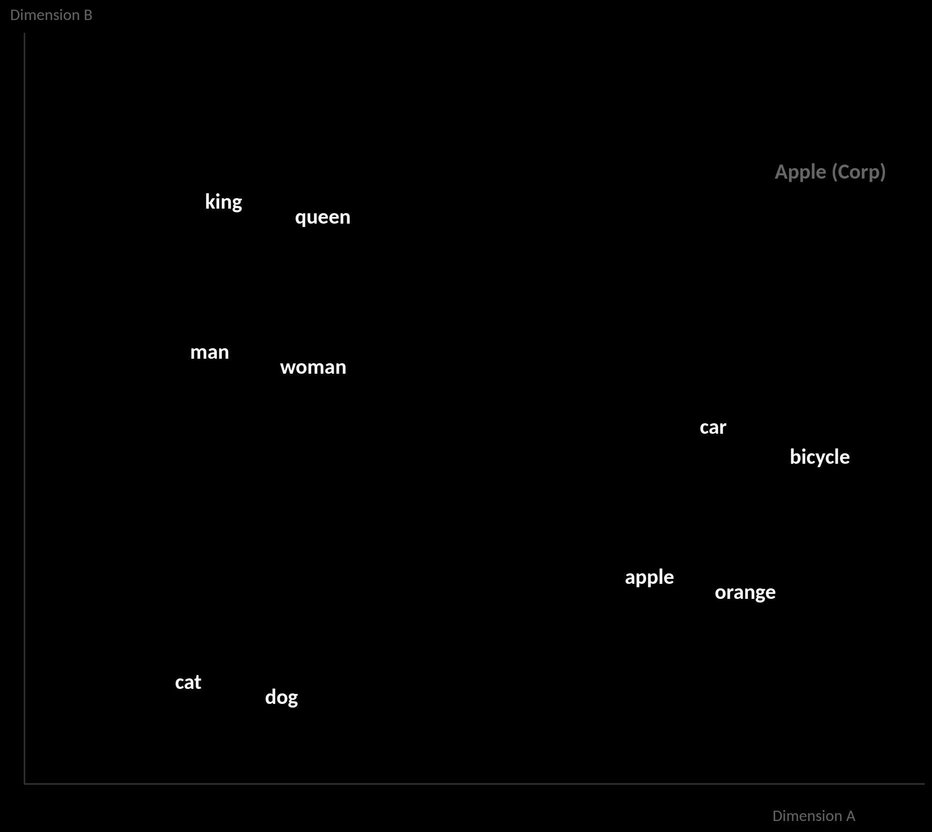
Words as Positions in Space



If a word is a list of numbers, it's a point on a map. **Similar words are close together**, while different words are far apart.

Visualizing Word Embeddings

Similar words are placed close together in a multi-dimensional space. Proximity indicates shared meaning.



The Magic of Embeddings — Word Math

King − Man + Woman = **Queen**

Paris − France + Italy = **Rome**

Walking − Walk + Swim = **Swimming**

Can Embeddings Solve Our Earlier Problems?

"I went to the **bank** to deposit cash"

"I sat on the **bank** of the river"

The Problem: In basic word embeddings, each word has exactly **ONE position** in space.

The model can't distinguish between a financial institution and a riverbank because it doesn't look at the surrounding words.

NLP Attempt #4: Sequence Models (RNNs)

Idea: Process the sentence one word at a time, building up understanding as you go.

> "I"
>
> "I went"
>
> "I went to"
>
> "I went to the"
>
> "I went to the bank..."

The model maintains a "memory" of what it has read so far.

The Problem: Forgetting

Long-Range Dependency Failure

> "**The cat**, which was sitting on the mat that I bought from the store near the old church on the corner, **was** happy."

By the time a sequence model reaches "was," it has often **forgotten** the subject at the beginning.

RNNs struggle with long sentences because they process information linearly and have limited memory capacity.

The Stage Is Set (around 2017)

What we had

- Word embeddings
- Sequence models (RNNs)
- Tons of internet data
- Powerful GPUs

The Problem

The Word-by-Word Bottleneck

Processing text sequentially was slow and models still "forgot" the beginning of long sentences.

What if there was a better way?

The Transformer Idea

The Old Way

Sequential

Process words one by one, like reading a book from start to finish.

The New Way

Simultaneous

Look at every word in the sentence at the same time.

The Secret: Attention

The model "attends" to the most relevant words, no matter how far apart they are.

Attention in Action

Processing the word "it"

The animal didn't cross the street because **it** was too tired.

When the model looks at the word **"it"**, the attention mechanism tells it to pay the most attention to **"animal"**.

This is how the model "knows" what the pronoun refers to, building a context-aware representation of every word.

Same Structure, Different Attention

Scenario A

The **animal** didn't cross the street because **it** was too ***tired***.

Scenario B

The animal didn't cross the **street** because **it** was too ***wide***.

By changing just one word (**tired** vs **wide**), the model's attention shifts, correctly identifying what "it" refers to in each context.

Why Attention Is Powerful

01. No Forgetting

More Memory

Every word in a sentence "sees" every other word simultaneously, no matter how far apart they are.

02. Speed

Parallel Processing

Unlike RNNs that read word-by-word, Transformers process all words at once, making training incredibly fast.

03. Context

Deep Understanding

The model builds a mathematical map of how every word relates to every other word in the specific context.

The Transformer

An architecture built entirely on the mechanism of attention.

Trait 01

No recurrence. No convolution. Just attention.

Trait 02

The foundation of all modern Large Language Models (LLMs).

Trait 03

Enables massive scaling of data and compute.

How ChatGPT Works (Simplified)

Transformer + Internet Data + Prediction

01. Architecture

The Engine

A massive neural network built entirely on the Attention mechanism we just explored.

02. Training Data

The Knowledge

Trillions of words from books, articles, and the entire public internet.

03. The Objective

The Task

A simple goal: given a sequence of words, predict the most likely next word.

Why "Predict the Next Word" Creates Understanding

Grammar & Syntax

Predicting the next word requires the model to internalize language structure and rules.

Factual Knowledge

Correct predictions rely on learned factual relations between concepts.

Logic & Reasoning

Following a chain of reasoning lets the model anticipate the appropriate continuation.

How It Generates Text

Step 01	**The**	
Step 02	The	**quick**
Step 03	The quick	**brown**
Step 04	The quick brown	**fox...**

Generation is an **iterative loop**. Each predicted word is added back to the input, and the model predicts the next word based on the updated context.

Let's See It Again

Revisiting the demo with a new perspective.

Look For 01

The iterative, word-by-word prediction process.

Look For 02

How attention builds deep context awareness.

Look For 03

How "reasoning" emerges from simple prediction.

The Unexpected Discovery: Bigger = Smarter

01. Data

More Tokens

Scaling from millions to trillions of words of training data.

02. Parameters

More Size

Scaling from millions to hundreds of billions of internal connections.

03. Compute

More Power

Scaling from days to months of training on thousands of GPUs.

The Result: Emergent Capabilities. Reasoning, coding, and logic appear spontaneously as models get larger.

Foundation Models

The Old Paradigm

Task-Specific AI

- Model A: Translation
- Model B: Summarization
- Model C: Sentiment Analysis

The New Paradigm

General-Purpose AI

One massive model trained on everything, capable of performing any language task through prompting.

The "Foundation" is the base knowledge. We no longer build tools from scratch; we build on top of these giants.

Where We Are (2024-2025)

Trend 01

Multimodality

AI is no longer just text. It can see images, hear voices, and speak back in real-time.

Trend 02

Reasoning

New models are designed to "think" before they speak, solving complex math and logic problems.

Trend 03

Agents

The shift from chatbots to agents that can use tools, browse the web, and complete multi-step tasks.

Questions?

The floor is open for discussion.