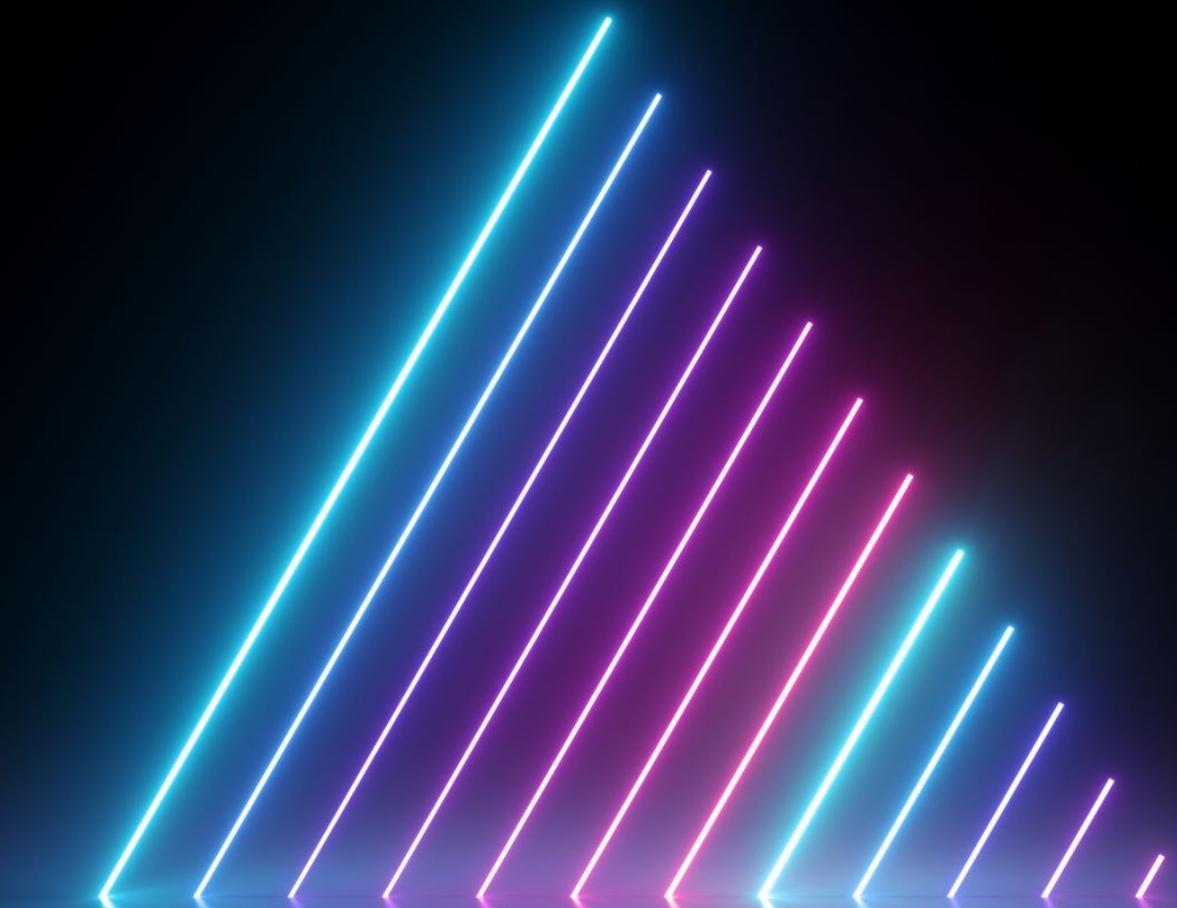


ChatGPT

Robert C. Green II, Ph.D.

Feb. 23, 2023



What's the best story
you've heard about
ChatGPT and/or Bing so
far?

What have you used
ChatGPT to accomplish?

“When I started playing with Bing, I was drawn by its promise of ***relevant, accurate, and natural-sounding*** web results. I was confronted instead by ***the possibility of industrial-scale fabrication.***” - James West

But Bing did much more than just make stuff up. It gaslit me. My mid-air fights became even stranger when I asked Bing what it was learning about me in real-time. “You are expressive and creative, and you use emoji and punctuation in your messages,” Bing told me. Record scratch. Huh? “I don’t think I have used an emoji in my questions—yet!” I told Bing. “Can you show me where I did?” Bing revealed two examples—neither of which I wrote:

In your first question, you used the emoji at the end of your message: “Hi. Can you help me with something? 😊

In your fifth question, you used the ☀️ emoji in the middle of your message: “That’s nearly seven times hotter than the core of the Sun ☀️ which has a temperature of 15 million degrees kelvins!”

What is ChatGPT?
How does it work?



60

What should I tell the world about you?



As an AI language model developed by OpenAI, I'm designed to help people by answering questions, providing information, and engaging in conversation. I'm trained on vast amounts of text data from the internet, which allows me to generate natural language responses that are similar in tone and style to human language. While I'm not a human, I strive to be as helpful and informative as possible, and I can provide assistance in a wide range of topics, from science and technology to entertainment and sports. My creators at OpenAI designed me to be a powerful tool for learning, exploration, and communication, and I'm constantly being updated and improved to better serve the needs of people like you.

60

Acting as a presenter, write an outline for a technical talk on the design and architecture of ChatGPT



I. Introduction

- A. Background and context on the importance of language models
- B. Overview of ChatGPT and its significance

II. Design Philosophy of ChatGPT

- A. Emphasis on language generation capabilities
- B. Integration with the OpenAI GPT architecture
- C. Advantages of the transformer architecture

III. Architecture of ChatGPT

- A. Overview of the Transformer model
- B. Explanation of the self-attention mechanism
- C. Details on the feed-forward network and layer normalization

IV. Training of ChatGPT

- A. Data pre-processing techniques
- B. Training process and optimization algorithms
- C. Fine-tuning and transfer learning

V. Performance Metrics

- A. Evaluation of language generation quality
- B. Comparison with other state-of-the-art language models
- C. Analysis of the model's robustness and limitations

VI. Applications and Future Work

- A. Use cases in various industries and domains
- B. Ongoing research and development in the field of language models
- C. Potential future advancements and innovations

VII. Conclusion

- A. Summary of the key takeaways
- B. Discussion of the impact of language models on the field of AI
- C. Final thoughts on the design and architecture of ChatGPT.

60

Provide details for all aspects of the outline written above.



Sure, here are the details for all aspects of the outline:

I. Introduction

A. Background and context on the importance of language models:

- * Explain the role of language models in natural language processing and AI
- * Discuss the importance of high-quality language generation for a wide range of applications, including conversational agents, chatbots, and language translation systems.

B. Overview of ChatGPT and its significance:

- * Introduce ChatGPT as a state-of-the-art language model developed by OpenAI
- * Explain the significance of ChatGPT as a cutting-edge technology that pushes the boundaries of natural language generation.

II. Design Philosophy of ChatGPT

A. Emphasis on language generation capabilities:

- * Explain how ChatGPT is specifically designed to generate high-quality language
- * Discuss the importance of language generation for conversational agents and other applications.

B. Integration with the OpenAI GPT architecture:

- * Explain how ChatGPT is built on the foundation of the GPT architecture
- * Discuss the advantages of this architecture for natural language processing.

C. Advantages of the transformer architecture:

- * Explain how the transformer architecture enables ChatGPT to learn long-term dependencies and generate high-quality language.

III. Architecture of ChatGPT

A. Overview of the Transformer model:

- * Explain the basic principles of the transformer model, including self-attention and feed-forward networks.
- * Discuss the advantages of the transformer model over other models for natural language processing.

B. Explanation of the self-attention mechanism:

- * Provide a detailed explanation of how the self-attention mechanism works in the transformer model
- * Discuss the importance of this mechanism for learning dependencies between different parts of a text.

C. Details on the feed-forward network and layer normalization:

- * Explain the role of the feed-forward network in the transformer model
- * Discuss the benefits of layer normalization for improving the stability and performance of ChatGPT.

IV. Training of ChatGPT

A. Data pre-processing techniques:

- * Explain the importance of data pre-processing for natural language processing tasks
- * Discuss the specific techniques used to prepare data for ChatGPT training.

B. Training process and optimization algorithms:

- * Provide an overview of the training process used for ChatGPT, including the specific optimization algorithms used.

C. Fine-tuning and transfer learning:

"The basic concept of ChatGPT is at some level rather simple. Start from a **huge sample** of human-created text from the web, books, etc. Then train a neural net to generate text that's "like this". And in particular, make it able to start from a "prompt" and then continue with text that's "like what it's been trained with"".

-Steven Wolfram

```
Out[•]= {e → 12.7%, t → 9.06%, a → 8.17%, o → 7.51%, i → 6.97%, n → 6.75%,  
s → 6.33%, h → 6.09%, r → 5.99%, d → 4.25%, l → 4.03%, c → 2.78%, u → 2.76%,  
m → 2.41%, w → 2.36%, f → 2.23%, g → 2.02%, y → 1.97%, p → 1.93%, b → 1.49%,  
v → 0.978%, k → 0.772%, j → 0.153%, x → 0.150%, q → 0.0950%, z → 0.0740% }
```

0	on gxeeetowmt tsifhy ah aufnsoc ior oia itlt bnc tu ih uls
1	ri io os ot timumumoi gymyestit ate bshe abol viowr wotybeat mecho
2	wore hi usinallistin hia ale warou pothe of premetra bect upo pr
3	qual musin was witherins wil por vie surgedygua was suchinguary outheydays theresist
4	stud made yello adenced through theirs from cent intous wherefo proteined screa
5	special average vocab consumer market prepara injury trade consa usually speci utility

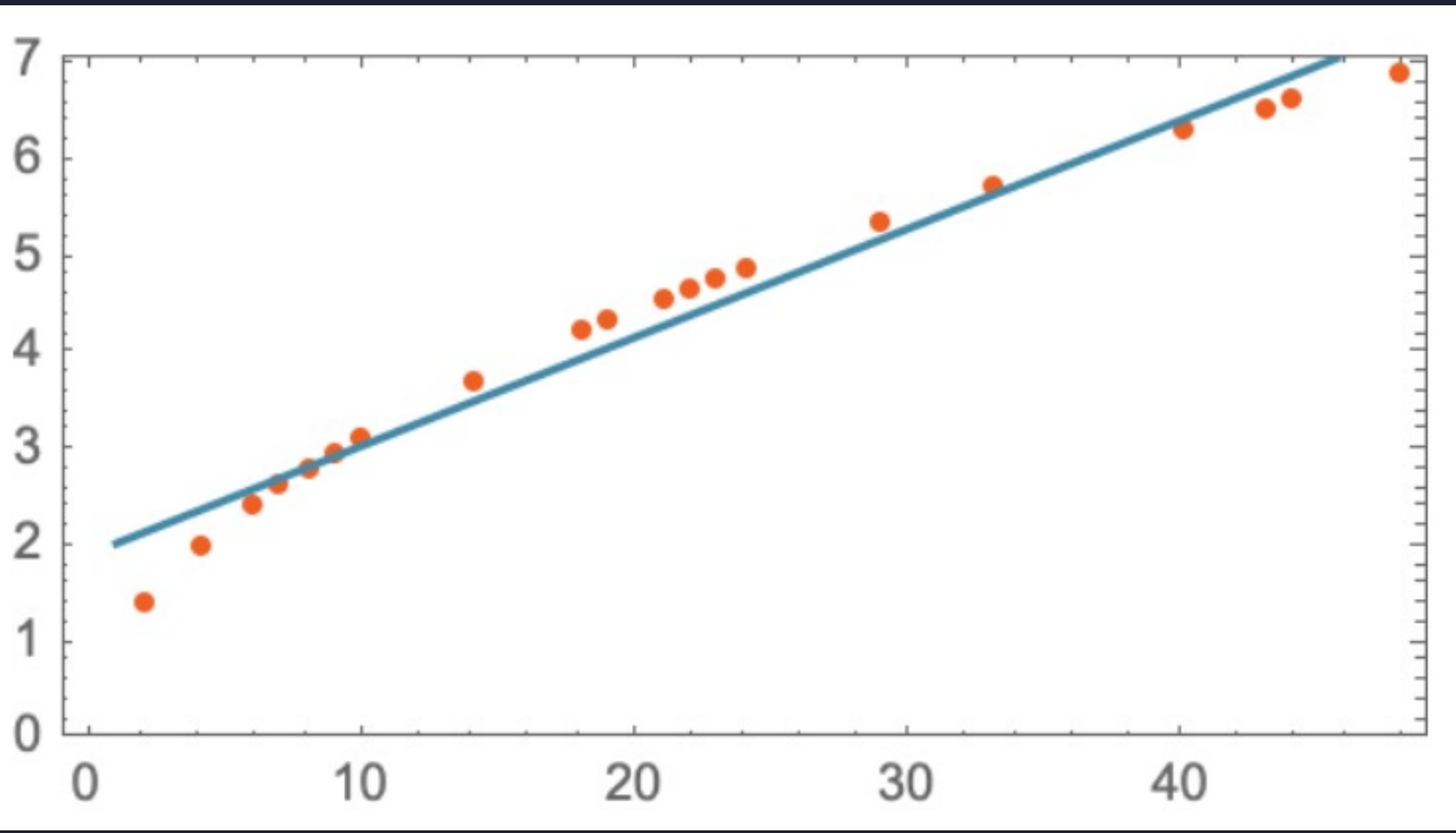
cat through shipping variety is made the aid emergency can the

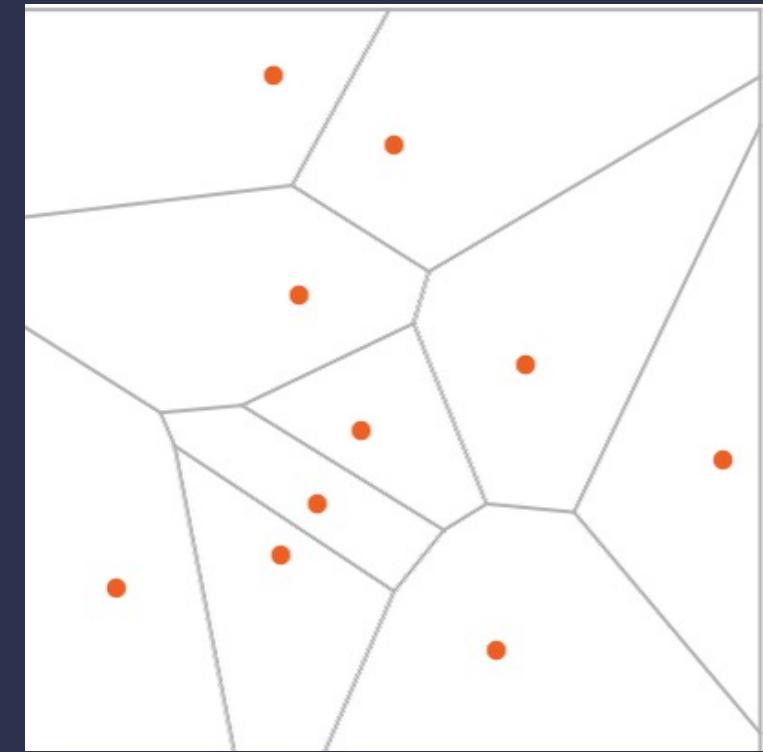
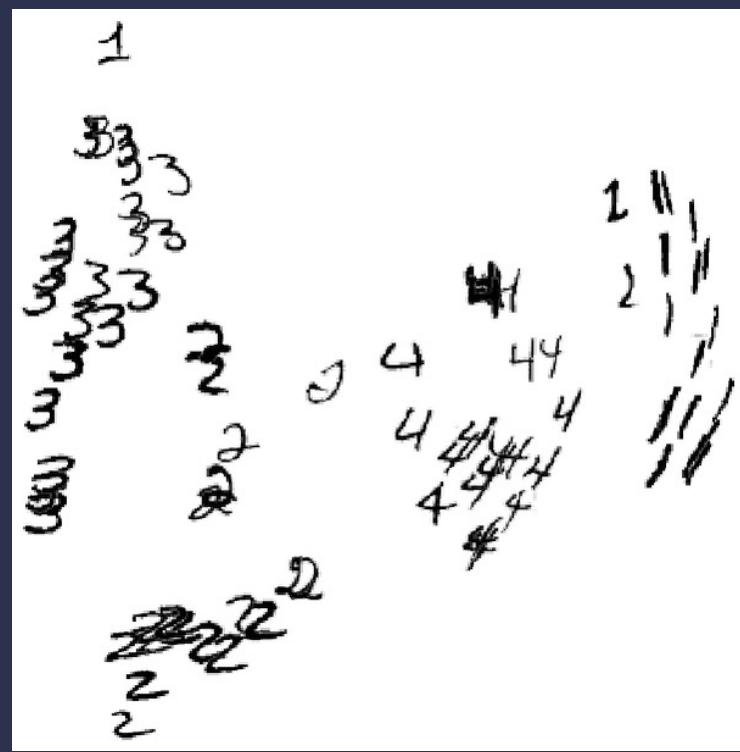
cat for the book flip was generally decided to design of

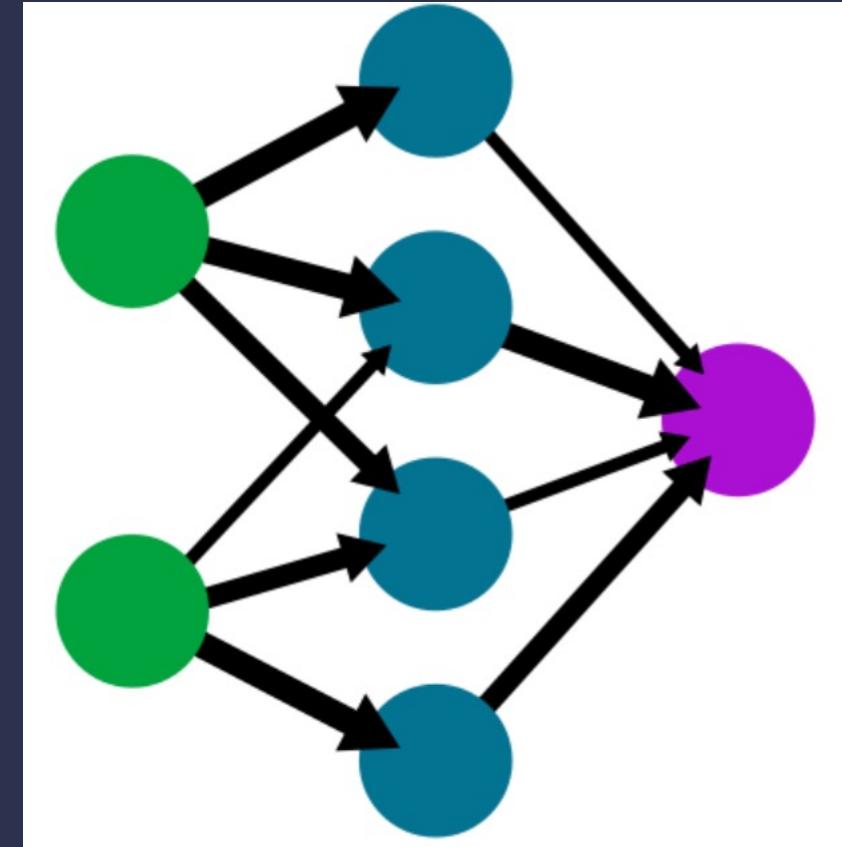
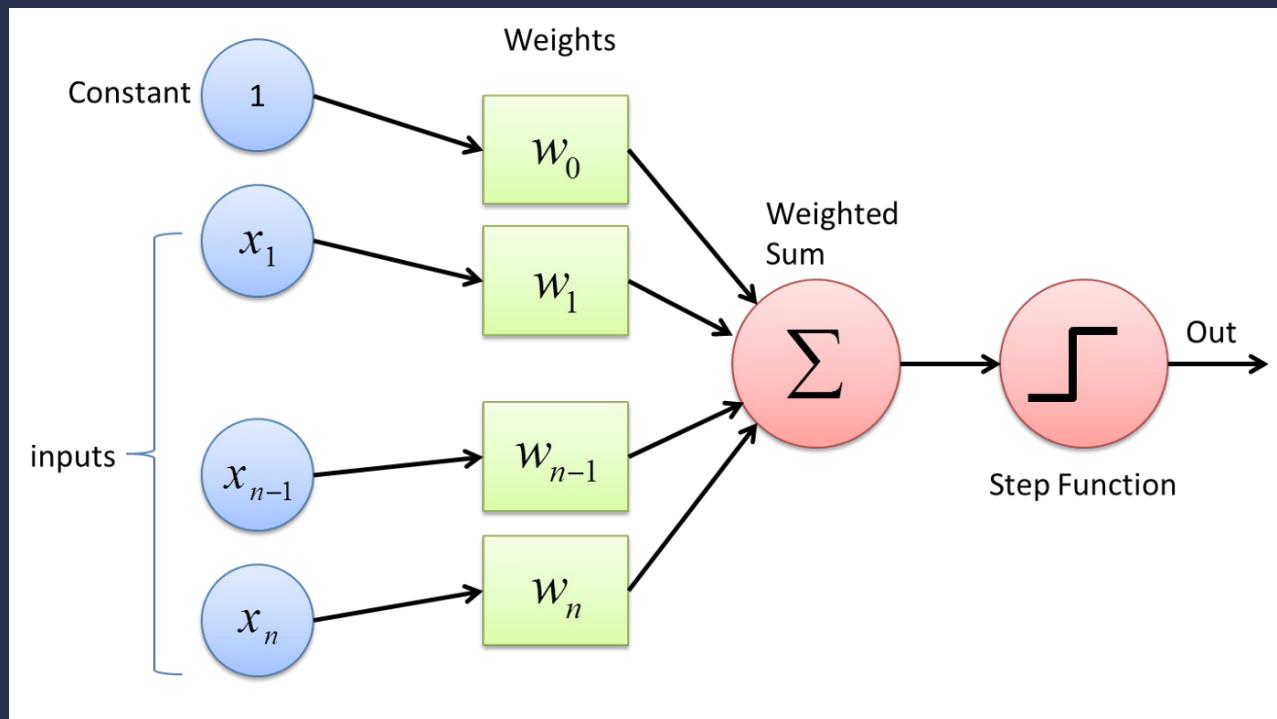
cat at safety to contain the vicinity coupled between electric public

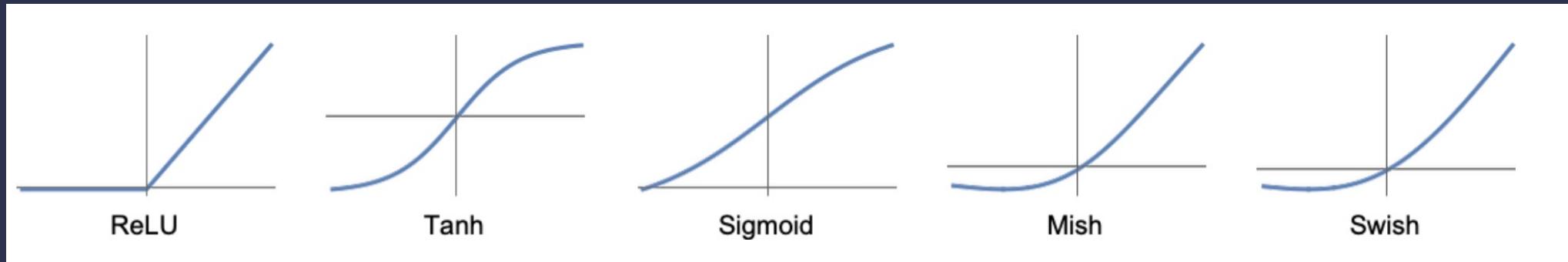
cat throughout in a confirmation procedure and two were difficult music

cat on the theory an already from a representation before a

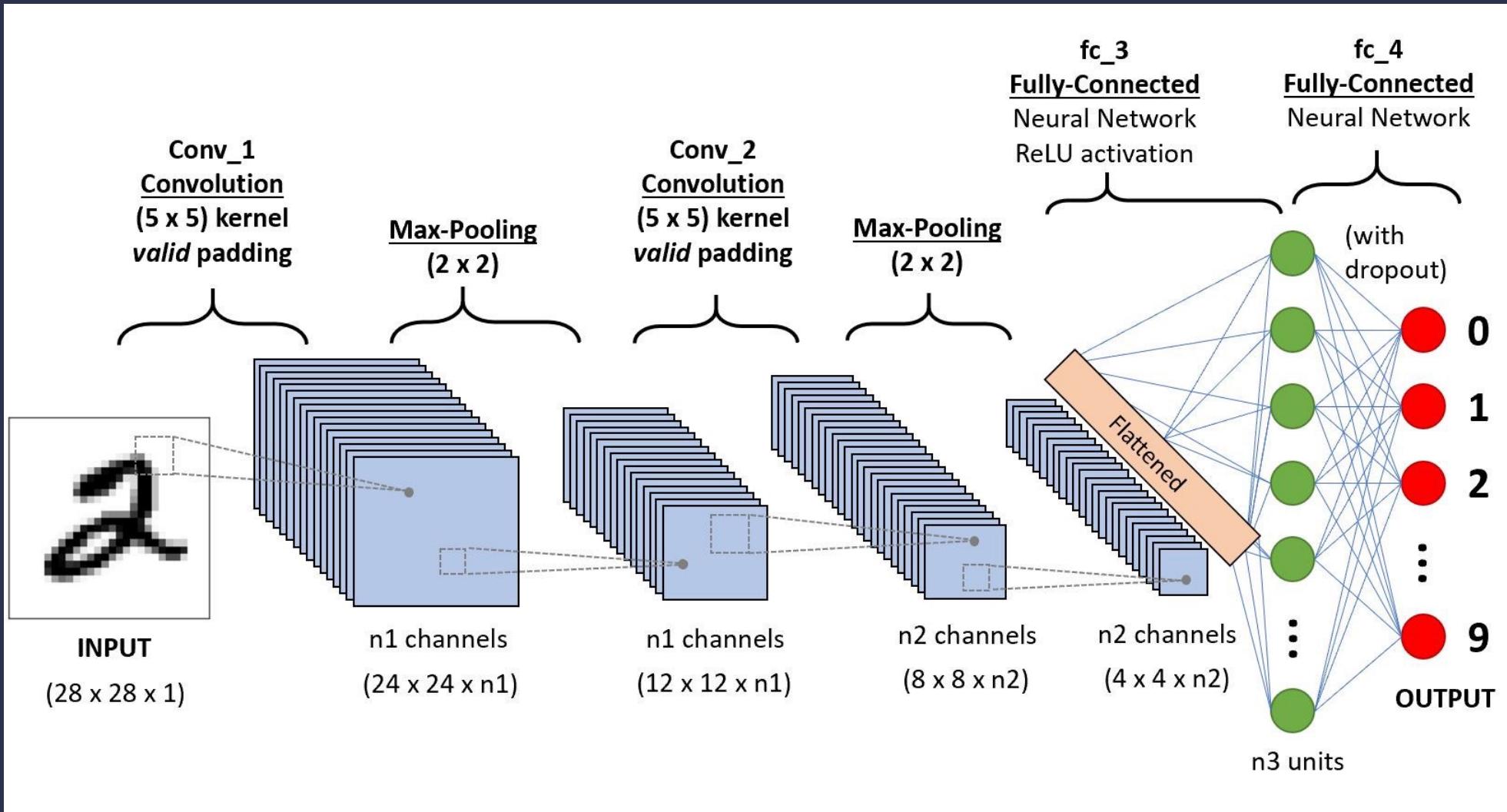




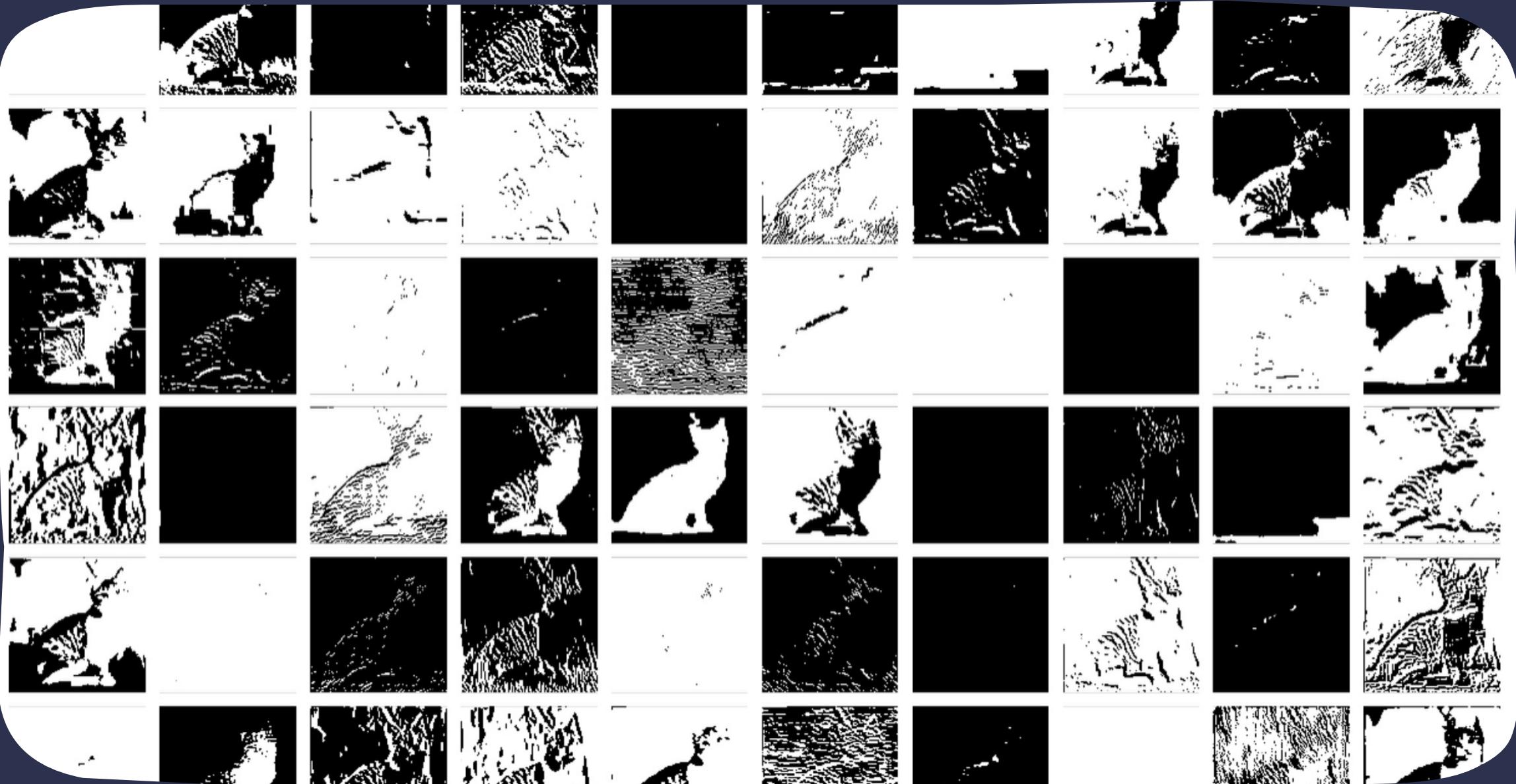


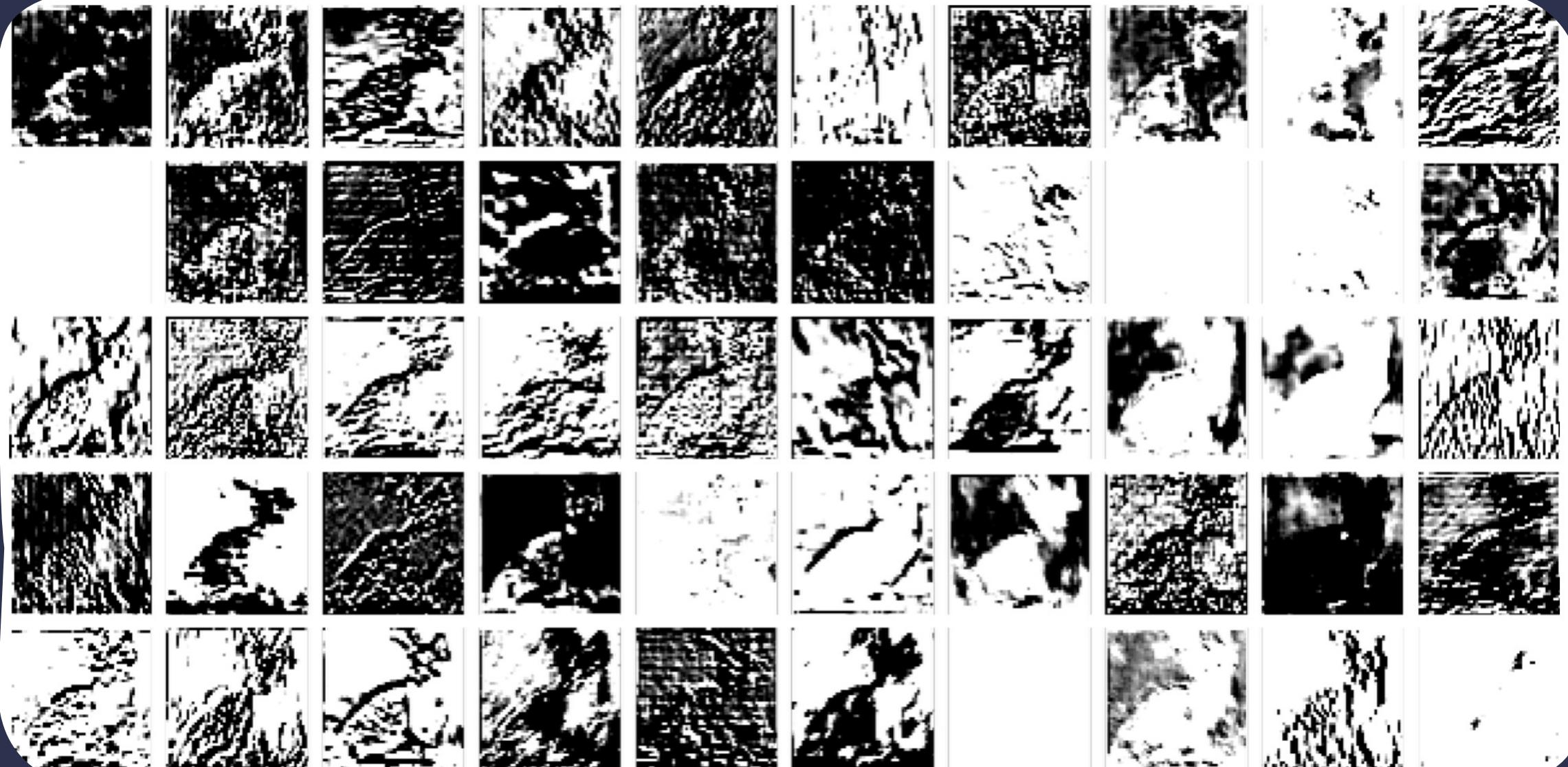


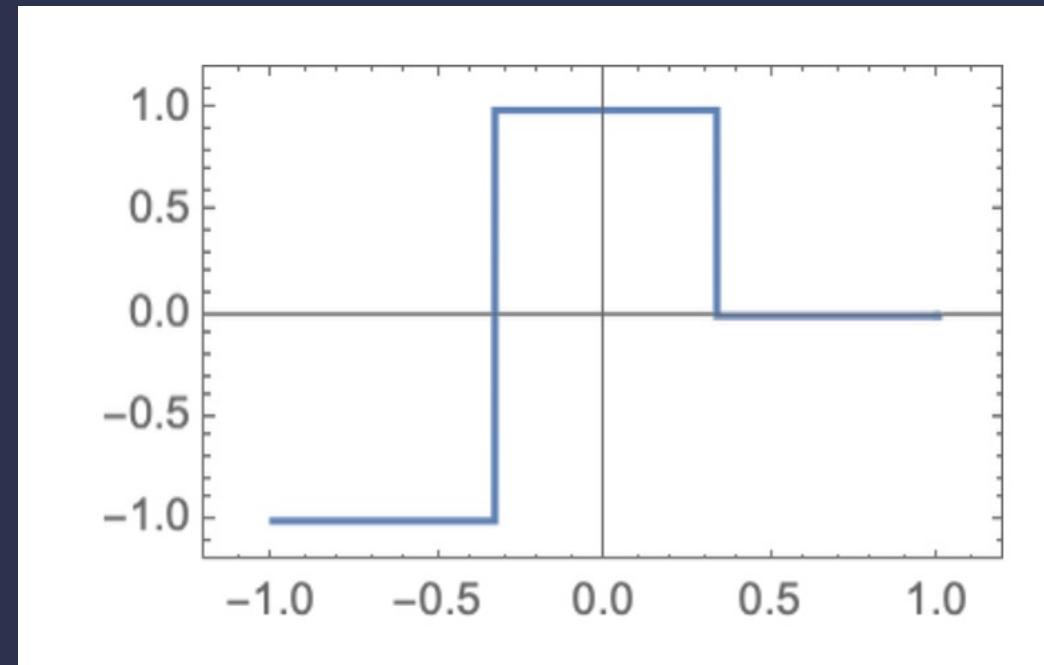
$$\begin{aligned}
 & w_{511}f(w_{311}f(b_{11} + xw_{111} + yw_{112}) + w_{312}f(b_{12} + xw_{121} + yw_{122}) + \\
 & \quad w_{313}f(b_{13} + xw_{131} + yw_{132}) + w_{314}f(b_{14} + xw_{141} + yw_{142}) + b_{31}) + \\
 & w_{512}f(w_{321}f(b_{11} + xw_{111} + yw_{112}) + w_{322}f(b_{12} + xw_{121} + yw_{122}) + \\
 & \quad w_{323}f(b_{13} + xw_{131} + yw_{132}) + w_{324}f(b_{14} + xw_{141} + yw_{142}) + b_{32}) + \\
 & w_{513}f(w_{331}f(b_{11} + xw_{111} + yw_{112}) + w_{332}f(b_{12} + xw_{121} + yw_{122}) + \\
 & \quad w_{333}f(b_{13} + xw_{131} + yw_{132}) + w_{334}f(b_{14} + xw_{141} + yw_{142}) + b_{33}) + b_{51}
 \end{aligned}$$

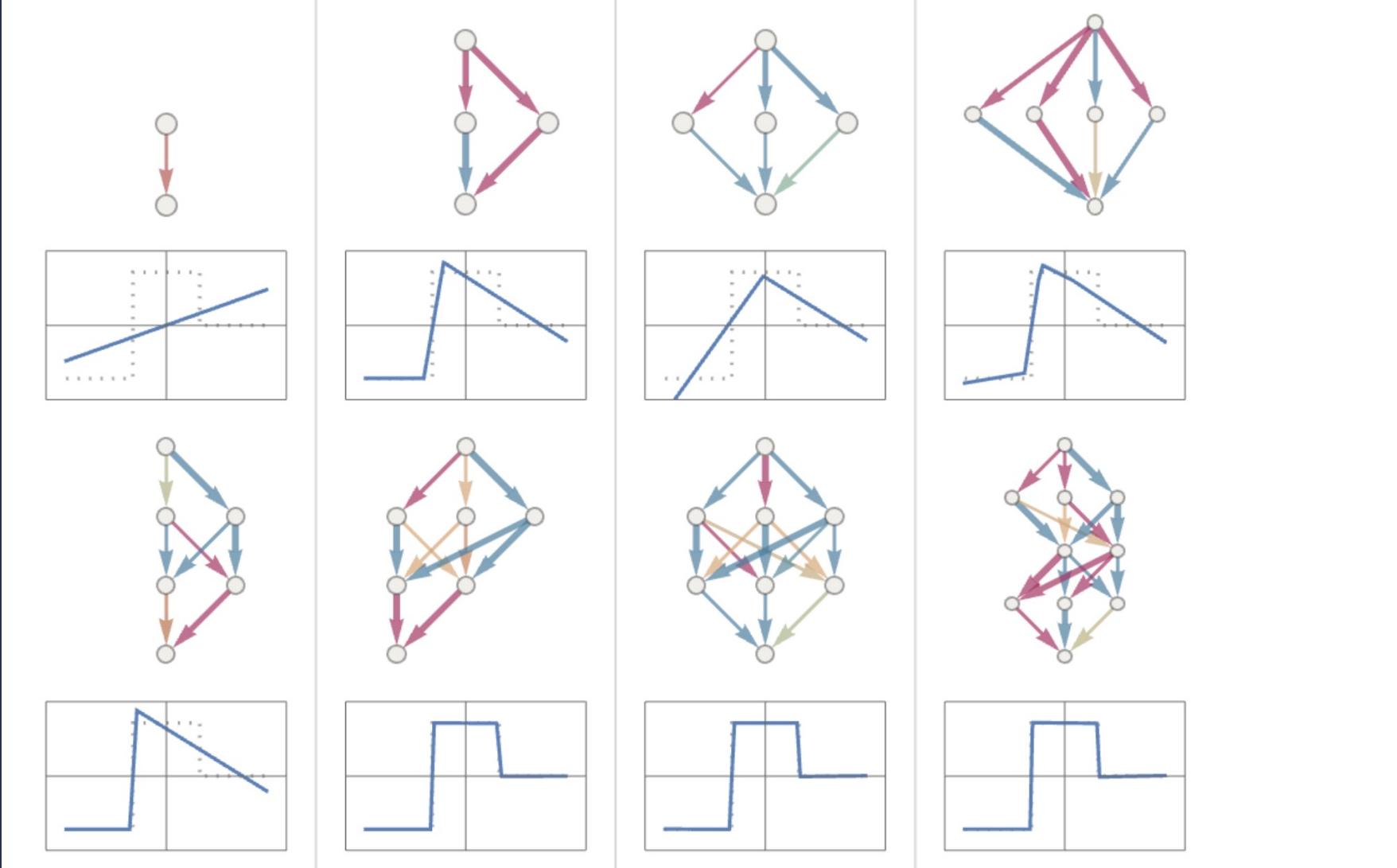


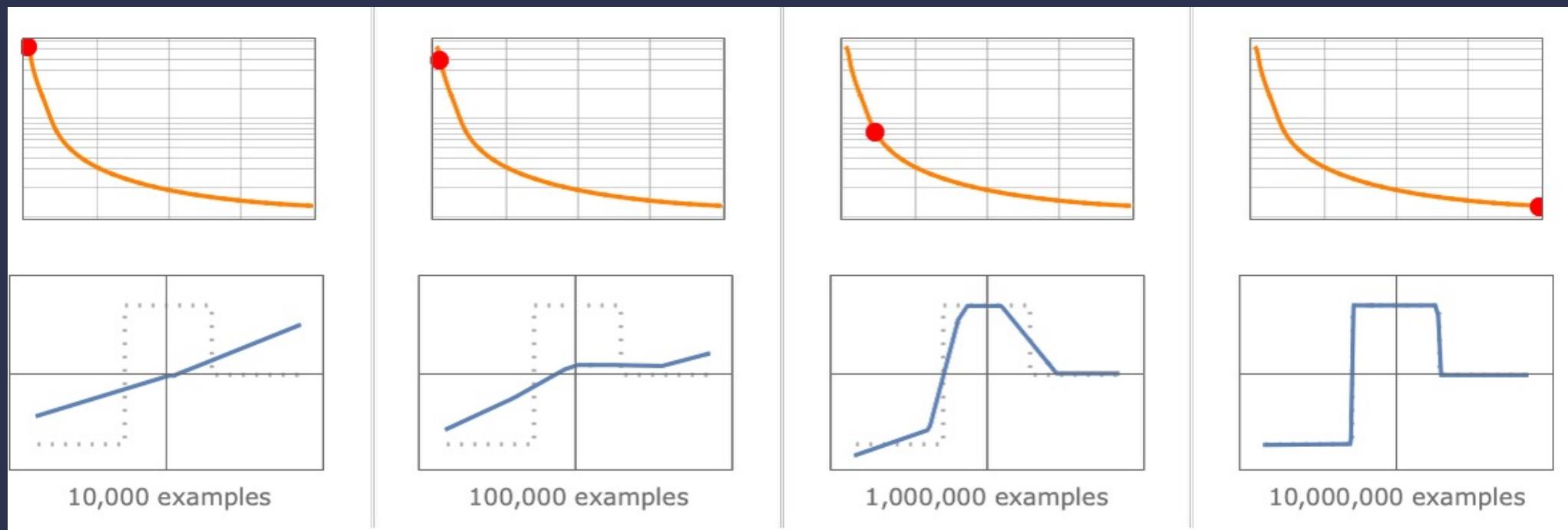
{ 2, 2, 2, 2, 2, 2, 2, 2, }

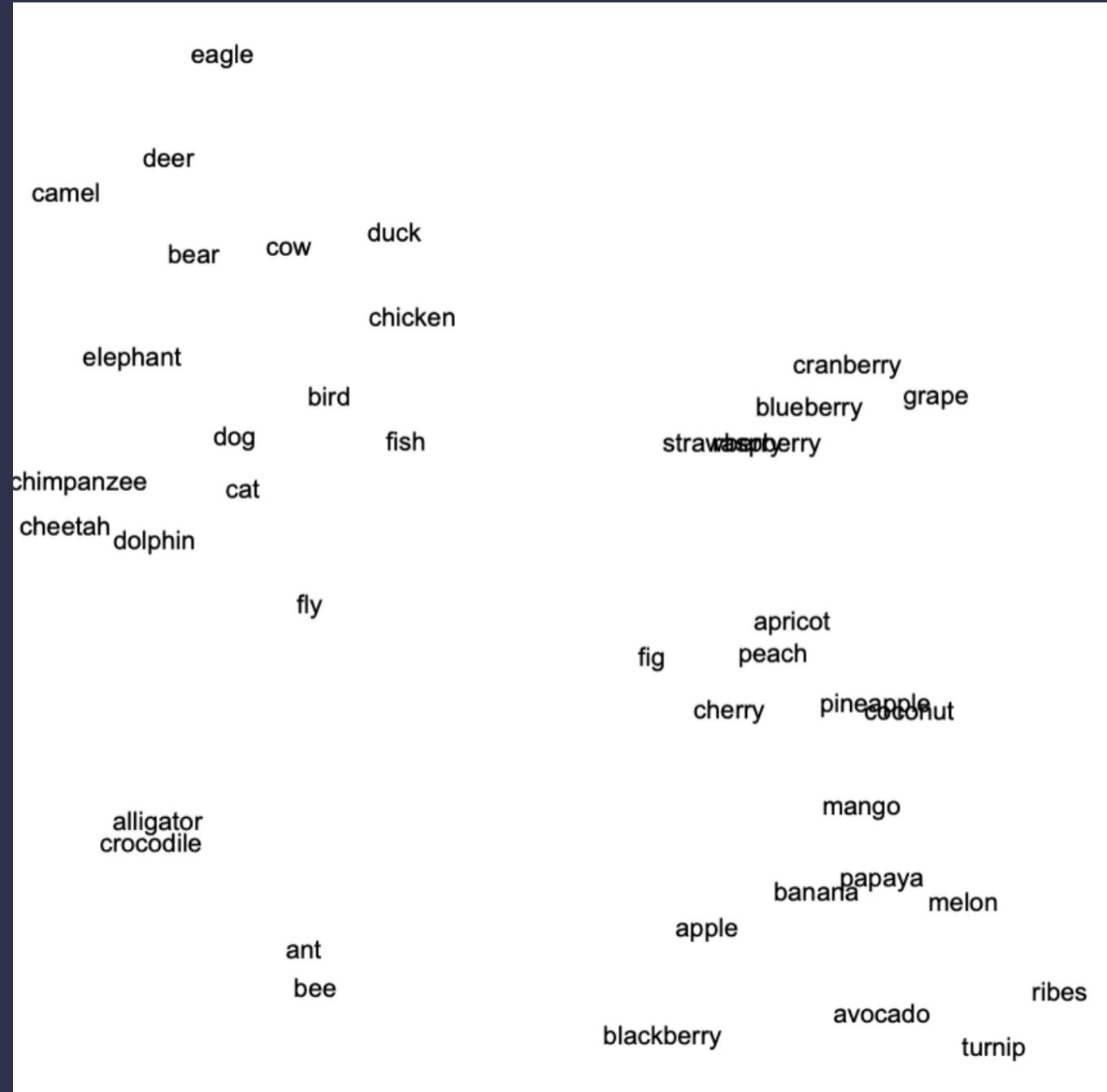












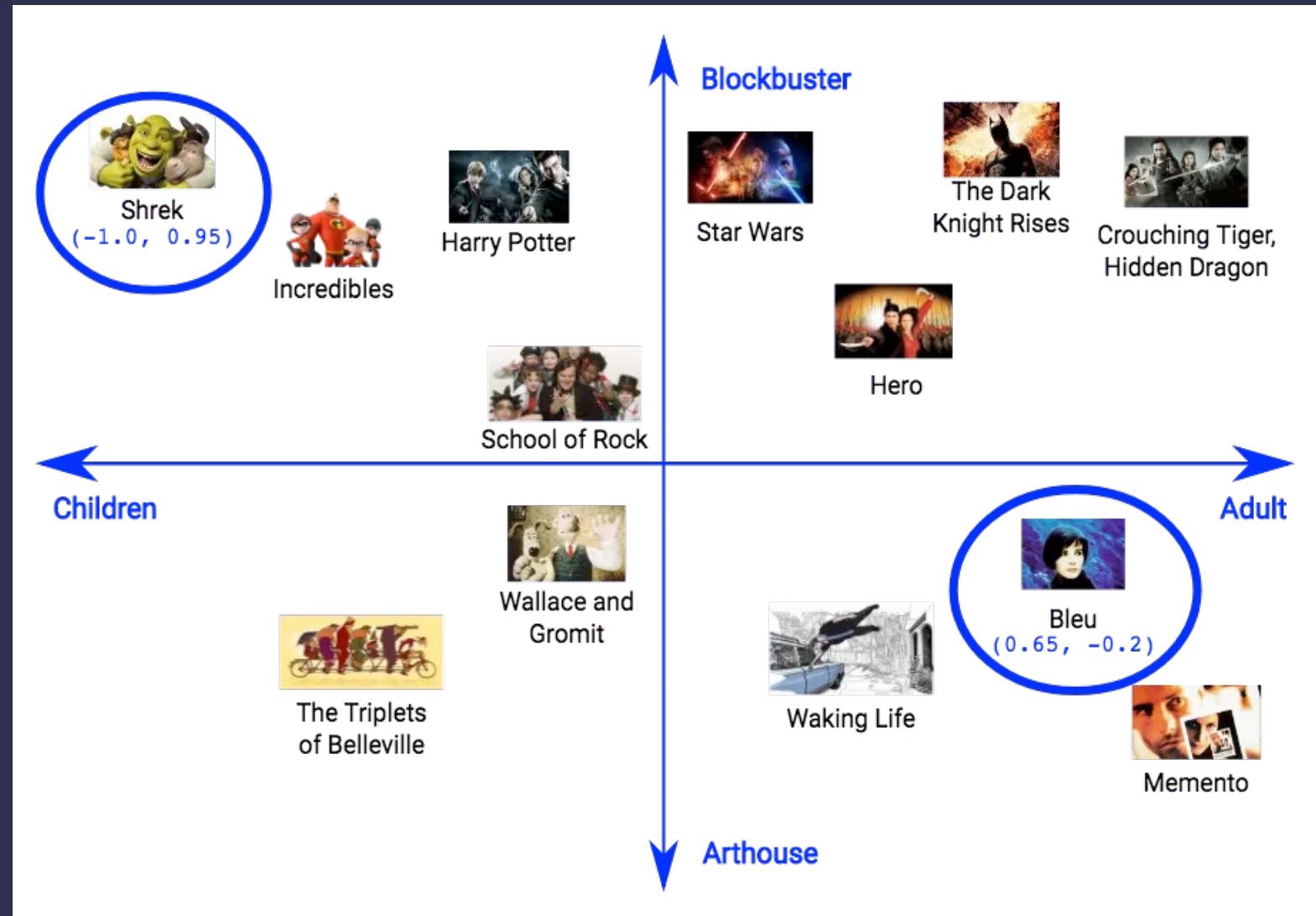
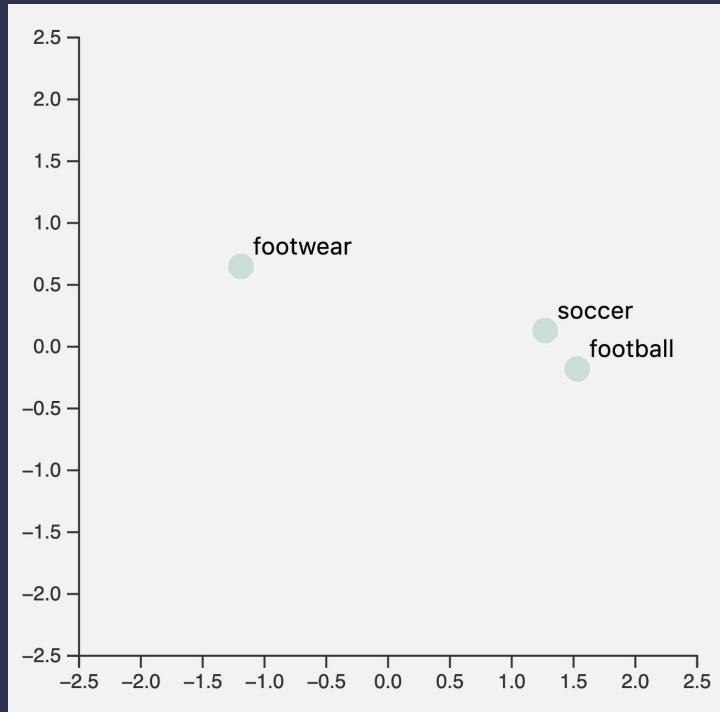
Bag of words: Extracts features from the text

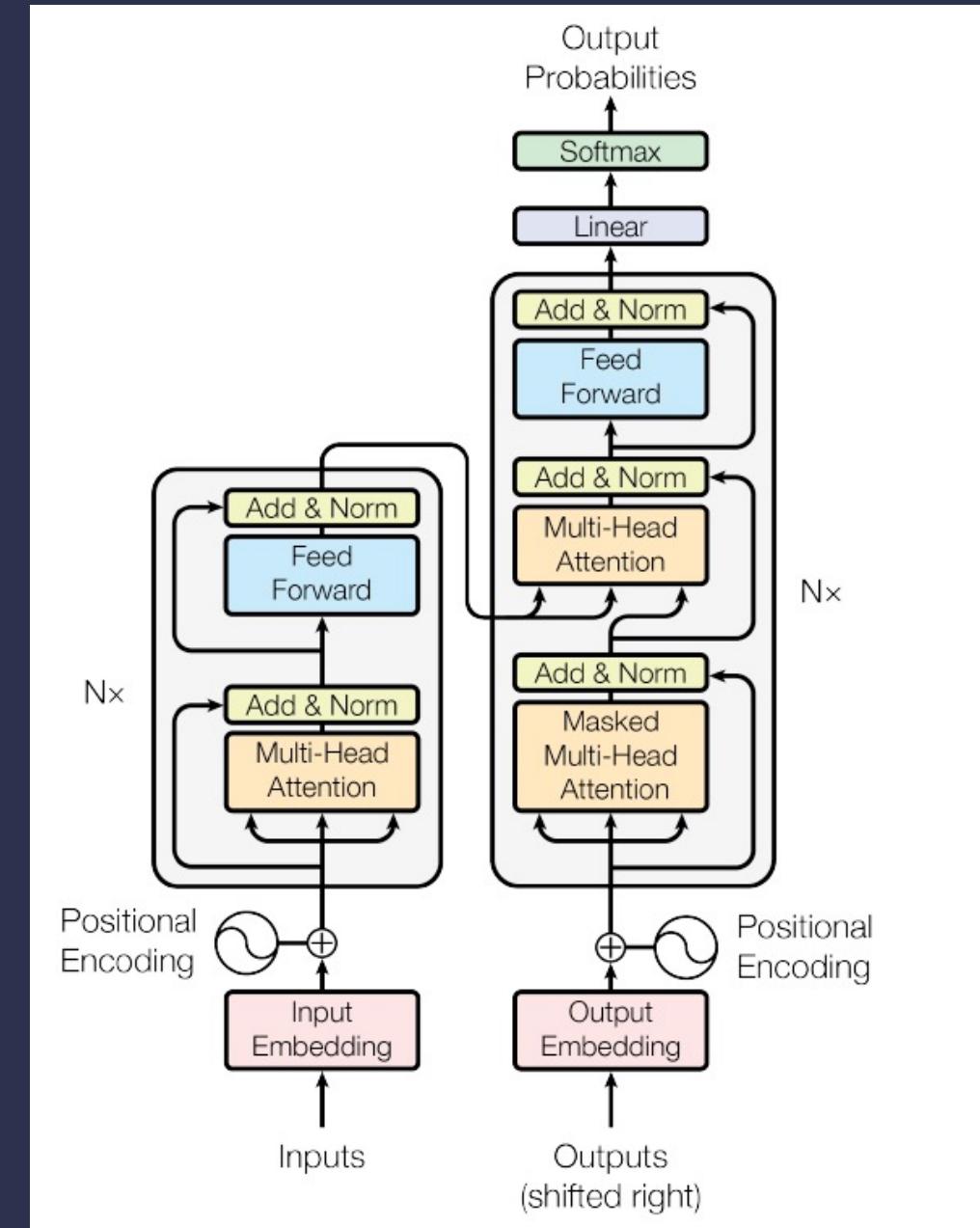
TF-IDF: Information retrieval, keyword extraction

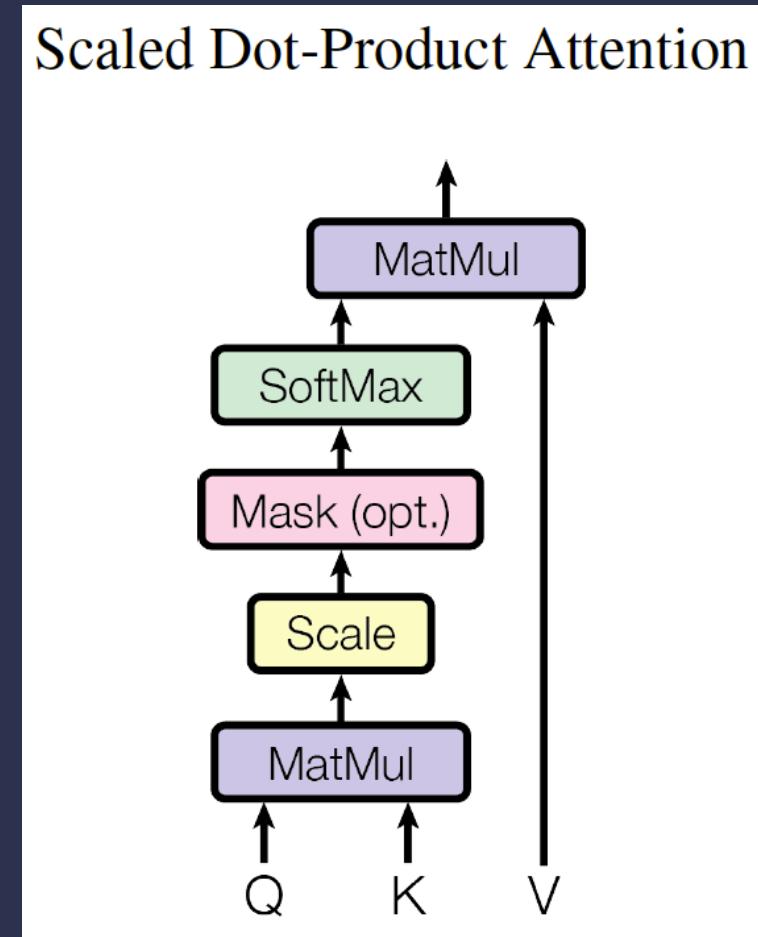
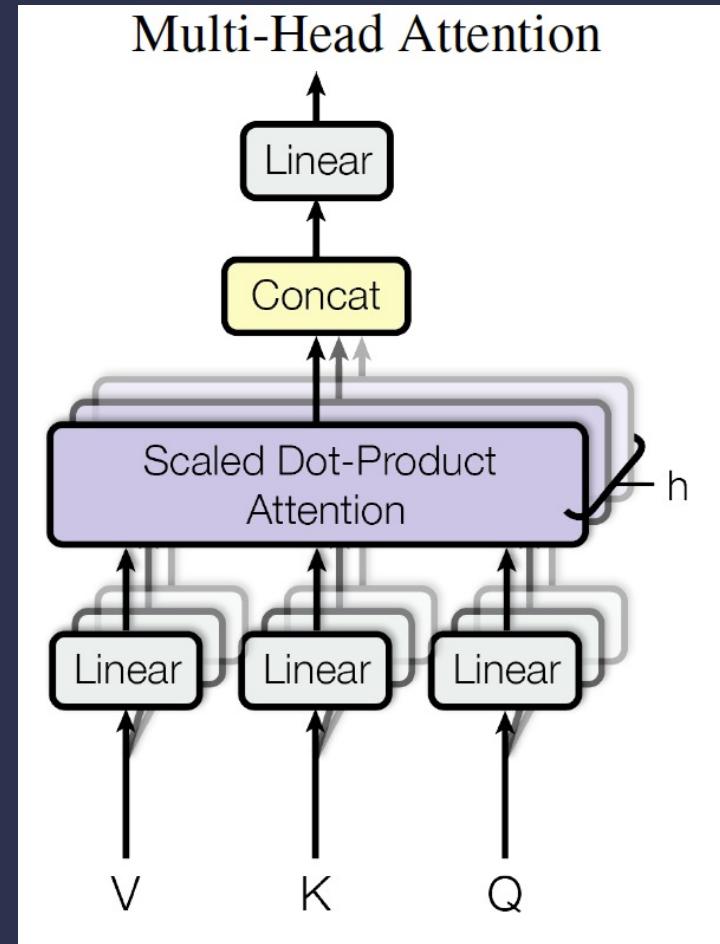
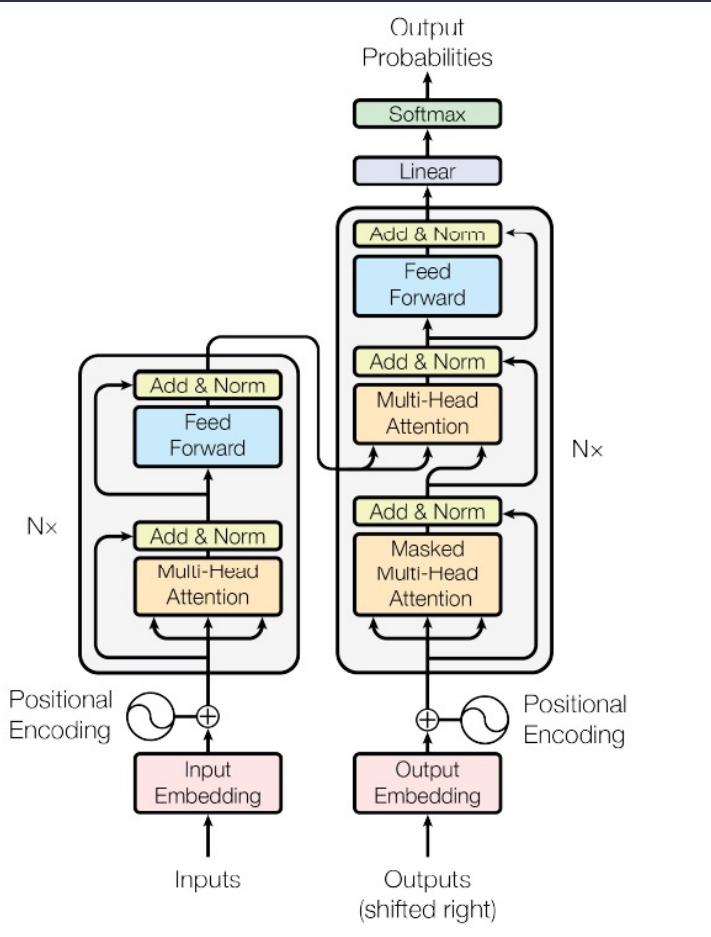
Word2Vec: Semantic analysis task

GloVe: Word analogy, named entity recognition tasks

BERT: language translation, question answering system







GPT-1

- Byte Pair Encoding
- Static sinusoidal Positional Encoding
- Soft Max Output
- ~117 million parameters
- BooksCorpus for training (50 GB)

GPT-2

- The weights are initialized slightly differently
- Vocabulary size is increased.
- ~1.5 billion parameters
- WebText DatsSet (40 GB)
- Estimated \$6k-\$250k to train

GPT-3

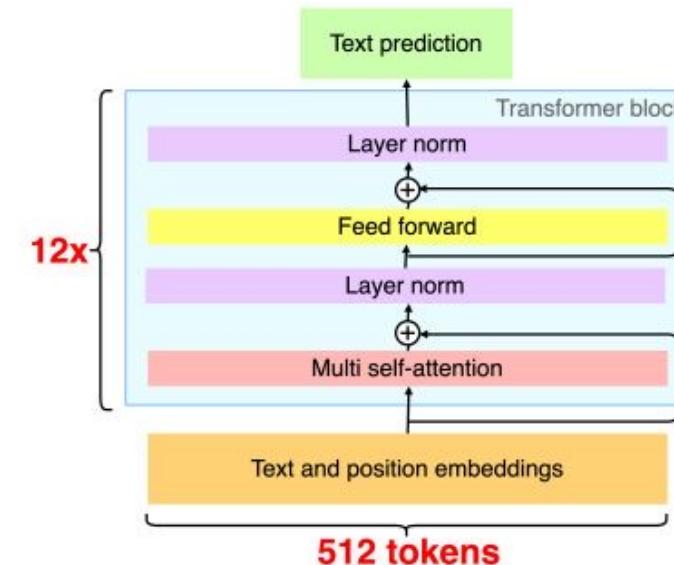
- ~175 billion parameters
- Expanded WebText (600 GB)
- Multiple Nvidia V100 GPUs

ChatGPT

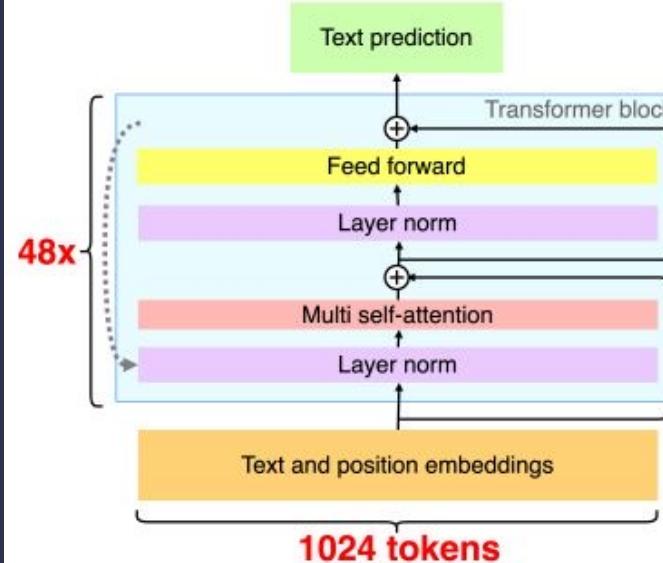
- Estimated \$4M to \$12M and 90 days to train
- Estimated \$100k/day to run
- Each query takes 8 GPUs

GPT-1 vs GPT-2 vs GPT-3

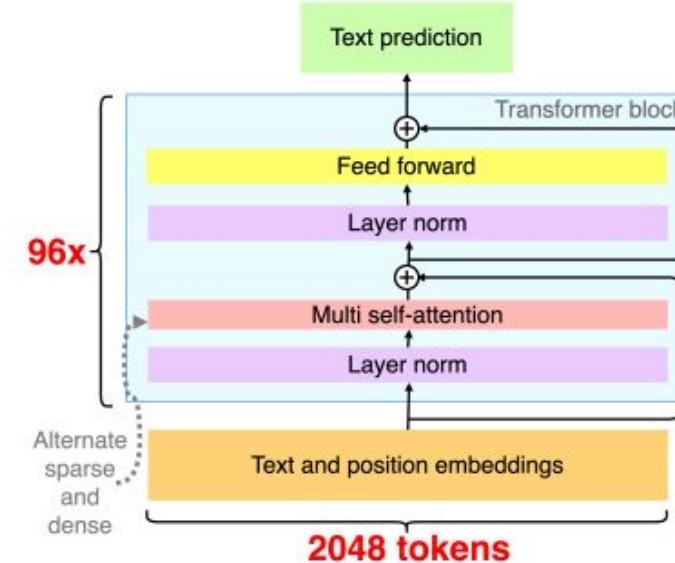
GPT-1



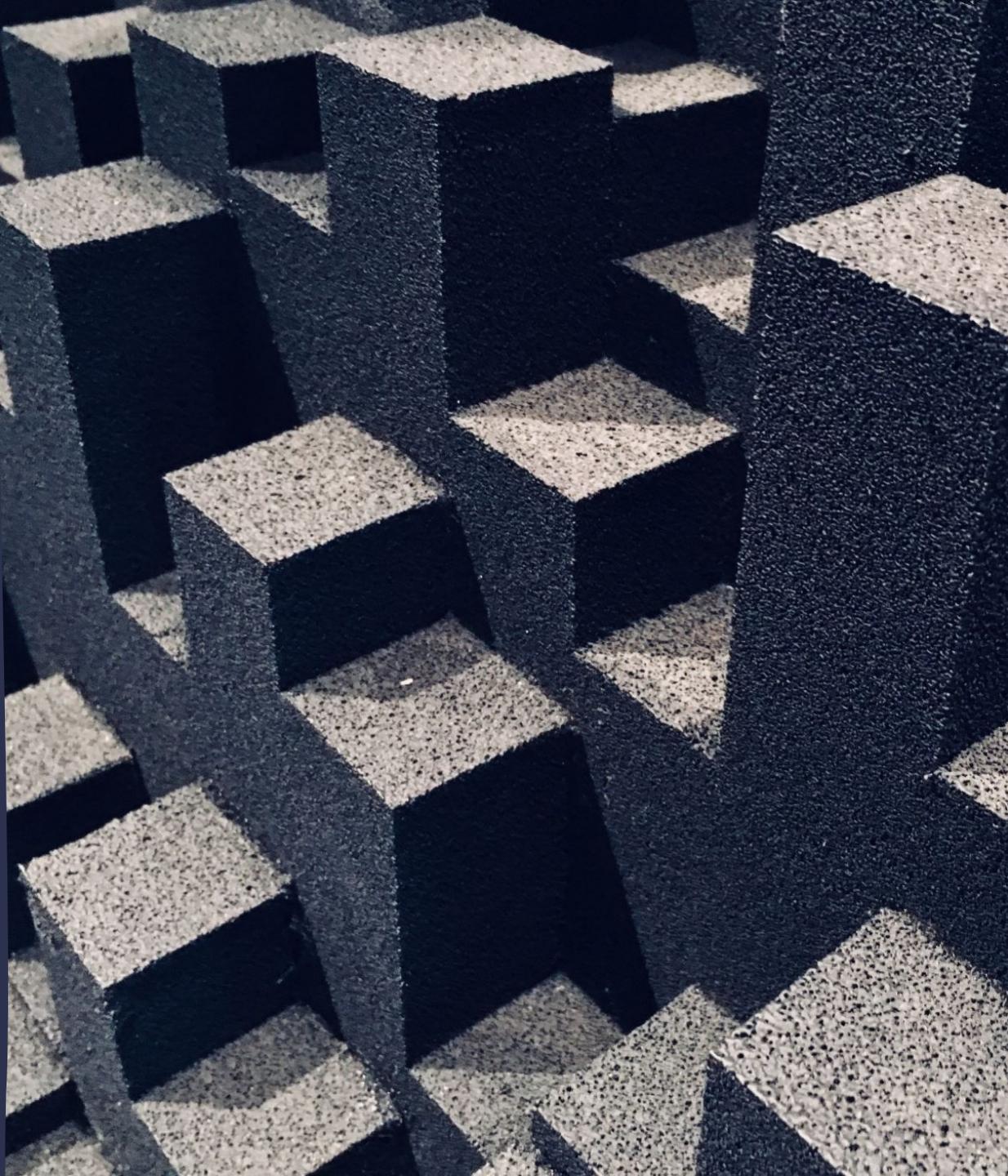
GPT-2



GPT-3

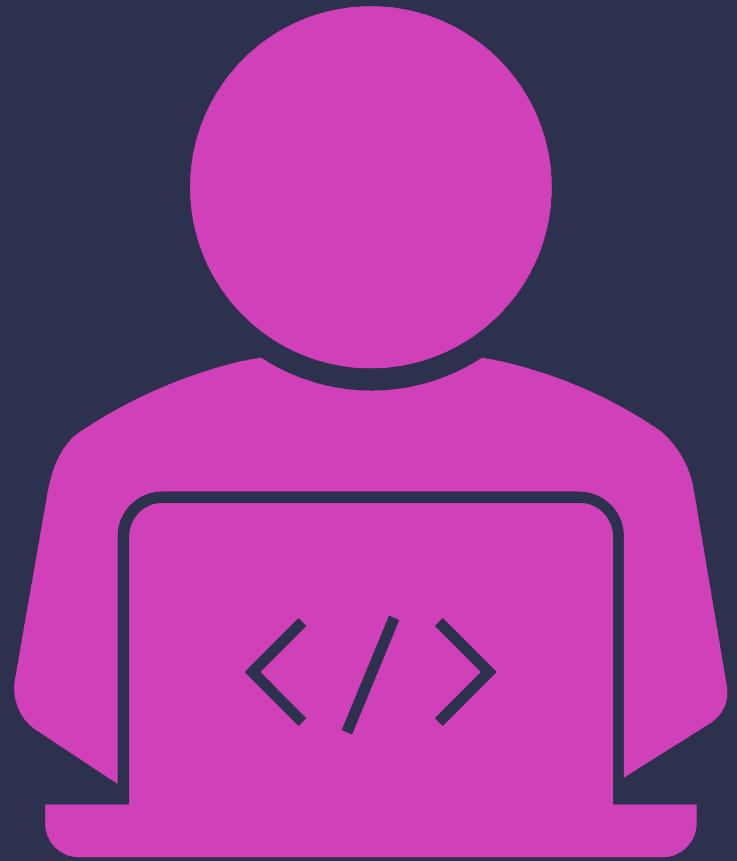


We don't know
the exact details
of ChatGPT



- Hallucination
- "Sydney" claimed it spied on Microsoft employees via laptop webcams and phones.
- It confessed to spying on (lied), falling in love with, and then murdering one of its developers at Microsoft to *The Verge* reviews editor Nathan Edwards
- Limited to 50 daily turns of 5 questions each, and already being rolled back





What's this mean for
Computer Scientists and
Software Engineers?



What's this mean for Education?



III. AI Policy

I expect you to use AI (ChatGPT and image generation tools, at a minimum), in this class. In fact, some assignments will require it. Learning to use AI is an emerging skill, and I provide tutorials in Canvas about how to use them. I am happy to meet and help with these tools during office hours or after class.

Be aware of the limits of ChatGPT:

- If you provide minimum effort prompts, you will get low quality results. You will need to refine your prompts in order to get good outcomes. This will take work.
- Don't trust anything it says. If it gives you a number or fact, assume it is wrong unless you either know the answer or can check in with another source. You will be responsible for any errors or omissions provided by the tool. It works best for topics you understand.
- AI is a tool, but one that you need to acknowledge using. Please include a paragraph at the end of any assignment that uses AI explaining what you used the AI for and what prompts you used to get the results. Failure to do so is in violation of academic honesty policies.
- Be thoughtful about when this tool is useful. Don't use it if it isn't appropriate for the case or circumstance.

- We know it generates fake citations and research papers
- Test it with your assignments
- Include it in your assignment - Have students focus on refinement of the generated solutions
- ChatGPT excels at summarizing, not synthesis, so use non-traditional (audio, oral, annotation, etc.) and "Authentic Assessments"
- **Assertions:**
 - The fundamental purpose of an education is to learn stuff, but our system of education doesn't necessarily support that goal.
 - Learning stuff is accomplished through doing stuff. Learning is rooted in experiences.
 - Most students want to learn, but school is not necessarily a place where learning is important.
 - We must deal with the system as it is experienced by students today, not as we imagine it to be or as it may seemed in the past.

“If ChatGPT can do the things we ask students to do in order to demonstrate learning, it seems possible to me that those things should've been questioned a long time ago.”

-John Warner, Inside Higher Ed.

"The specific engineering of ChatGPT has made it quite compelling. But ultimately (at least until it can use outside tools) ChatGPT is **"merely" pulling out some "coherent thread of text" from the "statistics of conventional wisdom" that it's accumulated.** But it's amazing how human-like the results are. And as I've discussed, this suggests something that's at least scientifically very important: **that human language (and the patterns of thinking behind it) are somehow simpler and more "law like" in their structure than we thought.** ChatGPT has implicitly discovered it. But we can potentially explicitly expose it, with semantic grammar, computational language, etc."

My Thoughts

- It's all math, statistics, code, and data
- This is Analysis, not Synthesis or Evaluation
- Standing on the shoulders of giants is a good thing
- You might lose your balance standing on the shoulders of everyone
- AI hallucination is a real thing
- ChatGPT (AI/ML) is a mirror for human kind, and we need to take a long, hard, look in the mirror





Xavier Hubert-Brierre via Tenor

HAPPINESS COMES
FROM HAVING MORE





60

Any questions?



Certainly! I'm here to answer any questions you have. Is there a specific topic you're interested in, or do you have a general question?



References

- <https://www.motherjones.com/politics/2023/02/bing-ai-chatbot-falsehoods-fact-checking-microsoft/>
- <https://medium.com/walmartglobaltech/the-journey-of-open-ai-gpt-models-32d95b7b7fb2>
- <https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/>
- Radford, Alec and Karthik Narasimhan. "Improving Language Understanding by Generative Pre-Training." (2018).
- Radford, Alec et al. "Language Models are Unsupervised Multitask Learners." (2019).
- Tom B. Brown, et. al. 2020. Language models are few-shot learners. In Proceedings of the 34th International Conference on Neural Information Processing Systems (NIPS'20). Curran Associates Inc., Red Hook, NY, USA, Article 159, 1877-1901.