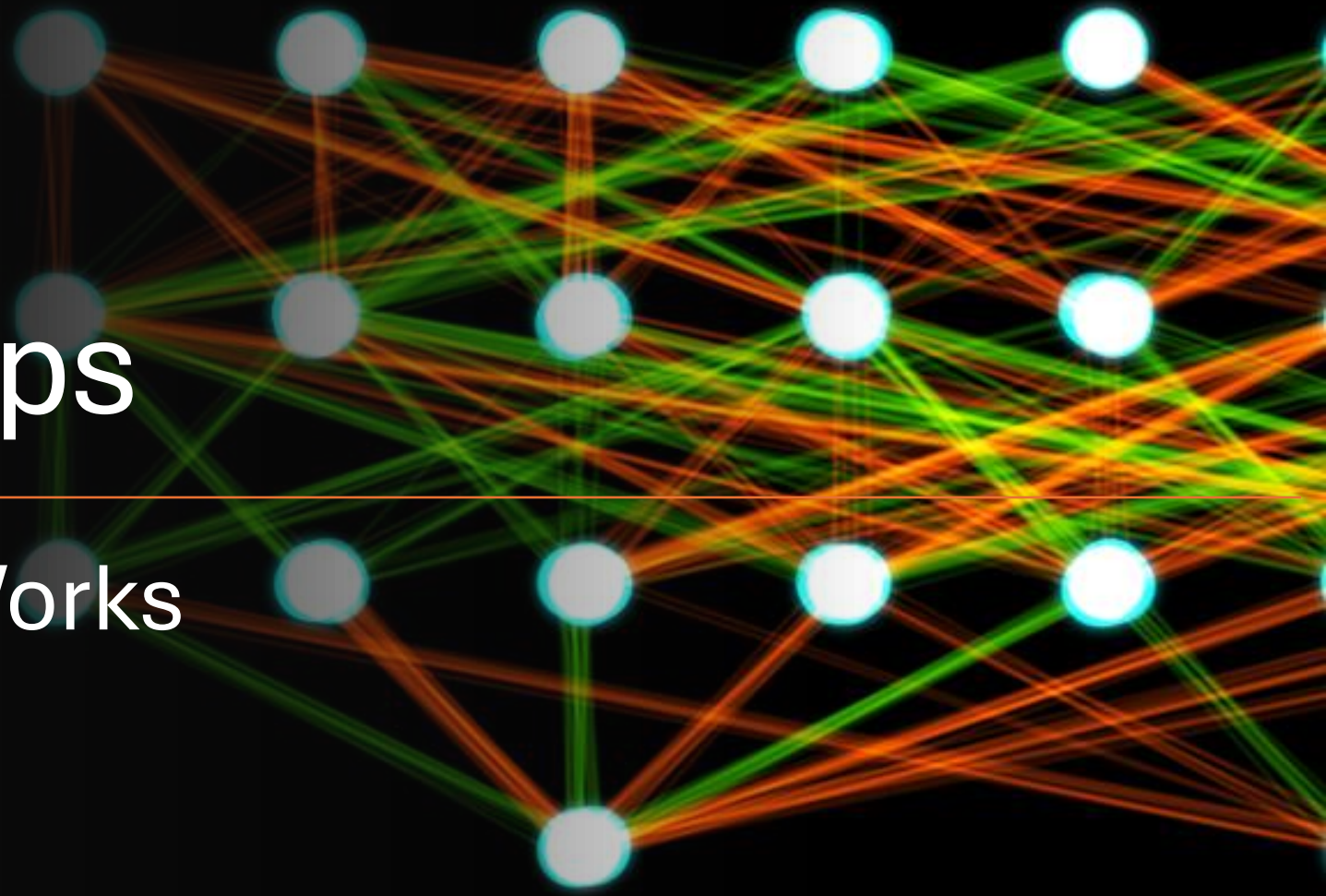
Autocorrect With Extra Steps

How Generative AI Works



Ian Burch
UNM Valencia

Why Learn How AI Works?

- Hype & anxiety make it hard to understand its capabilities and limitations
- Understanding makes it easier to detect and work with
- The tech and math is awesome

History of Language Models

- Research started in 1960s at MIT
- Early generators used strict rules
- ELIZA was a simulated therapist – responded to keywords
- SHRDLU could take limited actions in a virtual word

Statistical Generators

- Rule-based systems had obvious patterns & limitations
- Impossible to write a rule for every situation
- In the early 90s, research shifted to statistical methods
- Instead of writing rules, build systems to analyze and mimic existing text

Markov Chains

- Markov Chains are the simplest text generators
- Train it on a source by going word-by-word
- Build a table of what follows each specific word
- Build a sentence by choosing a word, then choosing the next word from the table

Markov Chain Example

Test Case: *Green Eggs & Ham* by Dr Seuss

Limited 50-word vocabulary makes a great test case

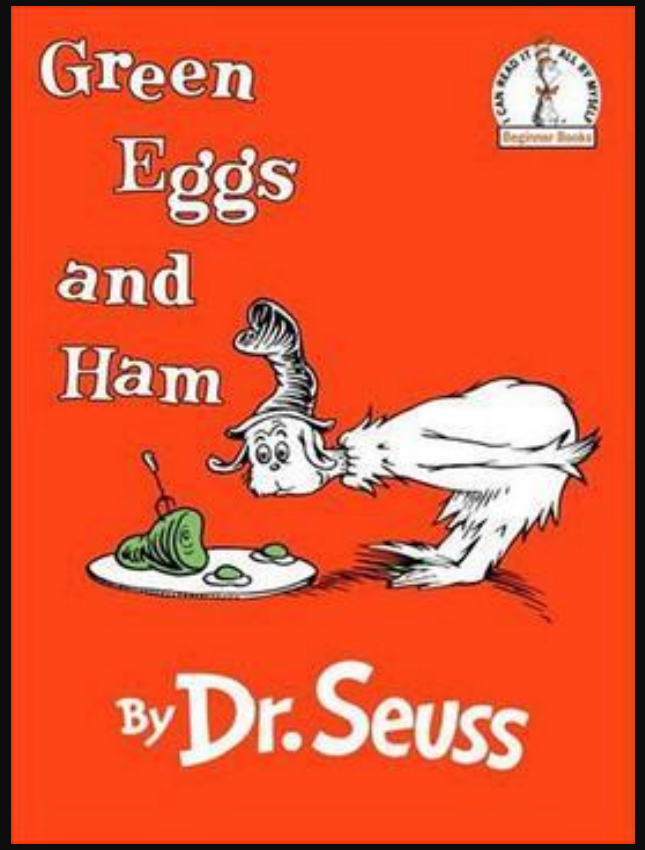


Table ["Green"]
- 100% "Eggs"

Table ["Could"]
- 48.2% "Not"
- 44.4% "You"
- 7.4% "Eat"

Table ["Ham"]
- 100% End Sentence

Improvements

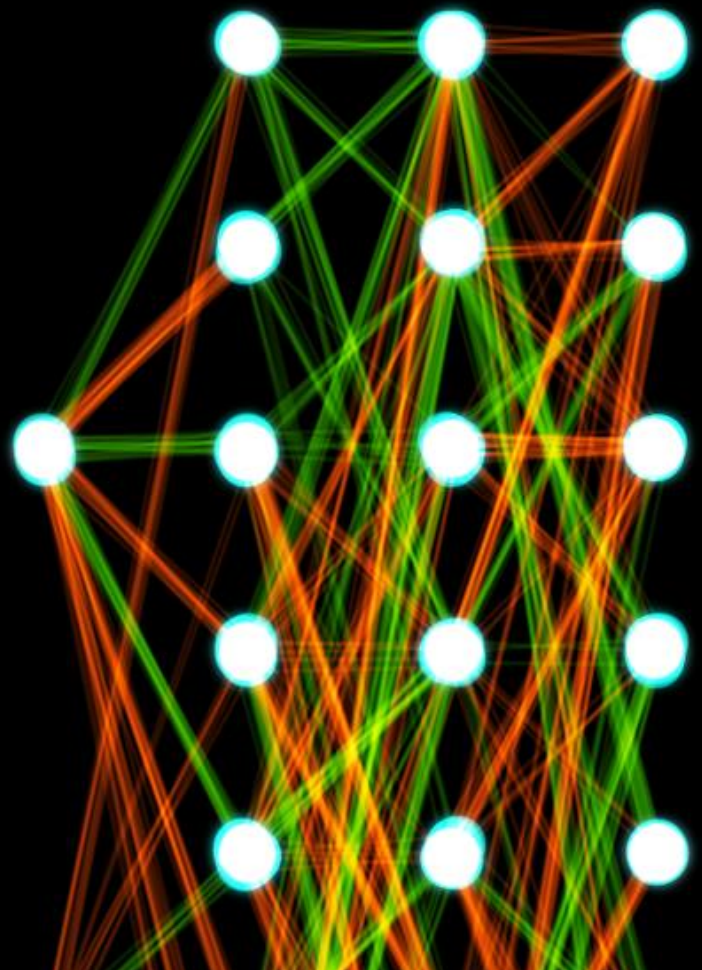
PROBLEMS:

- Can only use word patterns it's seen before
- Common phrases are randomized
- Words have no context

SOLUTIONS:

- Add more sources
- Look at multiple words, not just one
  - Ex: "Eggs and" always gives "Ham"
- Code some manual rules – parts of speech, common phrases

Neural Networks



- In the early 2000s, researches tried structuring systems like neurons
- Add more inputs, multiple layers, and random results to the lists
- Give feedback on what is good output vs bad output
- Let it tweak its own rules based on feedback

Large Language Models (Chat-GPT)

- Adds memory, attention, and context
- Context - Looks at prompt to limit data to the topic
- Attention - Focuses on relevant keywords
- Memory – Adjusts word probabilities as it goes for consistency

The *Large* in LLM

- LLMs require a huge amount of data to pretrain their neural networks
- Chat-GPT 4 cost >$100 million to train
- Training data sets can reach 500TB
- LLMs use the same energy as 15,000 - 30,000 U.S. household each day

What Does All This Mean?

- Each word is calculated by massive amounts of math calculated from massive amounts of examples
- Results become based on averages, giving safe results
- Does not have actual creativity or reasoning
  (Only remixes from existing works or examples)
- Will prioritize things that look right instead of being right
  (Citations are especially prone to this)

Valid Uses

- Search engine for common information
- Inspiration
- An outline of an essay
- A common consensus or outline of a topic

Bad Uses

- Searching for specific information or jargon
- Creative work
- An uncommon argument
- Anytime you need 100% accurate information

Detecting AI Works

- Arguments are "safe". Gives both sides instead of persuading
- Padding phrases: "delve into", "a tapestry of", "in the realm of"
- For art, fine details are missing or excessive. Look at hands & feet
- For photos, the scene may be too perfect

Conclusions

- Modern AI generators use tons of data and math to create content
- Fine for inspiration or searching for common, non-critical info
- Focuses on looking right over being right

Questions?