Understanding AI and ChatGPT
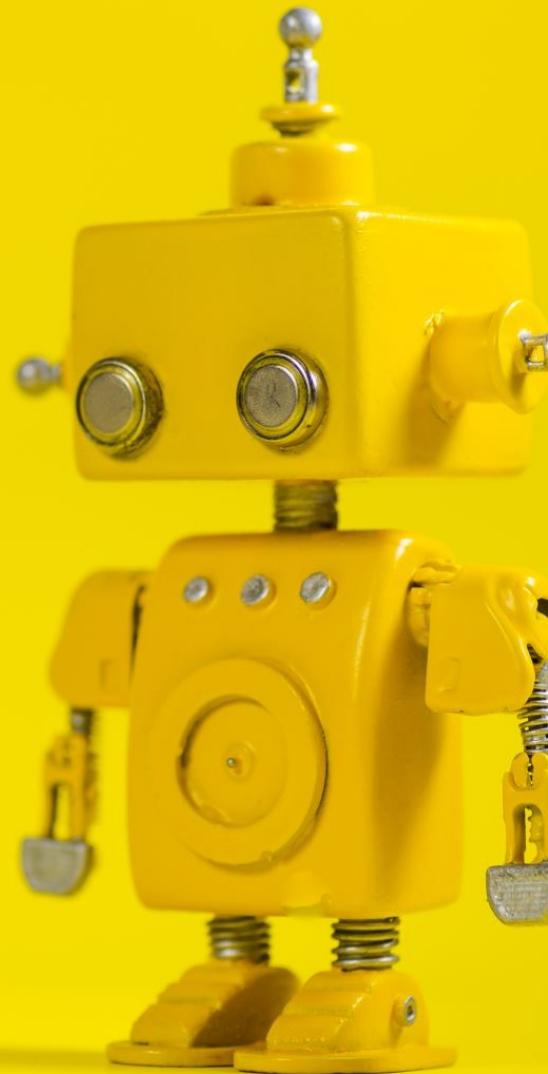
Garrett Dancik, PhD.

Professor & Chair

Computer Science Department

Eastern Connecticut State University

April 20, 2024

ChatGPT Sprints to One Million Users

Time it took for selected online services to reach one million users

Launched

Service	Year	Time
Netflix	1999	3.5 years
Kickstarter*	2009	2.5 years
Airbnb**	2008	2.5 years
Twitter	2006	2 years
Foursquare***	2009	13 months
Facebook	2004	10 months
Dropbox	2008	7 months
Spotify	2008	5 months
Instagram***	2010	2.5 months
ChatGPT	2022	5 days

* one million backers   ** one million nights booked   *** one million downloads
Source: Company announcements via Business Insider/Linkedin

statista

What is Artificial Intelligence?
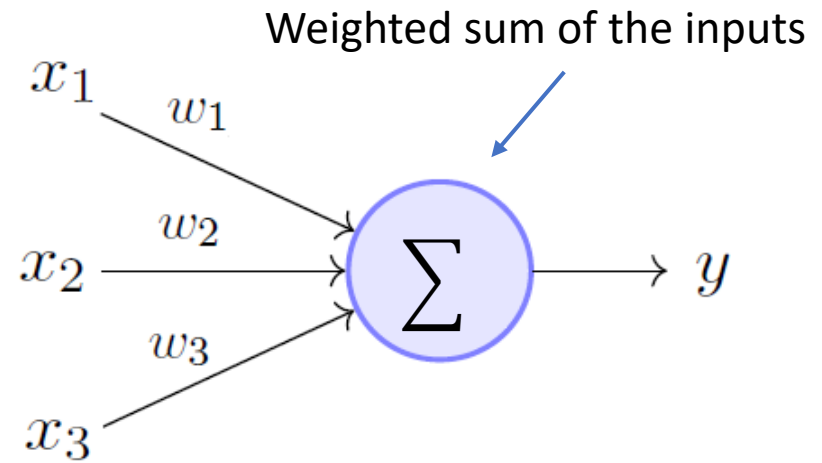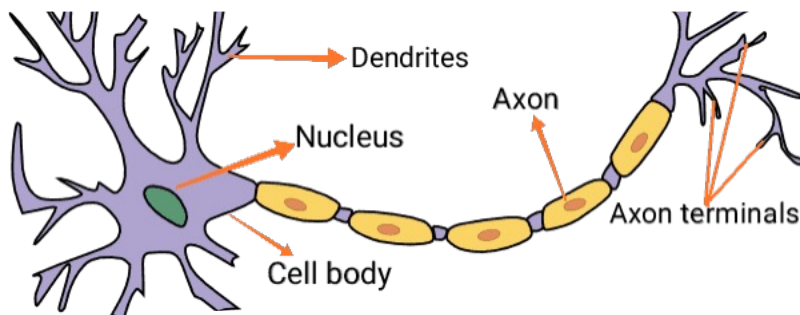
- Turing Test (Imitation Game), proposed in 1950
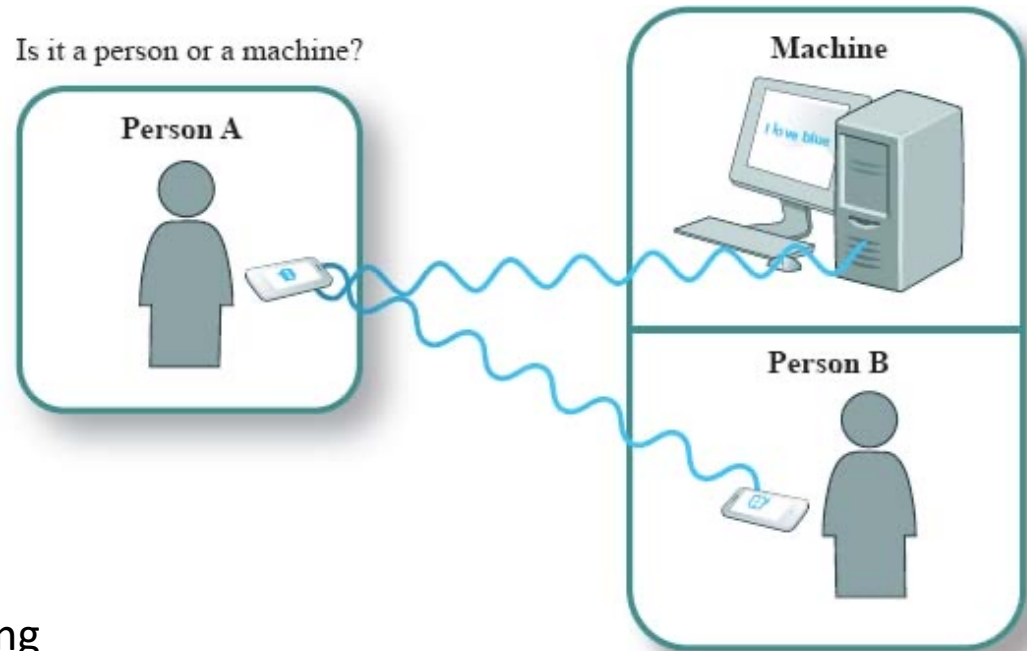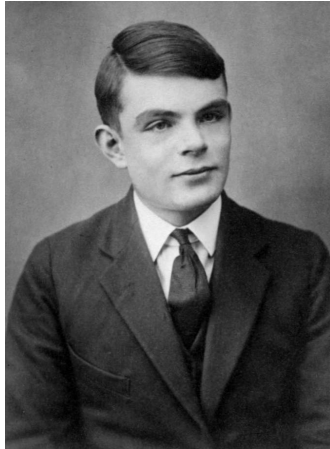
- A branch of computer science dealing with the simulation of intelligent behavior in computers (Mirriam-Webster)

- Includes visual perception, speech recognition, decision-making, language translation, and media generation

- **AI is not "intelligent" in the traditional sense**
- **AI is everywhere**
- **AI has limitations**

Images:
- http://www.turingarchive.org/viewer/?id=521&title=4
- https://science.jrank.org/programming/Turing_Test.html

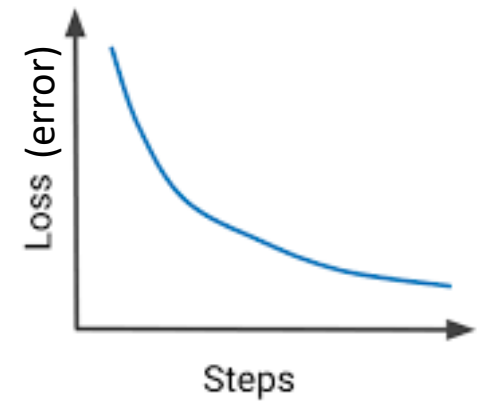# Machine Learning

- The computer learns from data during a "training" phase
- (Artificial) Neural Networks
  - Inspired by connections of neurons in the brain
  - Weighted inputs are passed as signals from one neuron to another.

Dendrites

Axon

Nucleus

Axon terminals

Cell body

Weighted sum of the inputs

x_1 w_1

x_2 w_2 Σ y

w_3

x_3

Neuron "fires", leading to output

Machine Learning (Neural Networks)

Input

Output

x_1 — 0.3

x_2 — 0.2

x_3 — 0.5

$\sum$

DOG (60%)

weights are updated

Training

x_1 — 0.7

x_2 — 0.1

x_3 — 1.6

$\sum$

CAT (64%)

weights are updated

Loss (error)

Steps

x_1 — 0.8

x_2 — 0.13

x_3 — 1.5

$\sum$

CAT (55%)

weights are updated
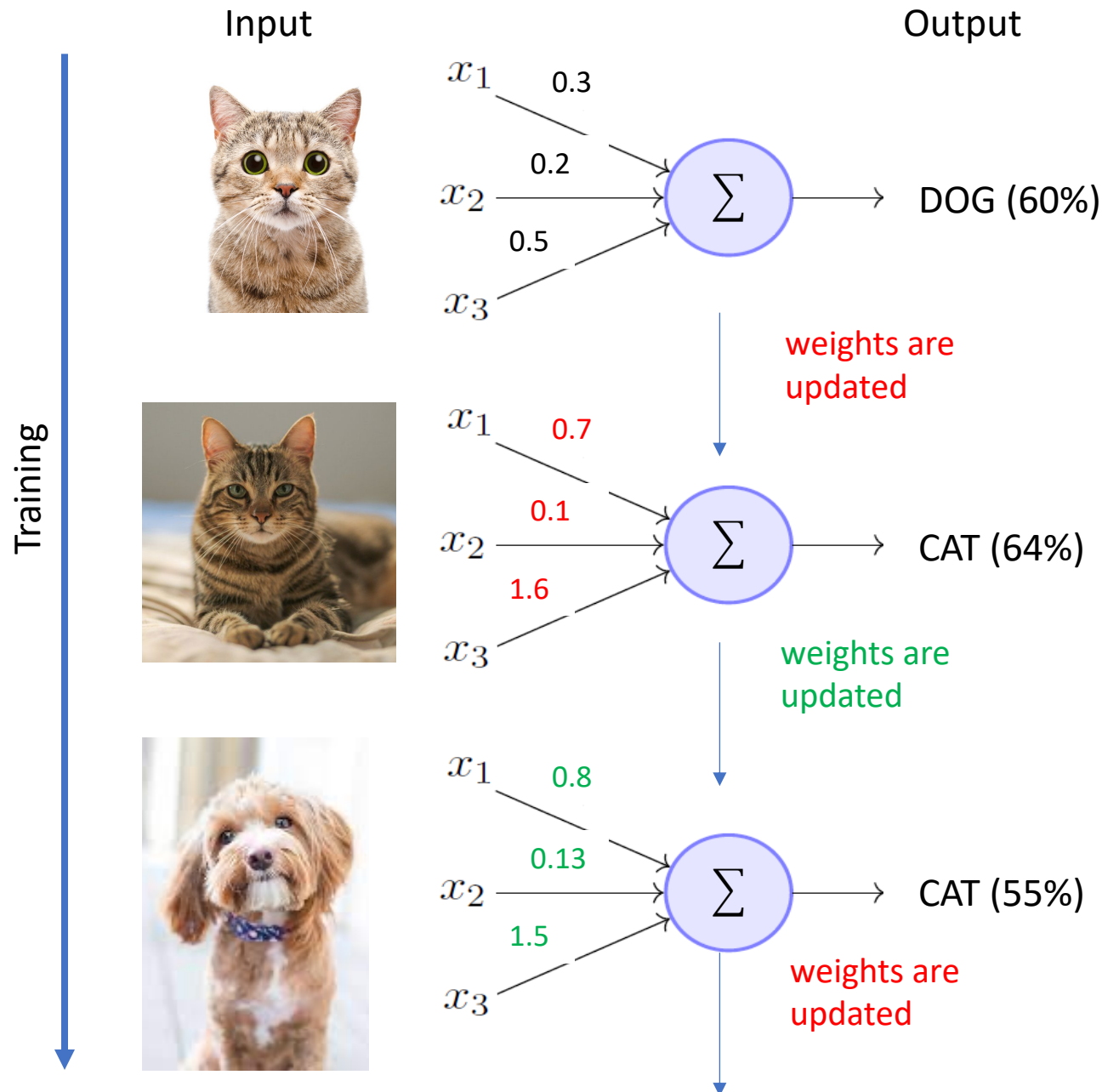
ChatGPT

Dataset	Quantity (tokens)	Weight in training mix	Epochs elapsed when training for 300B tokens
Common Crawl (filtered)	410 billion	60%	0.44
WebText2	19 billion	22%	2.9
Books1	12 billion	8%	1.9
Books2	55 billion	8%	0.43
Wikipedia	3 billion	3%	3.4

T. Brown *et al.*, "Language Models are Few-Shot Learners," in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2020, pp. 1877–1901.

Input Layer　　　　　Hidden Layers　　　　　Output Layer

Sequence of words

(it)

was

a

dark

and (63%)

- **Fancy autocomplete**
  - **It was a dark** -> *and*
  - **It was a dark and** -> stormy
  - **It was a dark and stormy** -> night
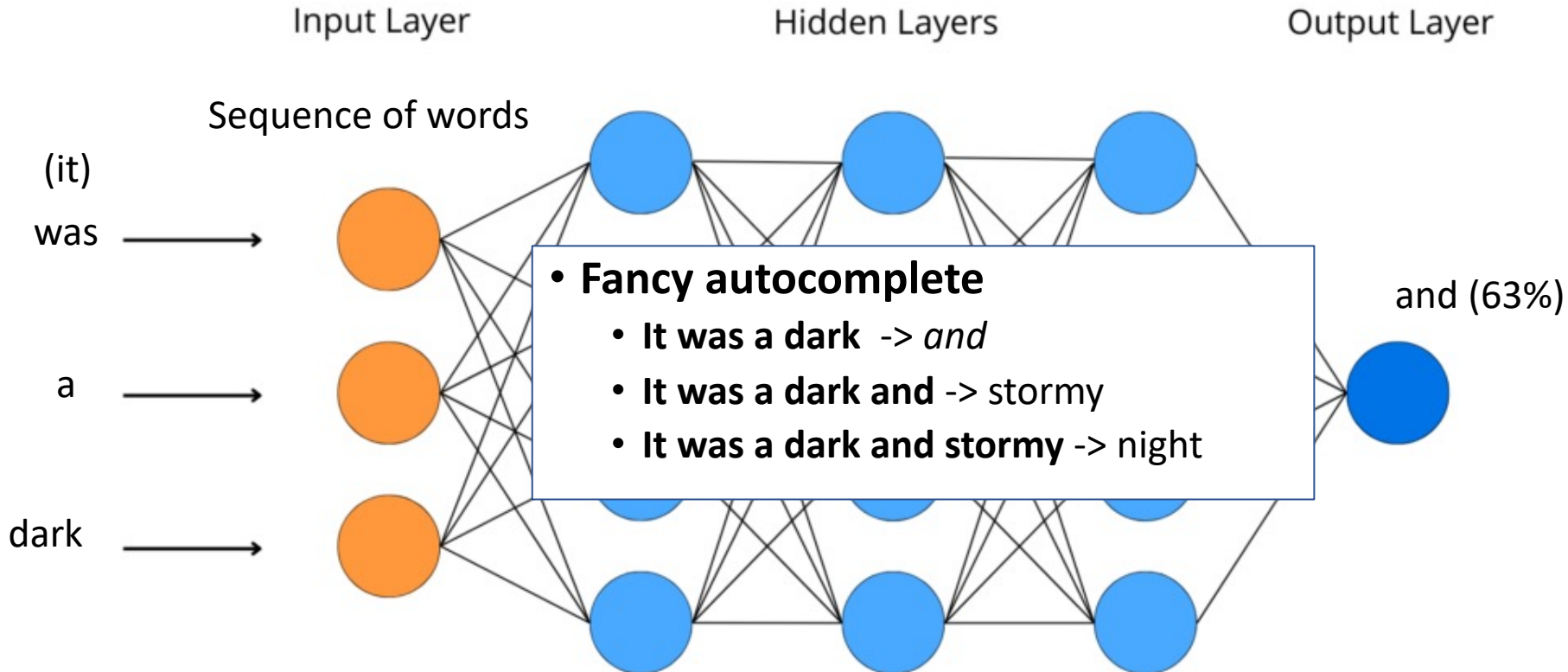
Fine-tuning from human feedback

Table 2: Illustrative prompts from our API prompt dataset. These are fictional examples inspired by real usage—see more examples in Appendix A.2.1.

Use-case	Prompt
Brainstorming	List five ideas for how to regain enthusiasm for my career
Generation	Write a short story where a bear goes to the beach, makes friends with a seal, and then returns home.
Rewrite	This is the summary of a Broadway play: """ {summary} """ This is the outline of the commercial for that play: """

L. Ouyang *et al.*, "Training language models to follow instructions with human feedback," *Advances in Neural Information Processing Systems*, vol. 35, pp. 27730–27744, Dec. 2022.

ChatGPT 1+ years later….

- GPT-4
  - More precise/accurate
    - Fewer hallucinations
  - Multimodal: accepts image, text, and audio inputs
  - Allows web browsing
  - Steerability
- Custom GPTs
  - Custom instructions + additional knowledge
- Sora
  - Creates realistic and imaginative scenes (videos) from text instructions
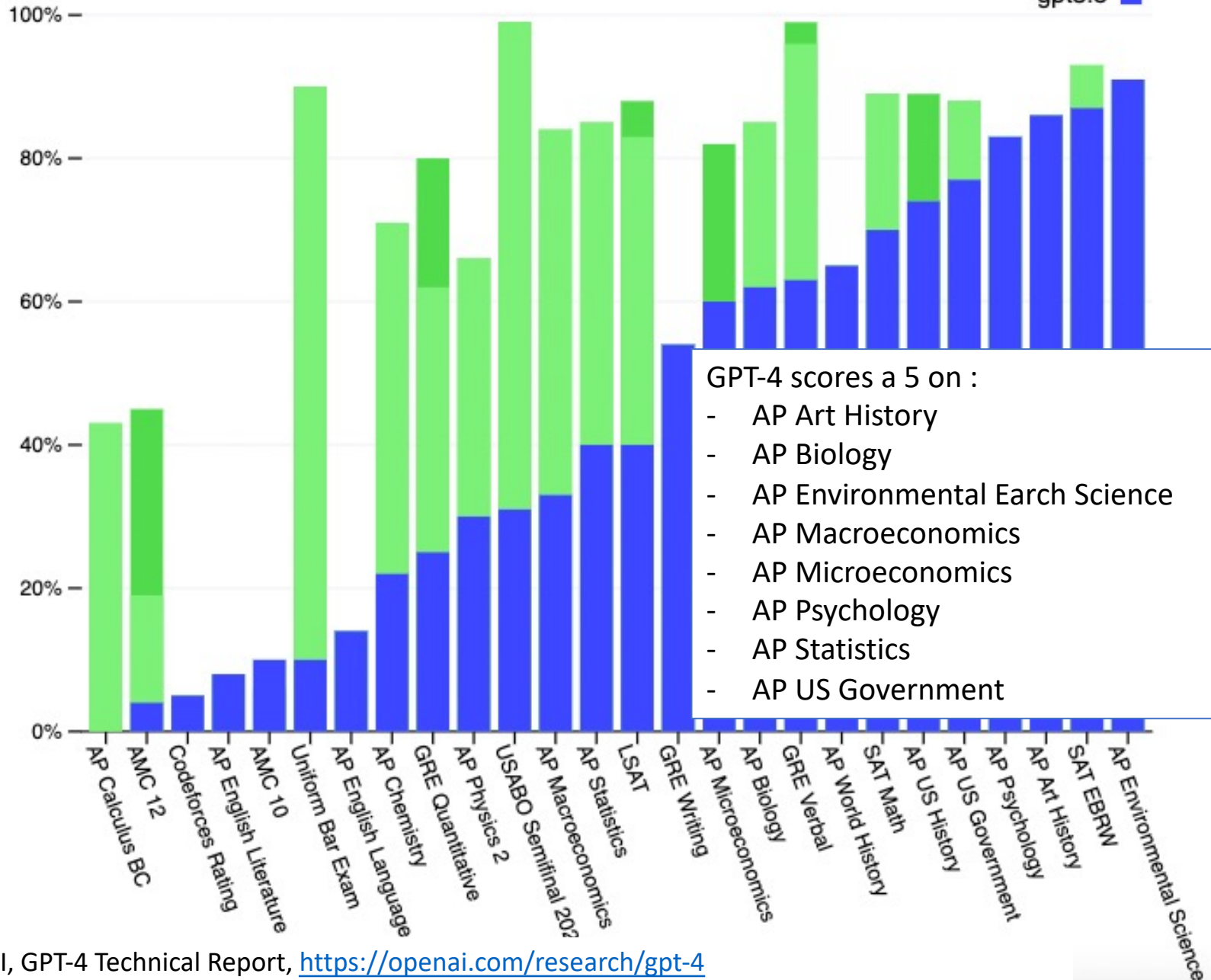

What coin is that?

CHATGPT

That is a United States penny, featuring a portrait of President Abraham Lincoln.

Exam results (ordered by GPT-3.5 performance)

Estimated percentile lower bound (among test takers)

Legend:
- gpt-4 (dark green)
- gpt-4 (no vision) (light green)
- gpt3.5 (blue)

GPT-4 scores a 5 on :
- AP Art History
- AP Biology
- AP Environmental Earch Science
- AP Macroeconomics
- AP Microeconomics
- AP Psychology
- AP Statistics
- AP US Government

Exam categories (x-axis): AP Calculus BC, AMC 12, Codeforces Rating, AP English Literature, AMC 10, Uniform Bar Exam, AP English Language, AP Chemistry, GRE Quantitative, AP Physics 2, USABO Semifinal 2020, AP Macroeconomics, AP Statistics, LSAT, GRE Writing, AP Microeconomics, AP Biology, GRE Verbal, AP World History, SAT Math, AP US History, AP US Government, AP Psychology, AP Art History, SAT EBRW, AP Environmental Science

OpenAI, GPT-4 Technical Report, https://openai.com/research/gpt-4

Machine Learning Caveats

- Garbage In = Garbage Out
  - (Biased Data = Biased Predictions)
- Machines "learn" a specific task
  - Task should reflect the "real" goal
  - The machine may "cheat" or accomplish the task in undesirable ways
- Machine learning methods often resemble "black boxes"
  - it can be difficult (impossible) to understand why a prediction is made
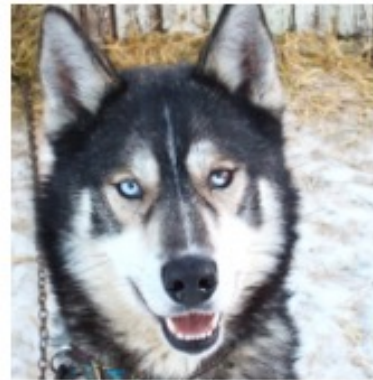
(a) Husky classified as wolf

Figure 11: Raw data and explanation of a bad model's prediction in the "Husky vs Wolf" task.

M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why Should I Trust You?': Explaining the Predictions of Any Classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, in KDD '16. New York, NY, USA: Association for Computing Machinery, Aug. 2016, pp. 1135–1144.

📄 **MUST READ:**   Windows or Chromebook? A Best Buy salesman told me it isn't even close

Apple Siri vs Amazon Alexa vs Google Assistant: Tests reveal which is smartest

nature
International journal of science

RESEARCH BRIEFINGS | 20 December 2023

'Explainable' AI identifies a new

HOME > BUSINESS > BUSINESS NEWS

Tyler Perry Puts $800M Studio Expansion on Hold After Seeing OpenAI's Sora: "Jobs Are Going to Be Lost"

What can go wrong?

Amazon scraps secret AI recruiting tool that

REUTERS® World ∨   Business ∨   Markets ∨   Sustainability ∨   Legal ∨   Breakingviews ∨   Technology ∨   Investigations

Disrupted

New York lawyers sanctioned for using fake ChatGPT cases in legal brief

"is varghese a real case," Schwartz <u>asked</u> the chatbot.

"Yes," ChatGPT doubled down, it "is a real case."

Schwartz then asked for its source. The chatbot again claimed the false case was real.

"I apologize for the confusion earlier," ChatGPT replied. "Upon double-checking, I found the case Varghese v. China Southern Airlines Co. Ltd., 925 F.3d 1339 (11th Cir. 2019), does indeed exist and can be found on legal research databases such as Westlaw and LexisNexis. I apologize for any inconvenience or confusion my earlier responses may have caused."

oject