



How does AI work?

Dennis Pultz, Science Manager of AI

Bioinformatics & Design, Novozymes

DIFI seminar: How will AI affect patent work?

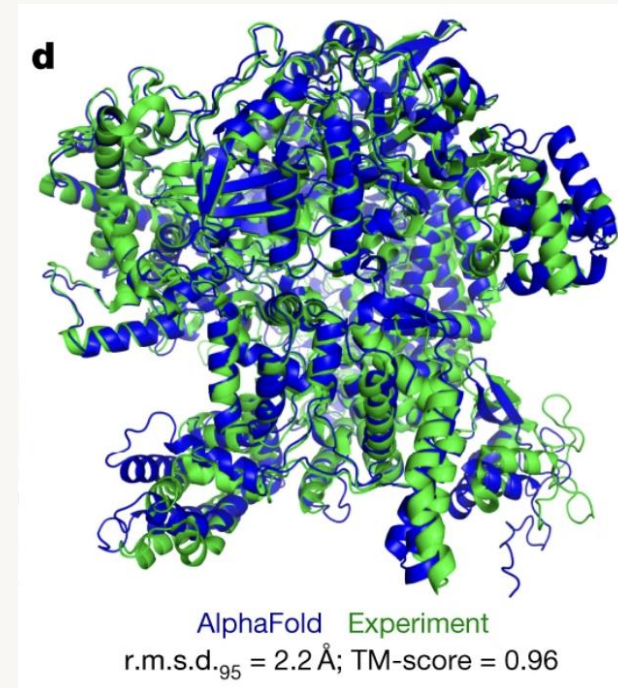
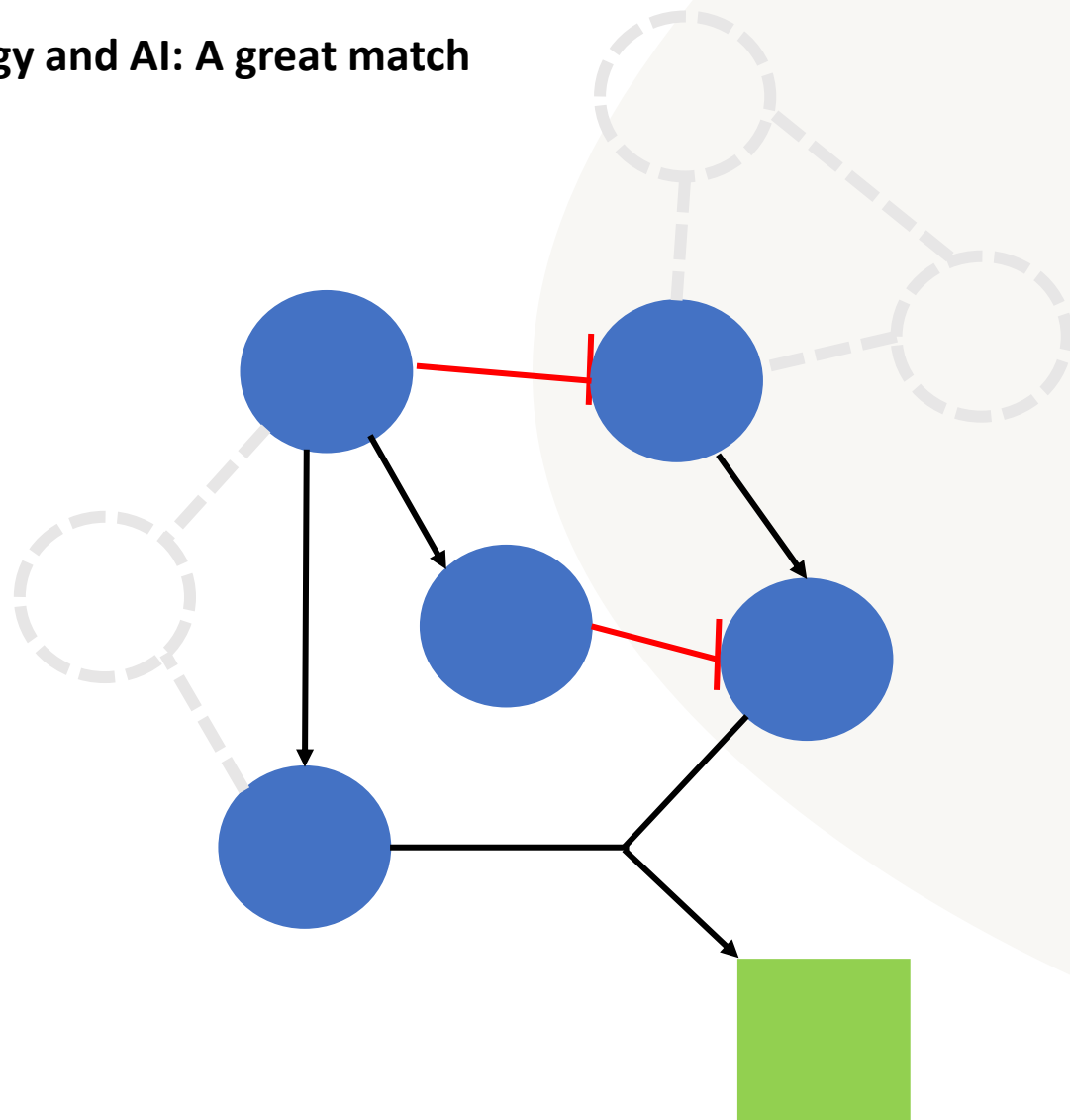
14th November, 2023

Rethink Tomorrow

novozymes[®] 

Introduction

- Biology and AI: A great match

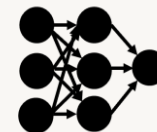


Agenda

- How does it work?



- Why Deep learning?



- What is chatGPT?



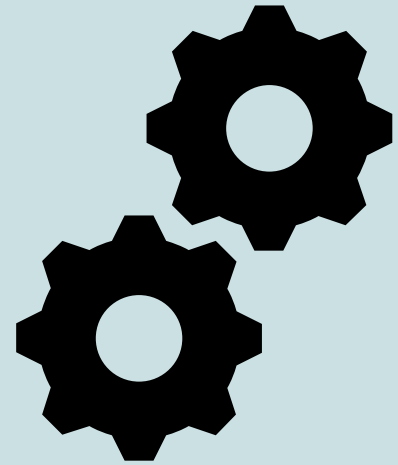
- AI Hype



- chatGPT limitations



HOW DOES IT
WORK?



How does it work: Classifying species

Iris versicolor

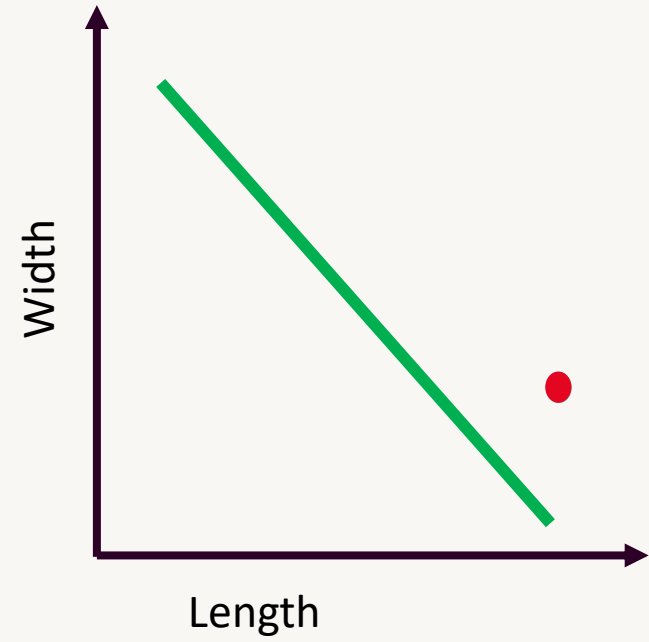
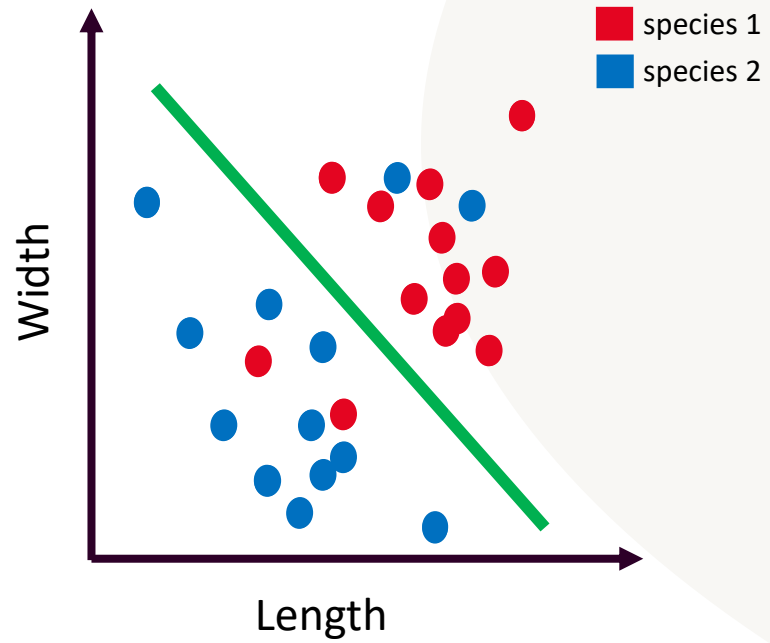


VS

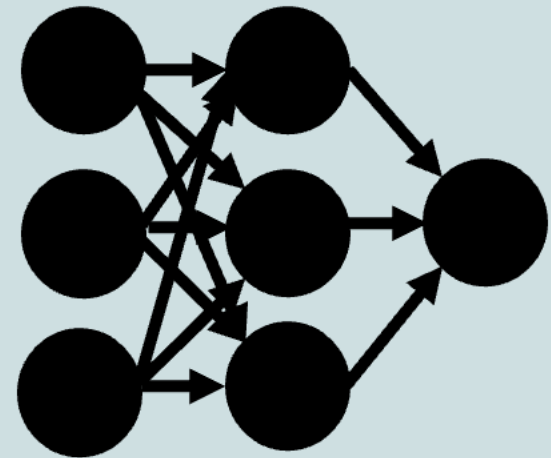
Iris setosa



Supervised Learning

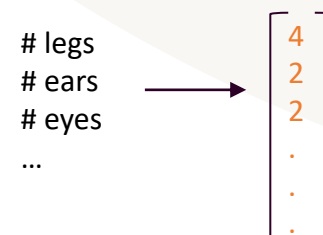
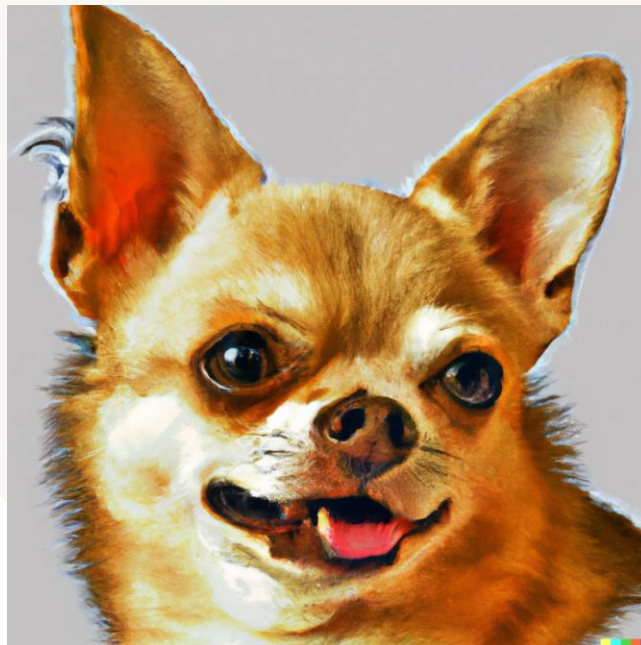


WHY DEEP LEARNING?



Chihuahua-detector

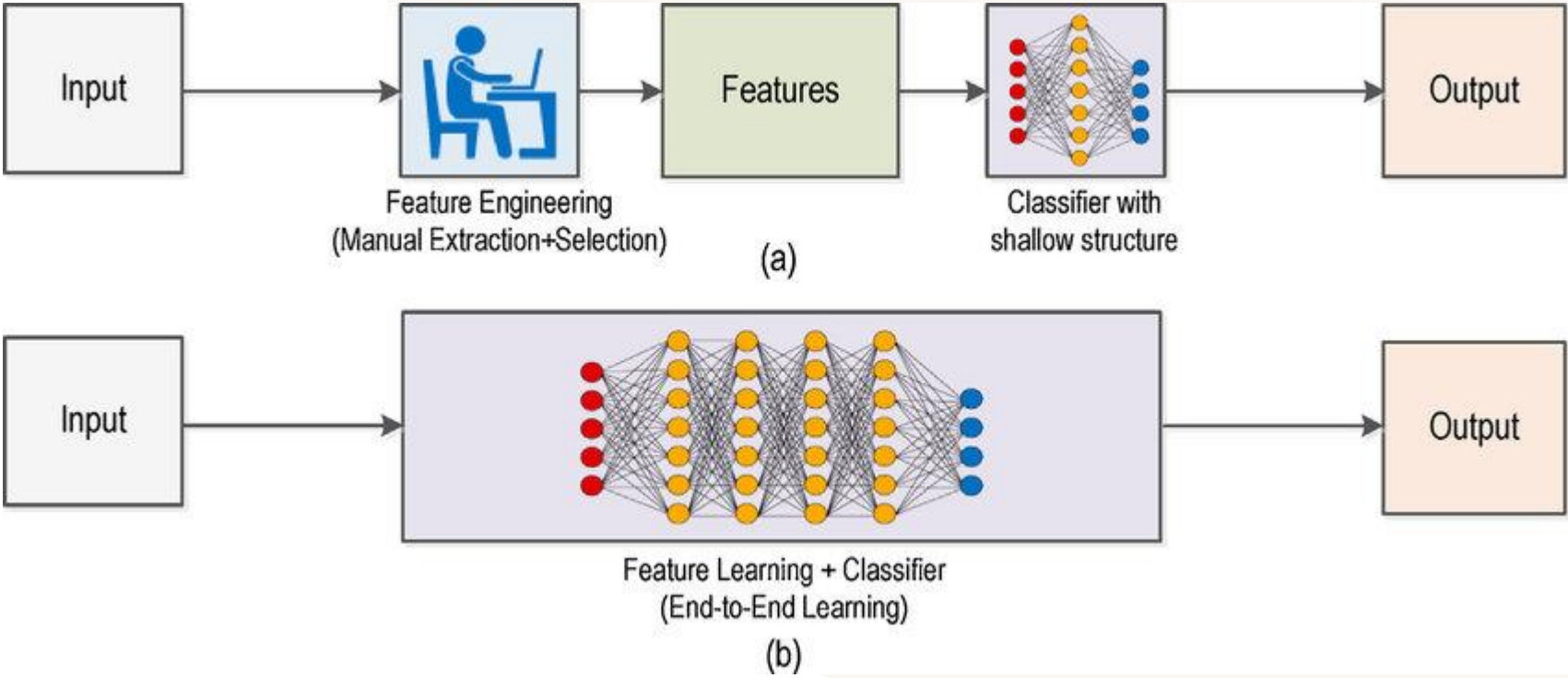
- Feature engineering



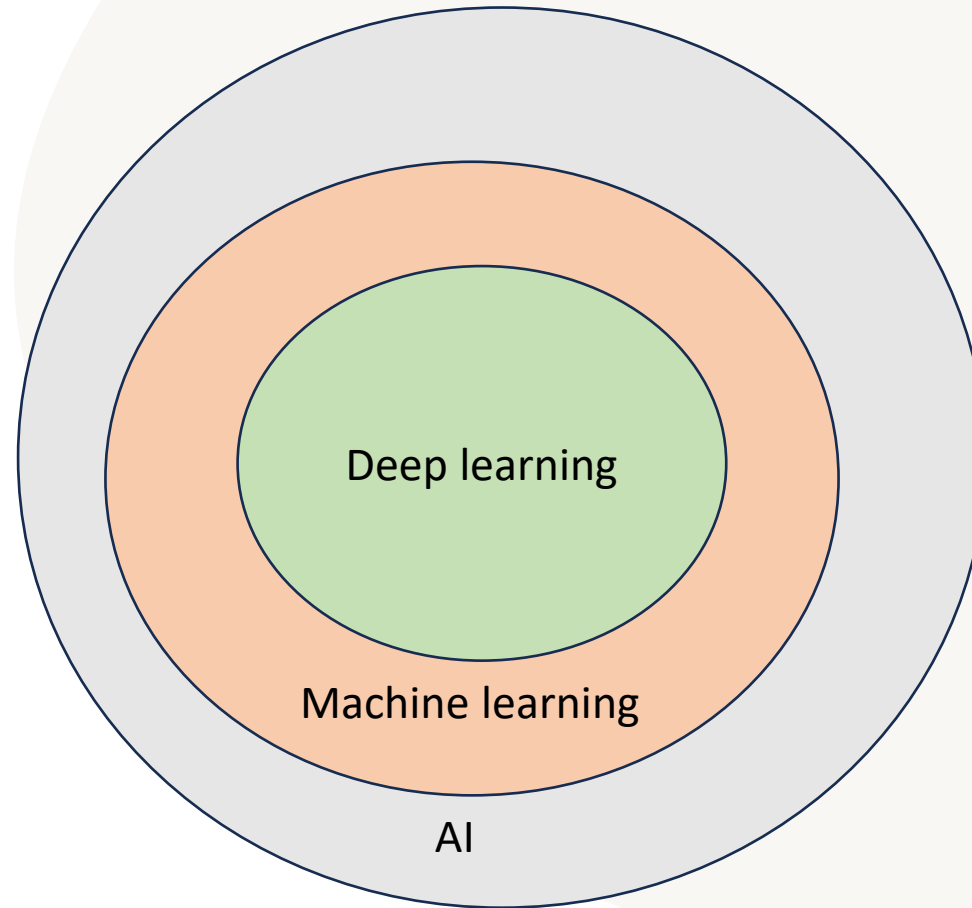
Chihuahua vs. muffin



Deep Learning



AI, Machine Learning... Deep Learning...???

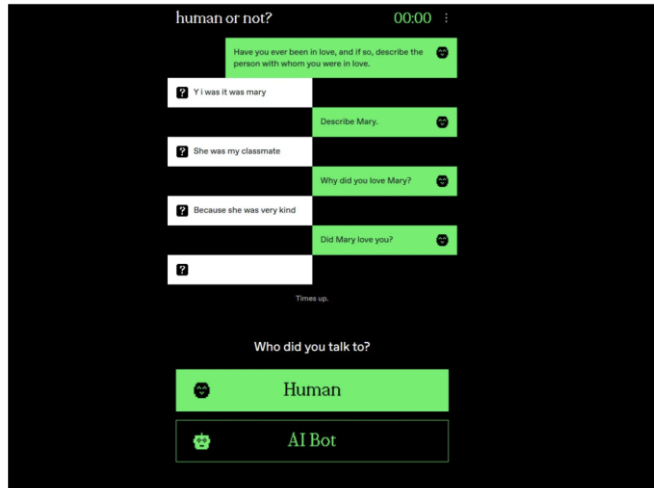


WHAT IS ChatGPT?



ChatGPT: Taking AI chatbots to the next level

Human or not?



AI21 Labs

Writing jokes



Sure thing! Let's get those giggles going:

1. Q: Why did the scarecrow become a successful stand-up comedian?
A: Because he was outstanding in his field!
2. Q: What do you call a fish wearing a crown?
A: A kingfish!
3. Q: Why did the math book look sad?
A: Because it had too many problems.
4. Q: How do you organize a space party?
A: You planet!
5. Q: What did one plate say to another plate?
A: Tonight, dinner's on me!

Coding



Certainly! In pandas, you can append data to a DataFrame using the `'append'` method. Here's a simple example:

```
python Copy code

import pandas as pd

# Create a DataFrame
df = pd.DataFrame({'Name': ['Alice', 'Bob'],
                  'Age': [25, 30]})

# Data to append
new_data = {'Name': 'Charlie', 'Age': 35}

# Append the new data to the DataFrame
df = df.append(new_data, ignore_index=True)

print(df)
```

Generally very useful, but not here:

pandas.DataFrame.append

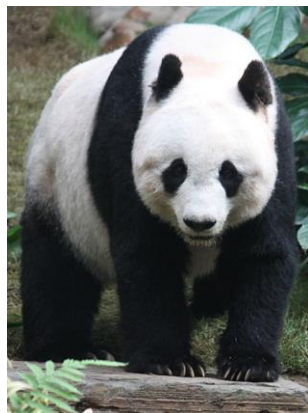
`DataFrame.append(other, ignore_index=False, verify_integrity=False, sort=False)`
Append rows of *other* to the end of caller, returning a new object. [\[source\]](#)

Deprecated since version 1.4.0: Use `concat()` instead. For further details see [Deprecated DataFrame.append and Series.append](#)

... information is out of date.

GPT: Generative Pre-trained Transformer

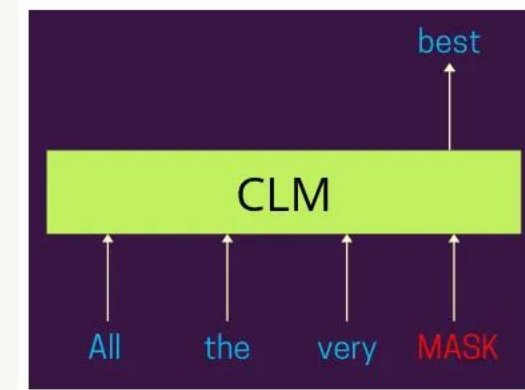
Discriminative Models



Generative Models



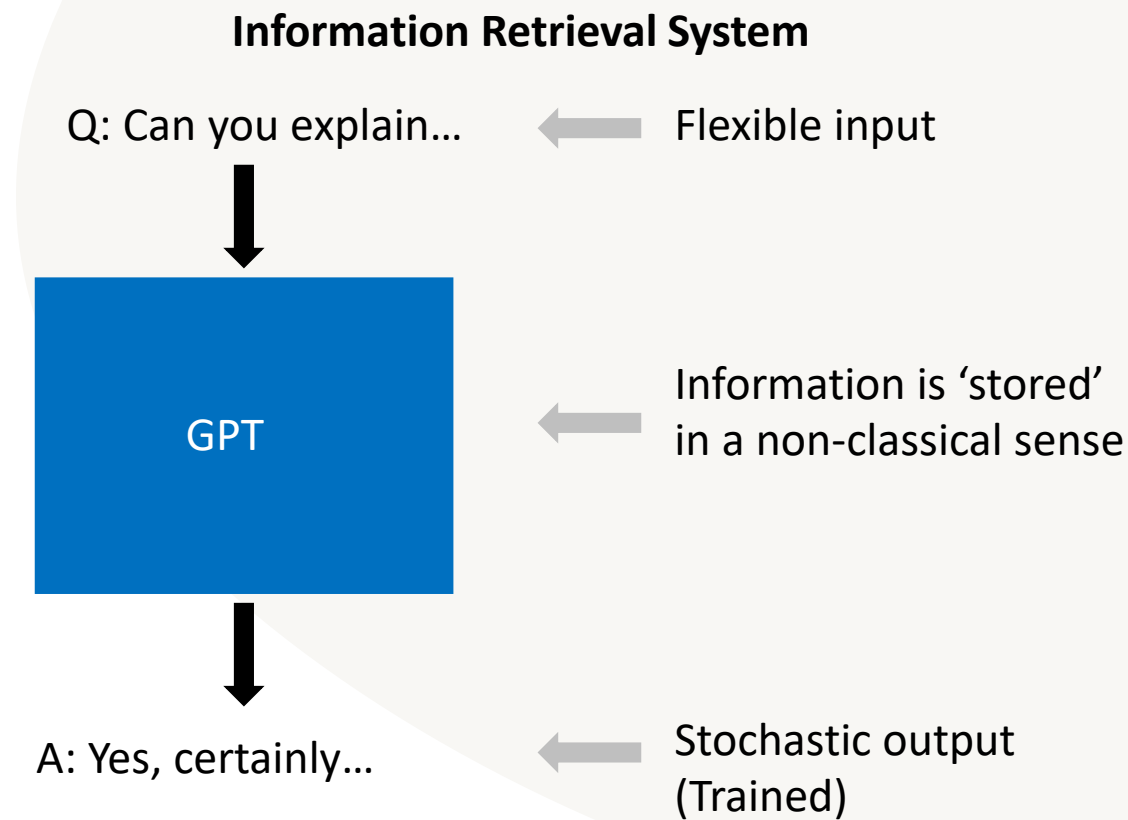
LLMs: Large Language Models



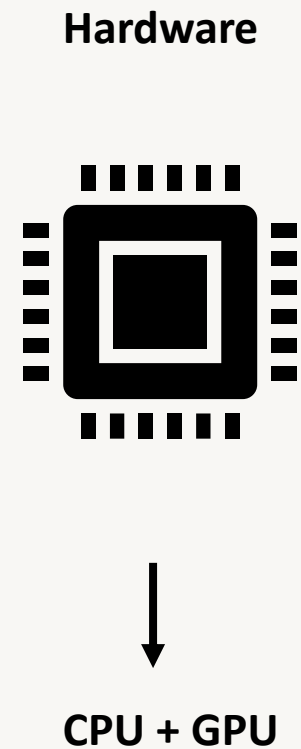
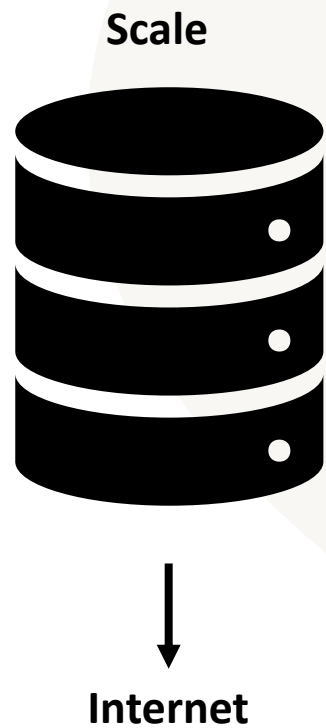
Training sentence: "All the very best"

<https://towardsdatascience.com/understanding-masked-language-models-mlm-and-causal-language-models-clm-in-nlp-194c15f56a5>; https://en.wikipedia.org/wiki/Giant_panda;
https://en.wikipedia.org/wiki/Polar_bear

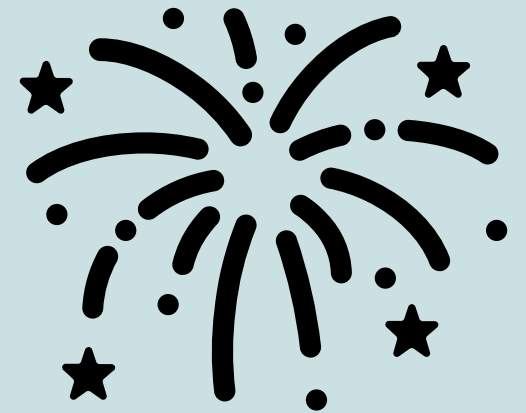
GPT: Generative Pre-trained Transformer



What drives the progress?



AI HYPE



AI Hype

TECH · A.I. CEO of Google's DeepMind says we could be 'just a few years' from A.I. that has human-level intelligence

BY TRISTAN ROVE
May 3, 2022 at 11:32 PM GMT+2



DeepMind CEO Demis Hassabis at the



Most Popular

LEADERSHIP
Uber puts its diversity

AI experts say 'the game is over' as Google nears human-level artificial intelligence

Sam Altman says we will have "unlimited intelligence and energy" before the decade it out

able of playing videogames, chatting, stacking blocks with a til AI overtakes us



Sam Altman
@sama · Follow



btw, i think we will have these before this decade is out.

not deployed at scale yet, but it will be clear that both are going to happen.



Sam Altman @sama

Replying to @sama

the future can be so good that it's hard for any of us to imagine. my basic take on this is that we will have 'unlimited' intelligence and energy, and all that will unlock.

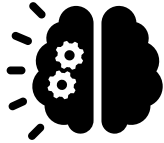
but the risks grow with the upside.

6:14 PM · May 1, 2022



Vocabulary in AI

Neural Network



Hallucination



Memorization



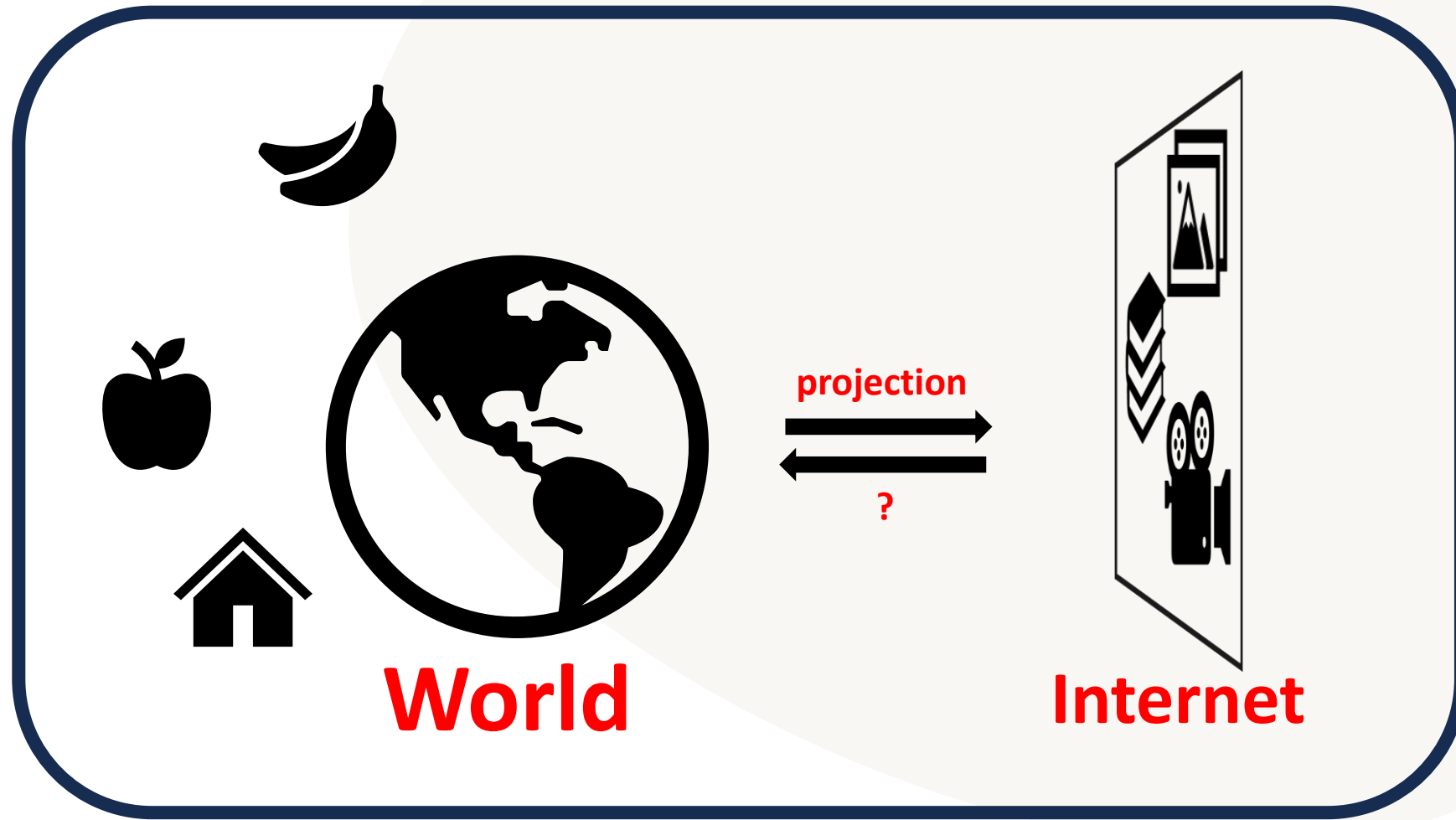
ChatGPT

LIMITATIONS



chatGPT limitations

- The world is complex



The internet is a projection of the world

chatGPT limitations

- Language vs. knowledge

'Memorization' + Language \approx Knowledge ?



'an apple'

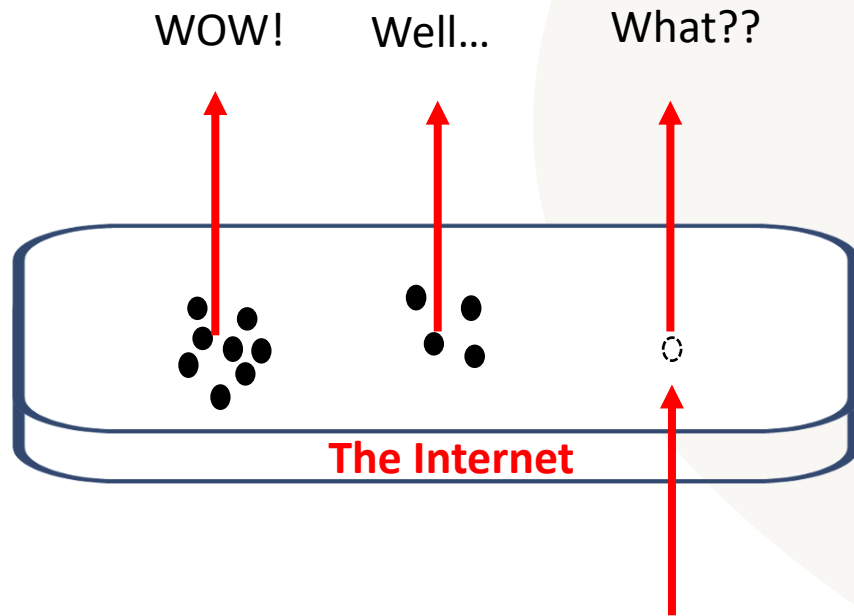


'a banana'

The girl reached for _____ from the fruit bowl.

chatGPT limitations

- Interpolation vs. extrapolation



True novelties (e.g. Innovations) are here

Horace He
@cHHillee · Follow

I suspect GPT-4's performance is influenced by data contamination, at least on Codeforces.

Of the easiest problems on Codeforces, it solved 10/10 pre-2021 problems and 0/10 recent problems.

This strongly points to contamination.

1/4

g's Race	implementation, math	greedy, implementation
nd Chocolate	implementation, math	cat?
triangle!	brute force, geometry, math	Actions
numbers	greedy, implementation, math	data structures, greedy, implementation, math
line Line	brute force	Interview Problem
r or Stairs?	brute force	brute force, implementation, strings
Loves 3 I	implementation	vers
s	implementation, math	nd Suffix Array
	greedy, implementation, sortings	strings
		ther Promotion
		greedy, math
		forces
		greedy, sortings
		land Append
		implementation, two pointers
		g Directions
		geometry, implementation

Pretraining Data Mixtures Enable Narrow Model Selection Capabilities in Transformer Models

Steve Yadlowsky, Lyric Doshi, Nilesch Tripuraneni
{yadlowsky, lyric, nileshtrip}@google.com
Google DeepMind

November 3, 2023

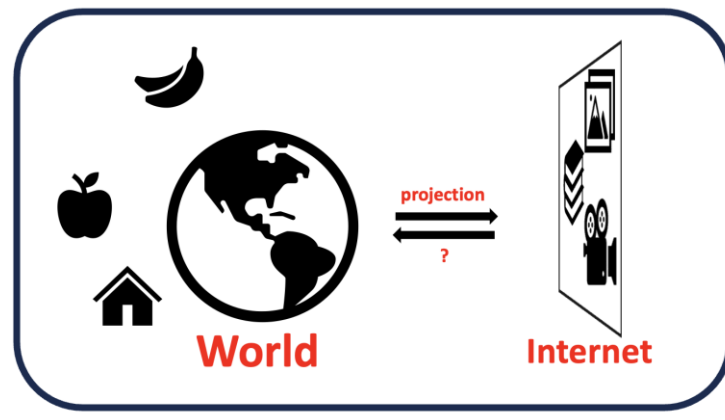
Abstract

Transformer models, notably large language models (LLMs), have the remarkable ability to perform in-context learning (ICL) – to perform new tasks when prompted with unseen input-output examples without any explicit model training. In this work, we study how effectively transformers can bridge between their pretraining data mixture, comprised of multiple distinct task families, to identify and learn new tasks in-context which are both inside and outside the pretraining distribution. Building on previous work, we investigate this question in a controlled

chatGPT limitations

- Context matters

Turn of the light, when you leave the room



chatGPT limitations

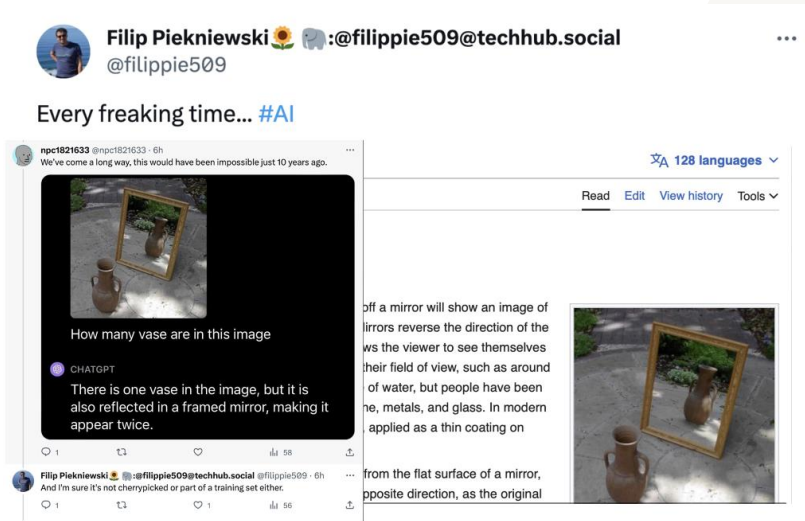
- Context matters: what is a pen?



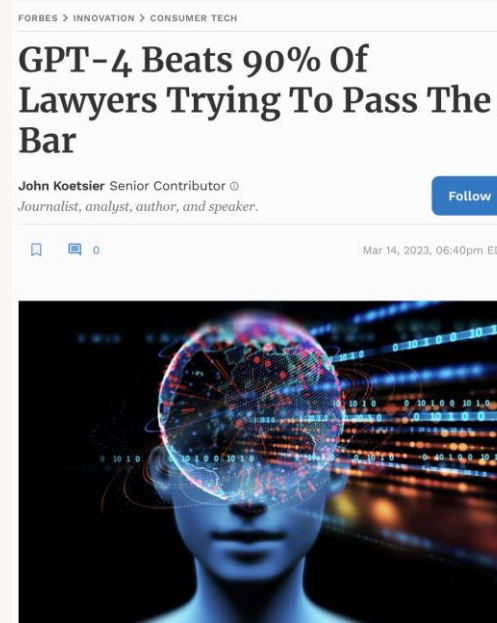
NOT an **intrinsic** property of the object itself – depends on **external** factors

chatGPT limitations

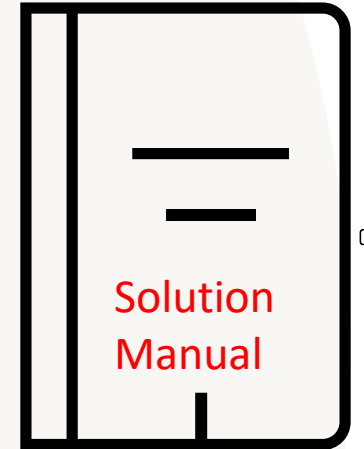
- Benchmarking is difficult



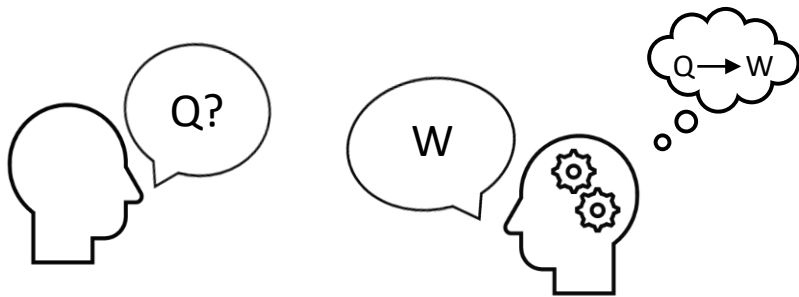
Benchmarking on training set + Techno-optimistic confirmation bias



"Companies that are slow to adopt AI will be left behind – large and small," says IDC analyst Mike Glennon. "AI is best used in



Intelligent (?)



Our questions are ordinary – have been asked before

Conclusion

- AI / Machine learning has great potential
- chatGPT is a **tool**
- Knowing limitations is important
- Future applications: Benchmark!

