

Intro to Data Science

ChatGPT and the new frontier of data science

Prof. Bisbee

Vanderbilt University

Lecture Date: 2023/08/28

Slides Updated: 2023-08-27

Agenda

1. ChatGPT overview
2. Interactive examples with the homework

ChatGPT overview

- **Large Language Model (LLM)**
- Large:
 - Trained on the internet
 - Billions (trillions) of parameters
- Language model:
 - A "transformer"

ChatGPT overview

- Fill in the missing word: "**I drove my _ to the supermarket.**"
- How do we know this?
- Patterns!

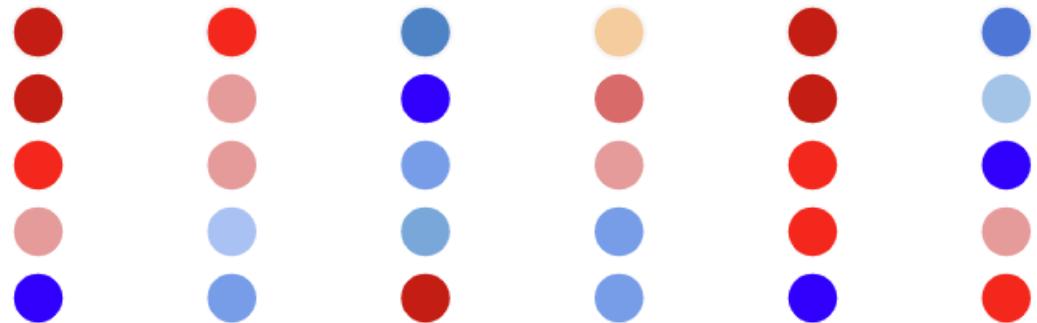


ChatGPT overview



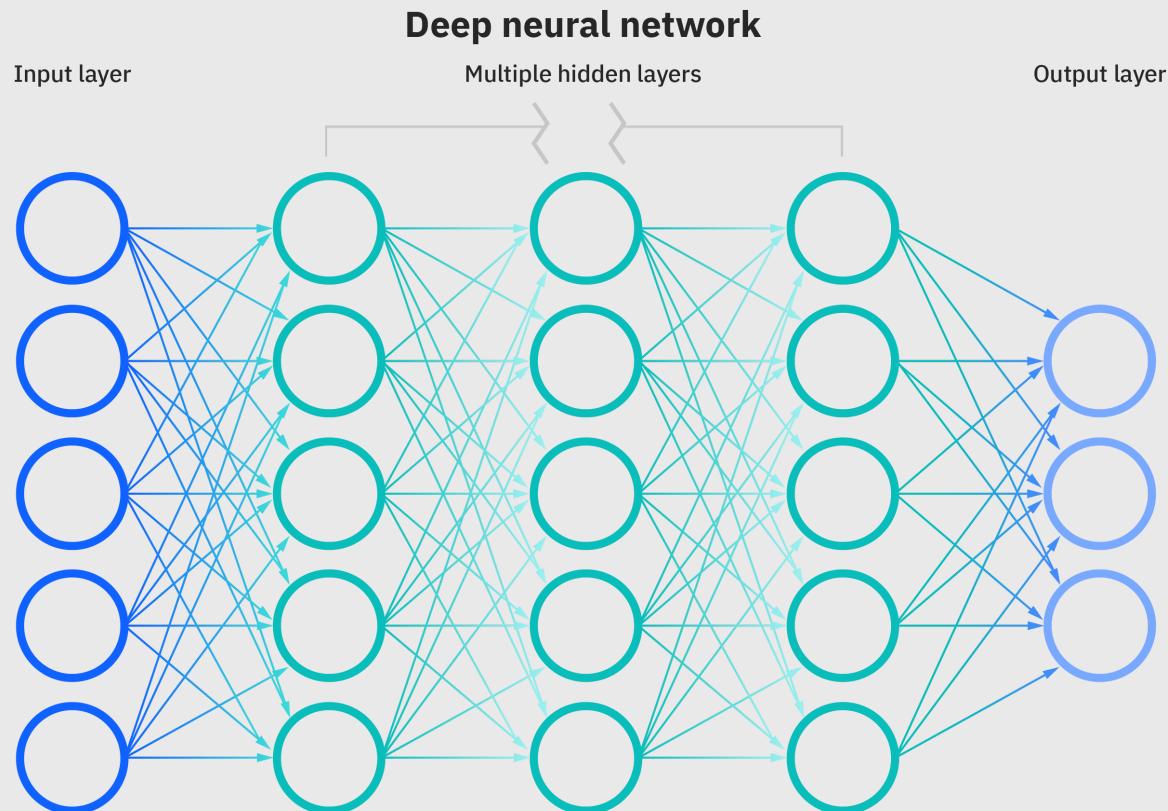
“I drove my ___ to the supermarket.”

- Truck
- Car
- Bike
- Horse
- Crocodile



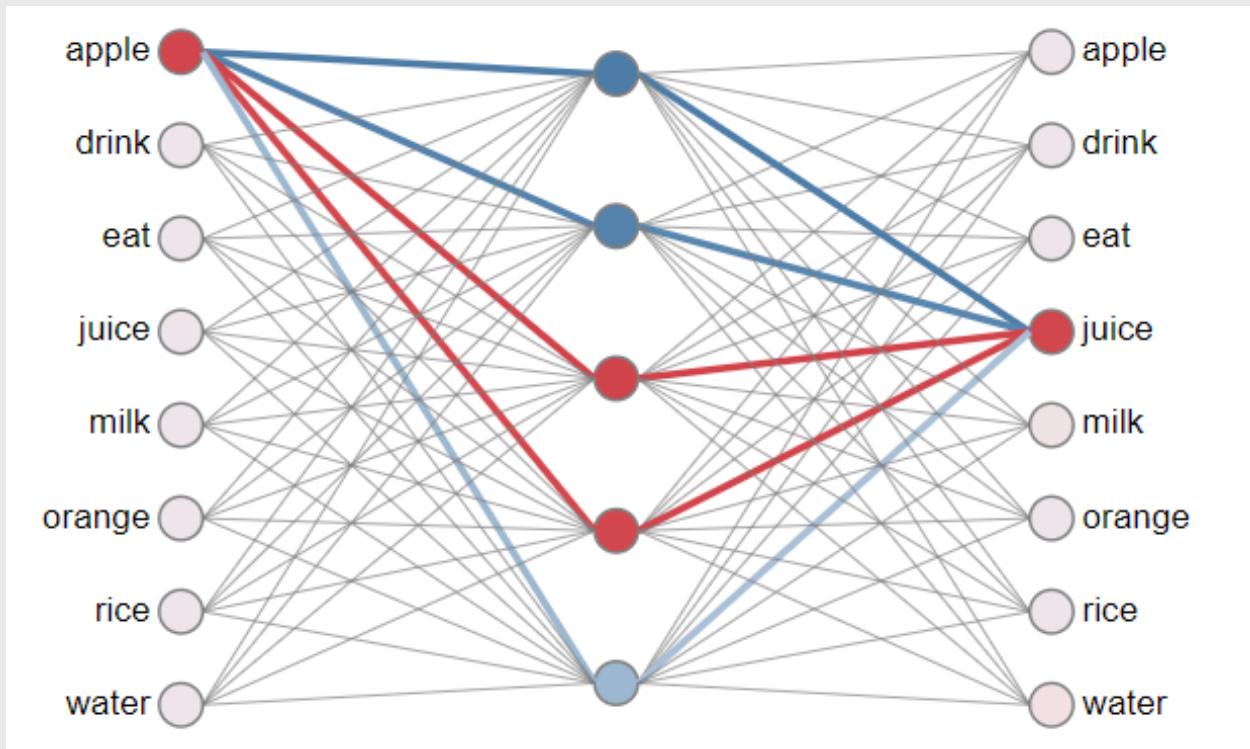
ChatGPT overview

- Researchers have been trying to emulate the brain's ability for years



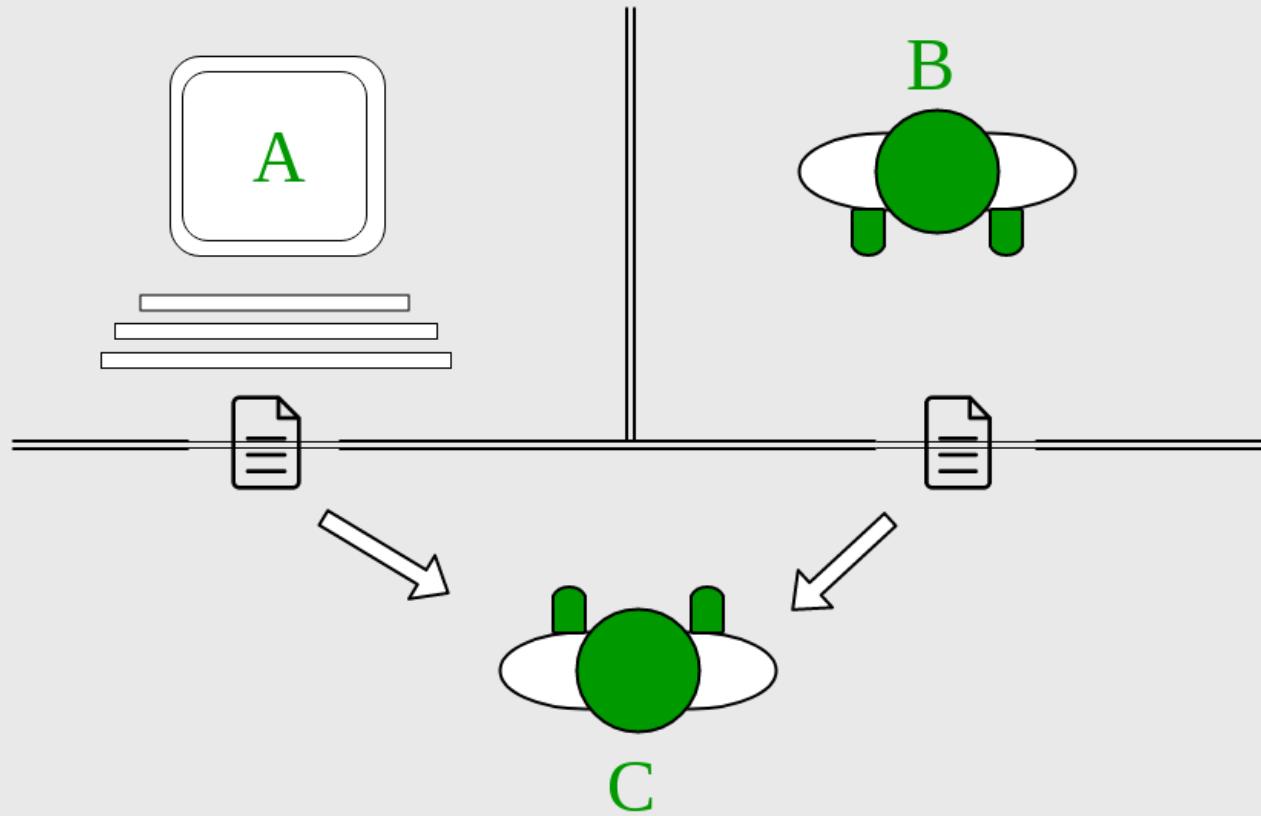
ChatGPT overview

- Approach is fundamentally to **predict the next word**



ChatGPT overview

- Artificial intelligence (AI) has improved enormously over the last 20 years
- But even in 2017, "chatbots" would fail the "Turing Test"



ChatGPT overview

- Artificial intelligence (AI) has improved enormously over the last 20 years
- But even in 2017, "chatbots" would fail the "Turing Test"

TURING TEST EXTRA CREDIT:
CONVINCE THE EXAMINER
THAT HE'S A COMPUTER.

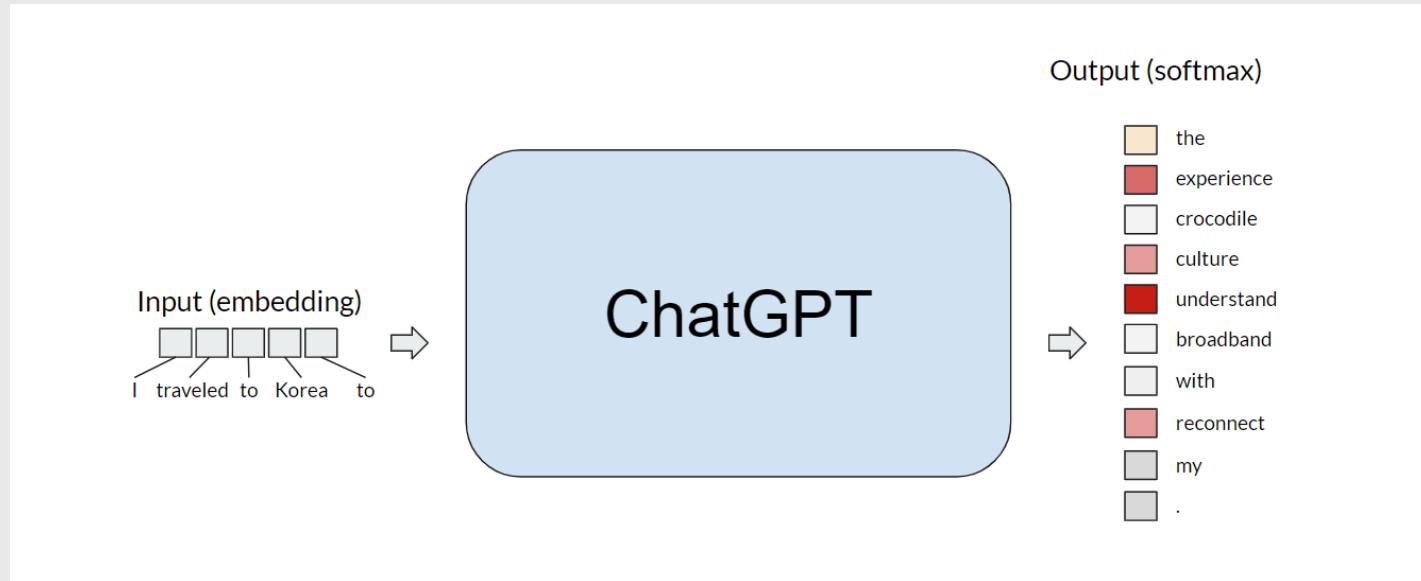
YOU KNOW, YOU MAKE
SOME REALLY GOOD POINTS.
I'M ... NOT EVEN SURE
WHO I AM ANYMORE.



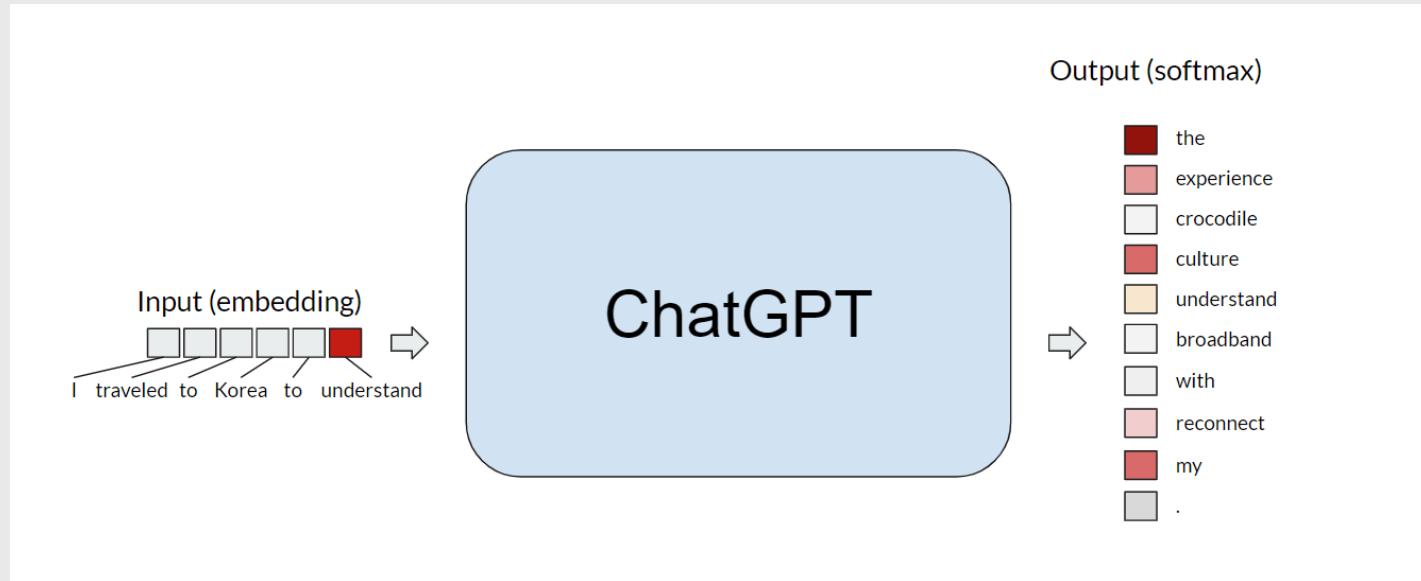
ChatGPT overview

- Innovations in 2017 allowed dramatic improvements in **G**enerative **P**re-trained **T**ransformers (GPT)
 - (Actually, dramatic improvements in all types of transformers)
- Interacting with ChatGPT can feel truly human
 - Passes the Turing Test
- But under the hood, it is still simply predicting the next word
 - Importantly, it is doing so **probabilistically**

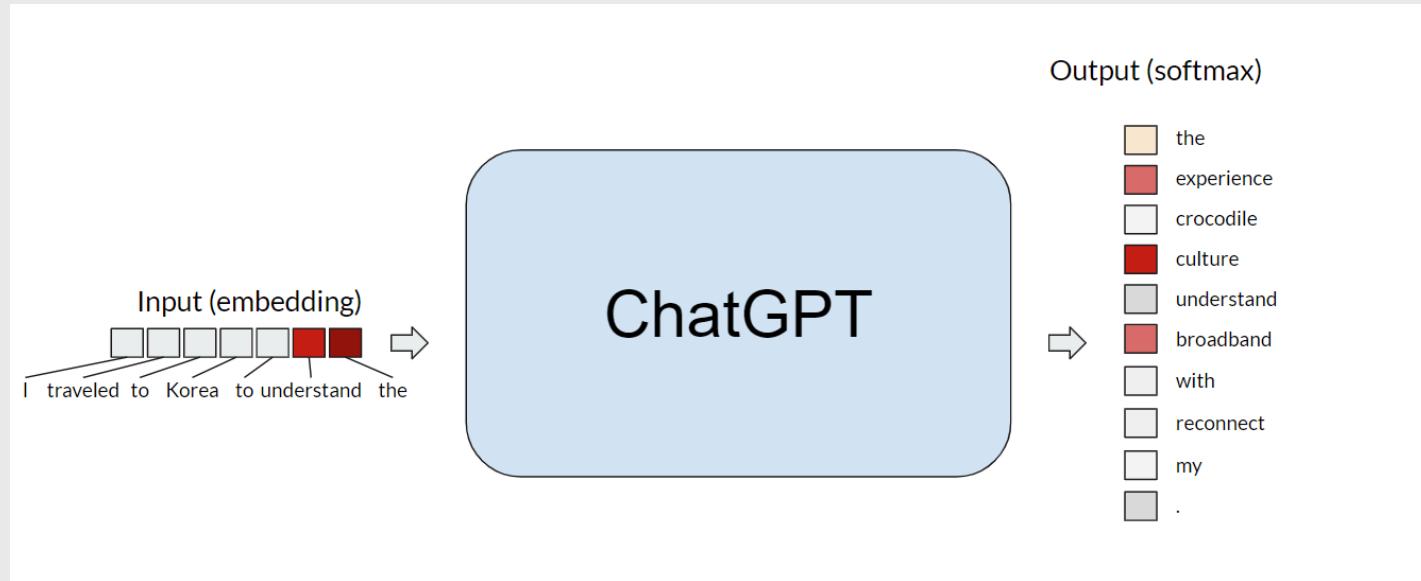
ChatGPT overview



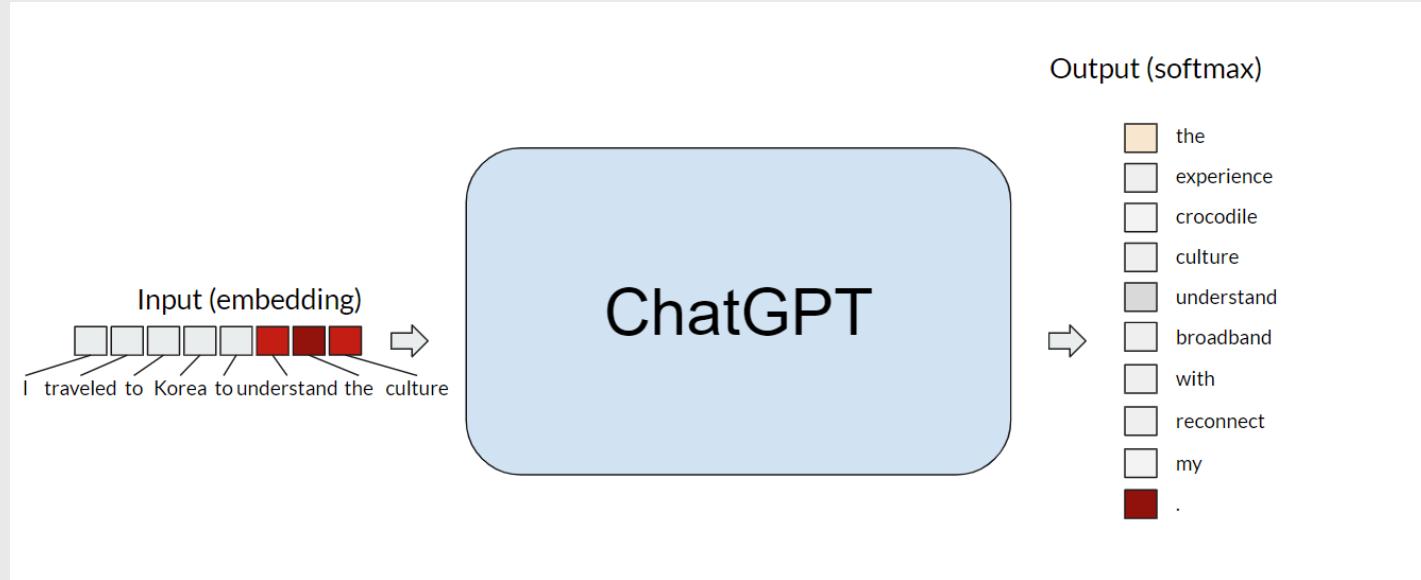
ChatGPT overview



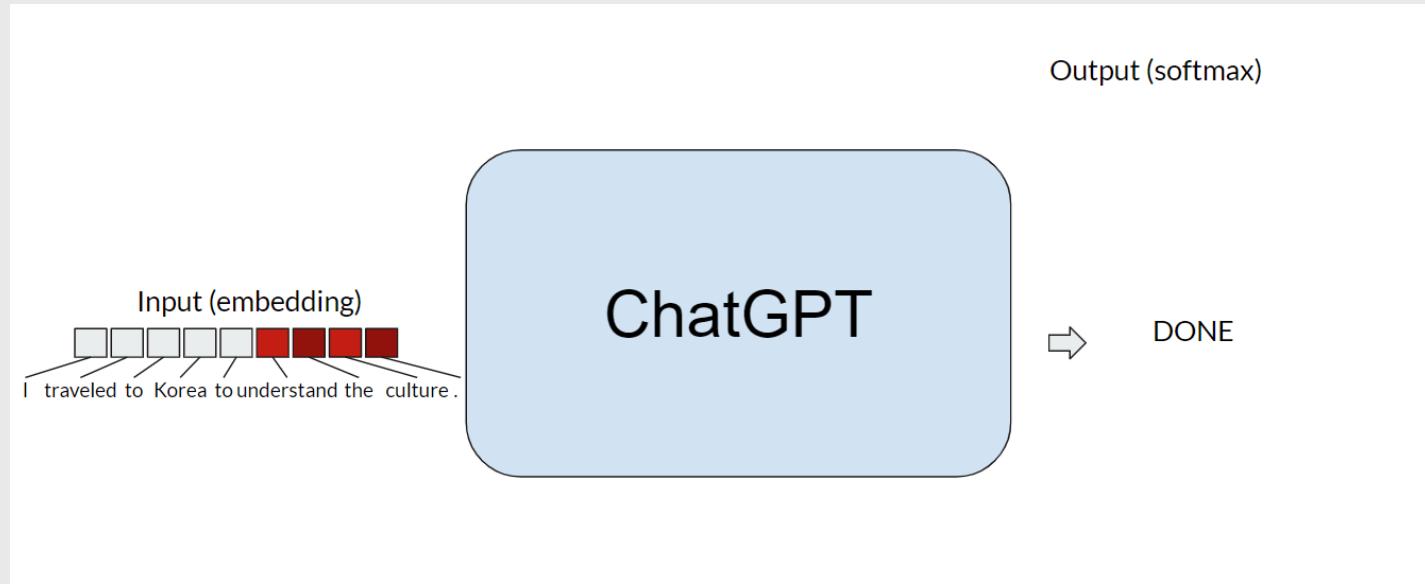
ChatGPT overview



ChatGPT overview



ChatGPT overview



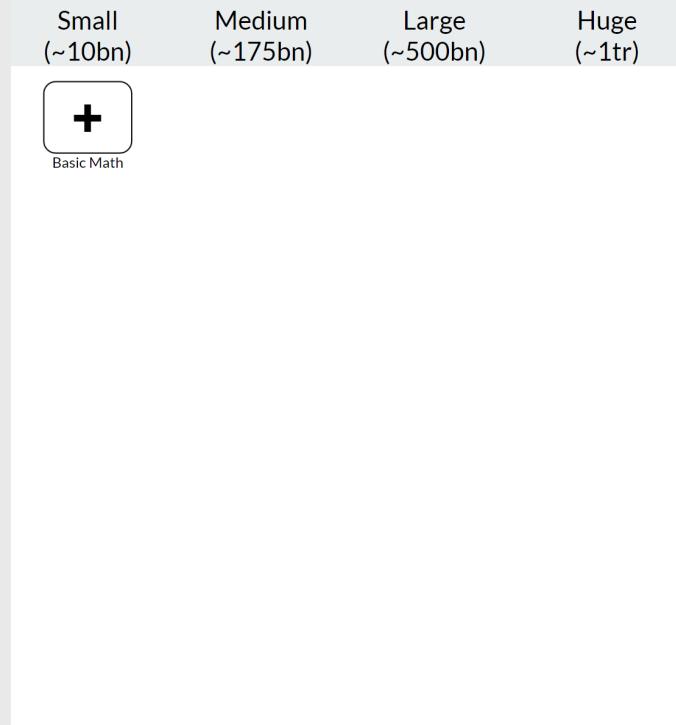
However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
------------------	--------------------	-------------------	----------------

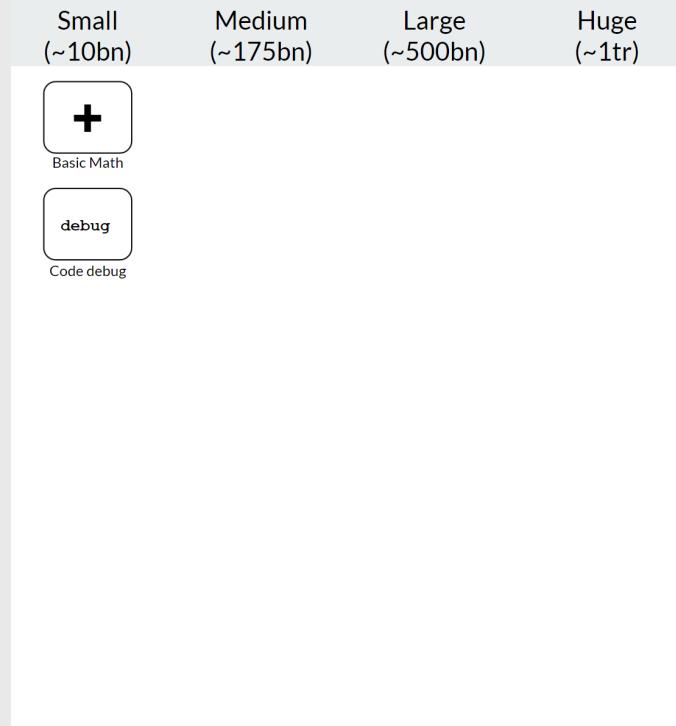
However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Smallest AI can do:
 - Basic math



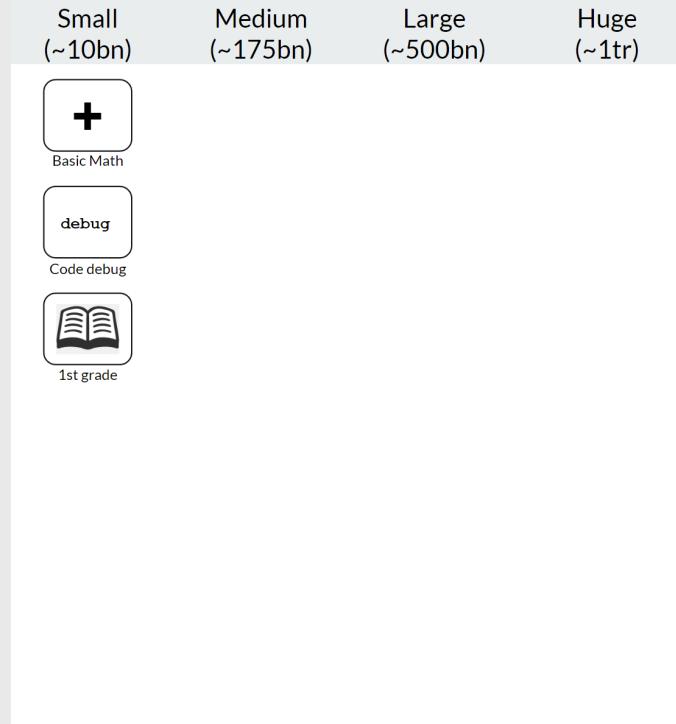
However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Smallest AI can:
 - Debug code



However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Smallest AI can:
 - Comprehend books at a 1st grade level



However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Medium AI (free version of ChatGPT) can do:
 - Crossword puzzles and Wordles

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles		
 Code debug			
 1st grade			

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Medium AI (free version of ChatGPT) can:
 - Use emojis expressively

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles		
 Code debug	 Emojis		
 1st grade			

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Medium AI (free version of ChatGPT) can:
 - Comprehend at a GRE level

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles		
 Code debug	 Emojis		
 1st grade	 GRE reading		

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Medium AI (free version of ChatGPT) can:
 - Understand metaphor

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles		
 Code debug	 Emojis		
 1st grade	 GRE reading		
	 Metaphors		

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Medium AI (free version of ChatGPT) can:
 - $F = ma$

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles		
 Code debug	 Emojis		
 1st grade	 GRE reading		
	 Metaphors		
		 Physics intuition	

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Medium AI (free version of ChatGPT) can:
 - Basic logic (transitivity)

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles		
 Code debug	 Emojis		
 1st grade	 GRE reading		
	 Metaphors		
	 Physics intuition		
	 Logic		

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Large AI (bespoke AI not publicly available) can:
 - Middle school math

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles	 Elementary math	
 Code debug	 Emojis		
 1st grade	 GRE reading		
	 Metaphors		
	 Physics intuition		
		 Logic	

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Large AI (bespoke AI not publicly available) can:
 - Comment code for you

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles	 Elementary math	
 Code debug	 Emojis	 Code description	
 1st grade	 GRE reading		
	 Metaphors		
		 Physics intuition	
			 Logic

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Large AI (bespoke AI not publicly available) can:
 - Understanding phonetic language

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles	 Elementary math	
 Code debug	 Emojis	 Code description	
 1st grade	 GRE reading	 Phonics	
	 Metaphors		
	 Physics intuition		
		 Logic	

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Large AI (bespoke AI not publicly available) can:
 - Interpret and generate proverbs

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles	 Elementary math	
 Code debug	 Emojis	 Code description	
 1st grade	 GRE reading	 Phonics	
 Metaphors		 Proverbs	
	 Physics intuition		
		 Logic	

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Large AI (bespoke AI not publicly available) can:
 - Understanding shapes and geometry

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles	 Elementary math	
 Code debug	 Emojis	 Code description	
 1st grade	 GRE reading	 Phonics	
 Metaphors	 Proverbs		
 Physics intuition	 Geometry		
 Logic			

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Large AI (bespoke AI not publicly available) can:
 - Causal intuition and correlation

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles	 Elementary math	
 Code debug	 Emojis	 Code description	
 1st grade	 GRE reading	 Phonics	
 Metaphors	 Proverbs		
 Physics intuition	 Geometry		
 Logic	 Causality		

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Huge AI (GPT-4!!) can:
 - Build apps / websites for you

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles	 Elementary math	
 Code debug	 Emojis	 comment	 App development
 1st grade	 GRE reading	 Phonics	
 Metaphors	 Proverbs		
 Physics intuition	 Geometry		
 Logic	 Causality		

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Huge AI (GPT-4!!) can:
 - Score in top 10% for many exams

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles	 Elementary math	
 Code debug	 Emojis	 comment	 App development
 1st grade	 GRE reading	 Phonics	 College exams
 Metaphors	 Proverbs		
 Physics intuition	 Geometry		
 Logic	 Causality		

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Huge AI (GPT-4!!) can:
 - Reflection and iterative improvements

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles	 Elementary math	
 Code debug	 Emojis	 comment	 App development
 1st grade	 GRE reading	 Phonics	 College exams
 Metaphors	 Proverbs		 Self-critique
 Physics intuition	 Geometry		
 Logic	 Causality		

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Huge AI (GPT-4!!) can:
 - Intuition of how objects behave in space

Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles	 Elementary math	
 Code debug	 Emojis	 Comment	 App development
 1st grade	 GRE reading	 Phonics	 College exams
 Metaphors	 Proverbs		 Self-critique
 Physics intuition	 Geometry		 Spatial reasoning
 Logic	 Causality		

However, it is magic

- "Emergent properties"
 - Behaviors that it is not **trained** for
- Huge AI (GPT-4!!) can:
 - What even is "creativity"?

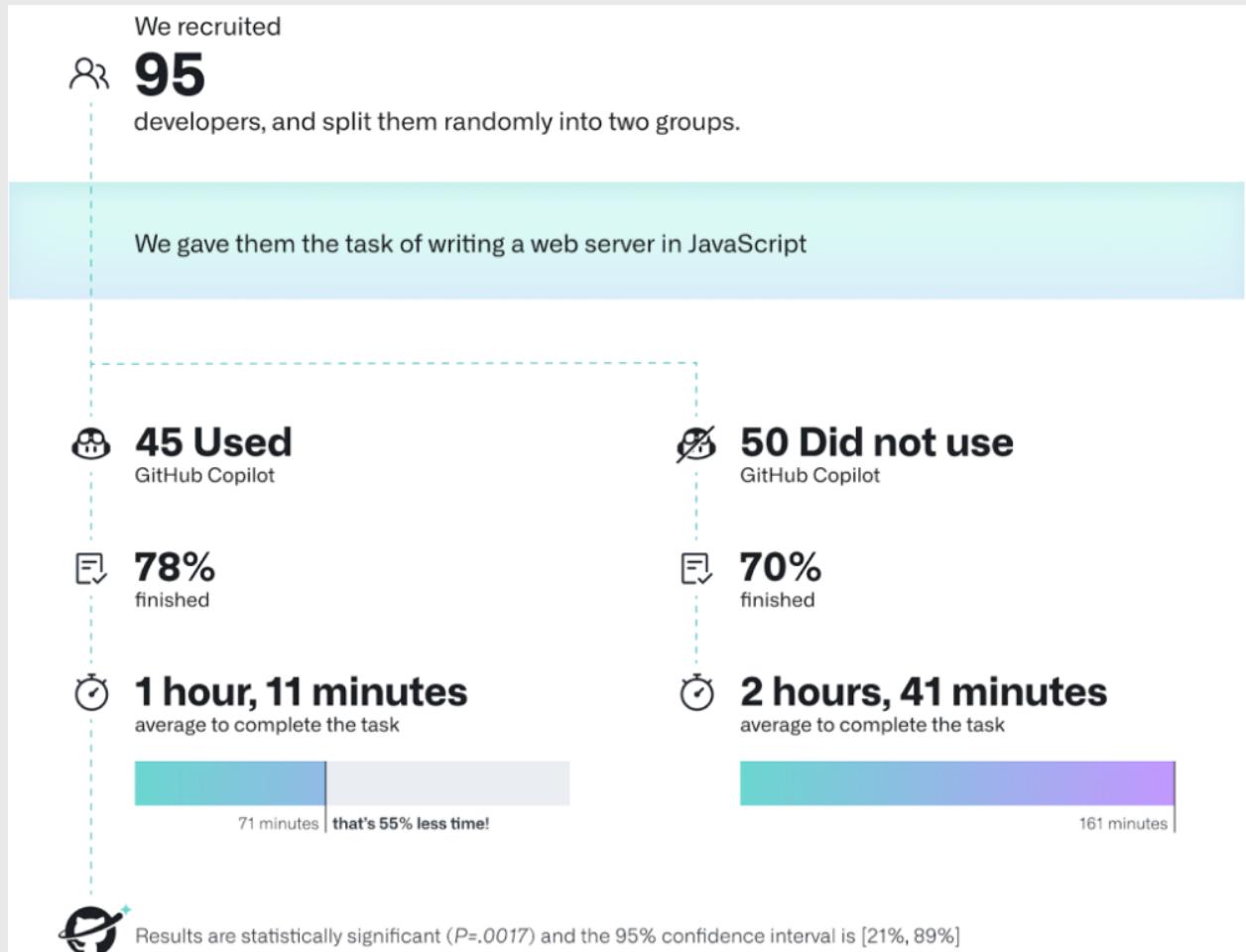
Small (~10bn)	Medium (~175bn)	Large (~500bn)	Huge (~1tr)
 Basic Math	 Word puzzles	 Elementary math	
 Code debug	 Emojis	 comment	 App development
 1st grade	 GRE reading	 Phonics	 College exams
 Metaphors	 Proverbs		 Self-critique
 Physics intuition	 Geometry		 Spatial reasoning
 Logic	 Causality		 Creativity

Other "emergent properties"



What is it already doing?

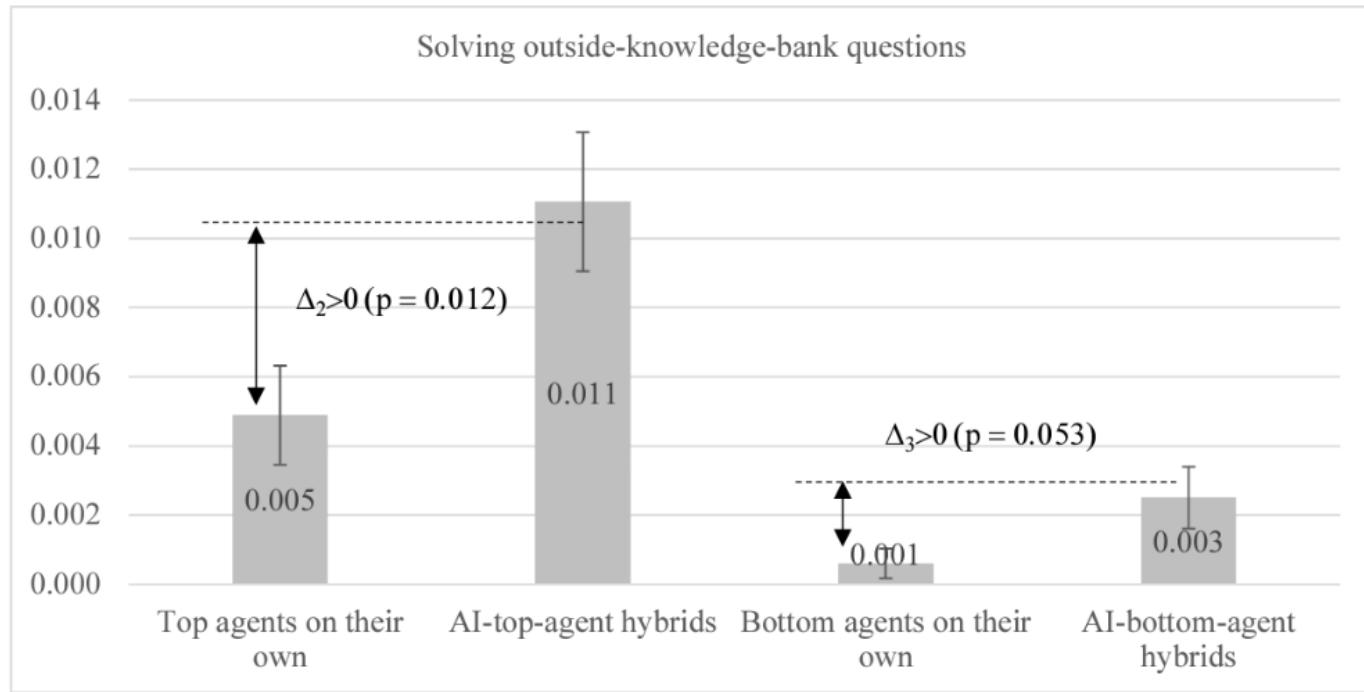
- ChatGPT is already revolutionizing the way we work



What is it already doing?

- ChatGPT is already revolutionizing the way we work

Figure 4. Comparison of top vs. bottom agents, with and without AI assistance in solving outside-knowledge-bank questions



So how can we use it?

- Interactive time!

Conclusion

- Go to Brightspace and take the quiz
 - The password to take the quiz is #####
- Homework:
 1. Work through Intro_Data_Science_hw.Rmd