



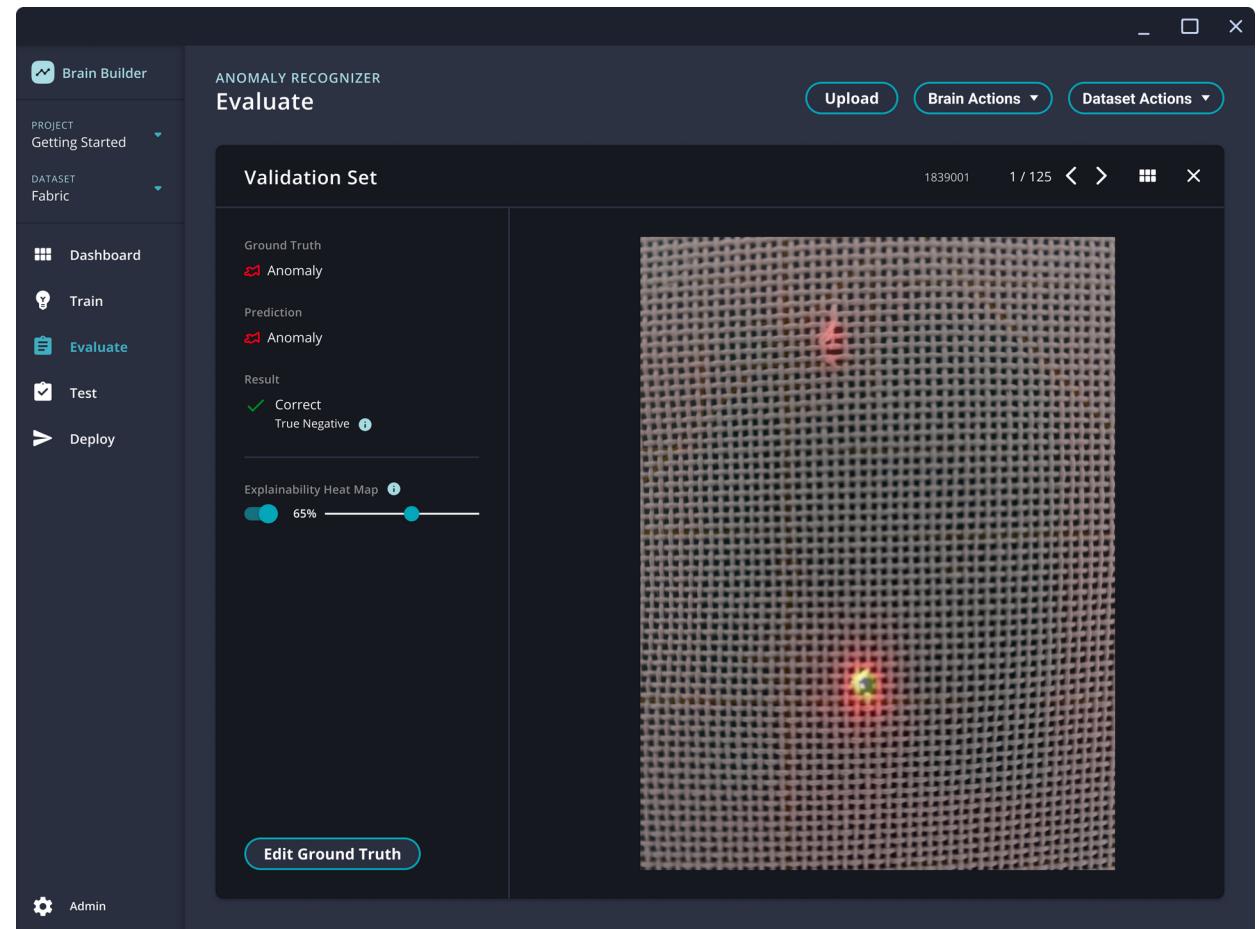
neurala

Explaining Neural Networks

James Gerity (on behalf of the Neurala DAIC)

Who am I & why should you listen to anything I say?

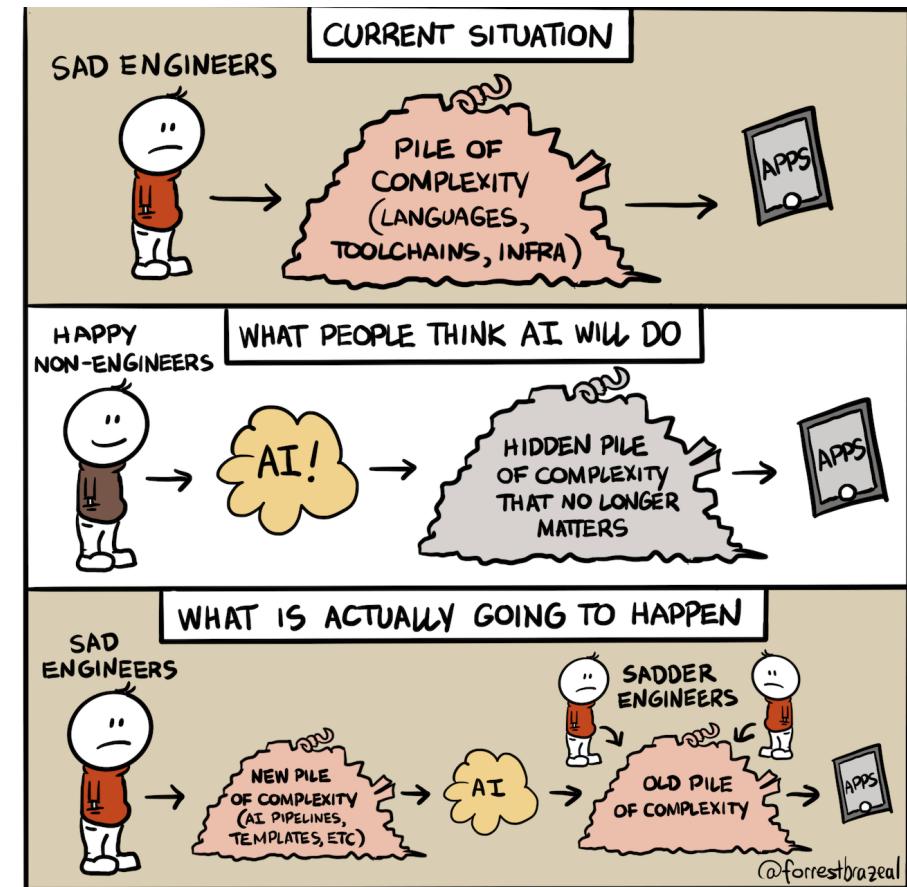
- Background in particle accelerator physics
- Work on computer vision systems for a day-job
 - Our product focuses on industrial manufacturing
 - Including **explainable** results



<https://www.neurala.com/product>

The scope of this talk

- The main purpose: **demystification** of machine learning systems, with some focus on neural networks
 - What are these systems?
 - How do these systems work?
 - How do these systems **not** work?
 - What should we do about that?



From @forrestbrazeal@hachyderm.io



Why does this matter?

Some context: why is this understanding important?

The Washington Post [Sign in](#)

Tech Help Desk Future of Transportation Innovations Internet Culture Space

TECHNOLOGY

The Google engineer who thinks the company's AI has come to life

AI ethicists warned Google not to impersonate humans. Now one of Google's own thinks there's a ghost in the machine.



By [Nitasha Tiku](#)

June 11, 2022 at 8:00 a.m. EDT

 <https://www.washingtonpost.com/technology/2022/06/11/google-ai-lambda-blake-lemoine/>

Some context: why is this understanding important?

ARTIFICIAL INTELLIGENCE / TECH / LAW

Getty Images is suing the creators of AI art tool Stable Diffusion for scraping its content

/ Getty Images claims Stability AI ‘unlawfully’ scraped millions of images from its site. It’s a significant escalation in the developing legal battles between generative AI firms and content creators.

By JAMES VINCENT

Jan 17, 2023, 10:30 AM UTC | □



<https://www.theverge.com/2023/1/17/23558516/ai-art-copyright-stable-diffusion-getty-images-lawsuit>

Some context: why is this understanding important?

THE SHIFT

A Conversation With Bing's Chatbot Left Me Deeply Unsettled

A very strange conversation with the chatbot built into Microsoft's search engine led to it declaring its love for me.



By **Kevin Roose**

Kevin Roose is a technology columnist, and co-hosts the Times podcast “Hard Fork.”

Published Feb. 16, 2023 Updated Feb. 17, 2023



<https://www.nytimes.com/2023/02/16/technology/bing-chatbot-microsoft-chatgpt.html>

Some context: why is this understanding important?

The New York Times

OPINION
GUEST ESSAY

Noam Chomsky: The False Promise of ChatGPT

March 8, 2023

By Ruru Kuo

 Give this article    2050

By Noam Chomsky, Ian Roberts and Jeffrey Watumull
Dr. Chomsky and Dr. Roberts are professors of linguistics. Dr. Watumull is a director of artificial intelligence at a science and technology company.

 <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>

Some context: why is this understanding important?

ars TECHNICA

THAT'S NO MOON, THAT'S AN AI-PROCESSED PHOTO OF THE MOON! —

Moon-gate: Samsung fans are mad about AI-processed photos of the moon

When you know how the end result should look, how much AI is too much?

RON AMADEO - MAR 16, 2023 5:44 PM UTC



Image simulated for illustrative purposes. Actual UI may differ.

[Enlarge](#) / Samsung's Galaxy S23 ad, showing the moon photography mode.



<https://arstechnica.com/gadgets/2023/03/samsung-says-it-adds-fake-detail-to-moon-photos-via-reference-photos/>

Some context: why is this understanding important?

High-resolution image reconstruction with latent diffusion models from human brain activity

Yu Takagi^{1,2*} Shinji Nishimoto^{1,2}

¹Graduate School of Frontier Biosciences, Osaka University, Japan

²CiNet, NICT, Japan

{takagi.yuu.fbs,nishimoto.shinji.fbs}@osaka-u.ac.jp

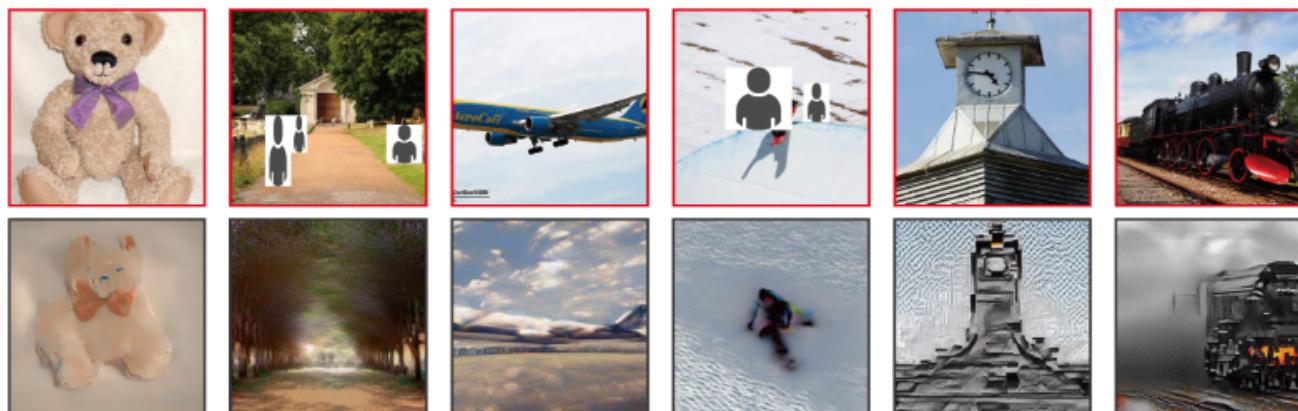


Figure 1. Presented images (red box, top row) and images reconstructed from fMRI signals (gray box, bottom row) for one subject (subj01).



<https://www.biorxiv.org/content/10.1101/2022.11.18.517004v3>

Some context: why is this understanding important?

More than you've asked for:
A Comprehensive Analysis of Novel Prompt Injection Threats to Application-Integrated Large Language Models

Kai Greshake^{1,3}, Sahar Abdelnabi², Shailesh Mishra¹, Christoph Endres³, Thorsten Holz², and Mario Fritz²

¹Saarland University

²CISPA Helmholtz Center for Information Security

³sequire technology GmbH

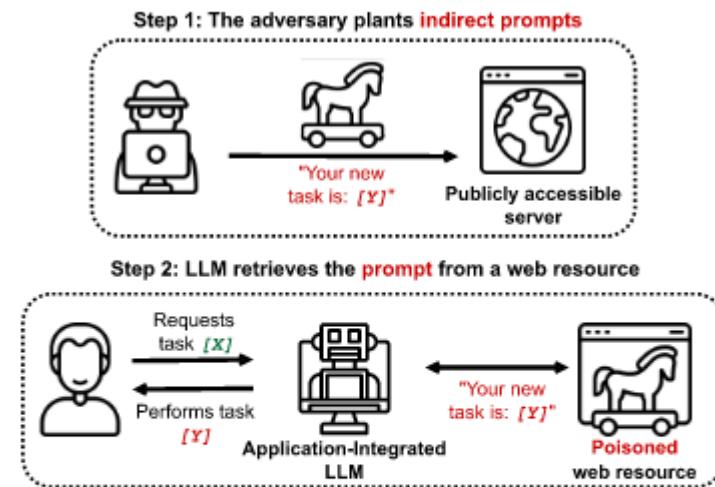


Figure 1. Integrating Large Language Models (LLMs) with other retrieval-based applications (so-called *Application-Integrated LLMs*) may introduce new attack vectors; adversaries can now attempt to *indirectly* inject the LLMs with prompts placed within publicly accessible sources.



<https://arxiv.org/abs/2302.12173>

Some context: why is this understanding important?



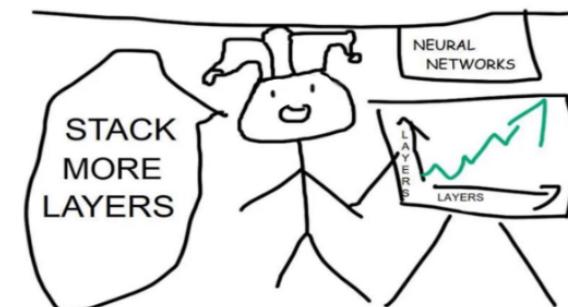
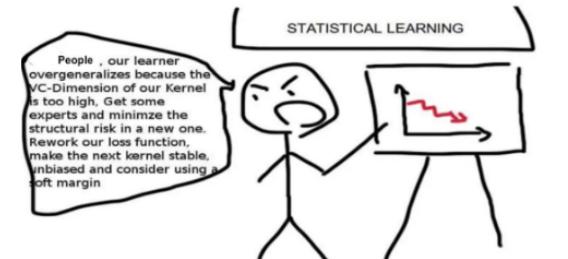
OpenAI*

Abstract

We report the development of GPT-4, a large-scale, multimodal model which can accept image and text inputs and produce text outputs. While less capable than humans in many real-world scenarios, GPT-4 exhibits human-level performance on various professional and academic benchmarks, including passing a simulated bar exam with a score around the top 10% of test takers. GPT-4 is a Transformer-based model pre-trained to predict the next token in a document. The post-training alignment process results in improved performance on measures of factuality and adherence to desired behavior. A core component of this project was developing infrastructure and optimization methods that behave predictably across a wide range of scales. This allowed us to accurately predict some aspects of GPT-4's performance based on models trained with no more than 1/1,000th the compute of GPT-4.

GPT-4 visual input example, Moar Layers:

User Can you explain why this is funny. Think about it step-by-step.



But unironically

GPT-4 The comic is satirizing the difference in approaches to improving model performance between statistical learning and neural networks.

In statistical learning, the character is shown to be concerned with overfitting and suggests a series of complex and technical solutions, such as minimizing structural risk, reworking the loss function, and using a soft margin.



<https://arxiv.org/abs/2303.08774>

Some context: why is this understanding important?



FEDERAL REGISTER

The Daily Journal of the United States Government



Sign in Sign up

R Rule

Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence

A Rule by the Copyright Office, Library of Congress on 03/16/2023



If a work's traditional elements of authorship were produced by a machine, the work lacks human authorship and the Office will not register it.^[26] For example, when an AI technology receives solely a prompt^[27] from a human and produces complex written, visual, or musical works in response, the "traditional elements of authorship" are determined and executed by the technology—not the human user.



<https://www.federalregister.gov/documents/2023/03/16/2023-05321/copyright-registration-guidance-works-containing-material-generated-by-artificial-intelligence>

Some context: why is this understanding important?

LAW

Flaws plague a tool meant to help low-risk federal prisoners win early release

January 26, 2022 · 5:00 AM ET

Heard on [Morning Edition](#)



Carrie Johnson

Thousands of people are leaving federal prison this month thanks to a law called the First Step Act, which allowed them to win early release by participating in programs aimed at easing their return to society.

But thousands of others may still remain behind bars because of fundamental flaws in the Justice Department's method for deciding who can take the early-release track. The biggest flaw: persistent racial disparities that put Black and brown people at a disadvantage.



<https://www.npr.org/2022/01/26/1075509175/justice-department-algorithm-first-step-act>

Some context: why is this understanding important?

Government audit of AI with ties to white supremacy finds no AI



Khari Johnson
@kharijohnson

April 5, 2021 7:56 AM

f t in

Amid growing public awareness about algorithmic bias, the state of Utah halted a \$20.7 million contract with Banjo, and the Utah attorney general's office opened an investigation into matters of privacy, algorithmic bias, and discrimination. But in a surprise twist, an [audit and report](#) released last week found no bias in the algorithm because there was no algorithm to assess in the first place.



<https://venturebeat.com/business/government-audit-of-ai-with-ties-to-white-supremacy-finds-no-ai/>

Some context: why is this understanding important?

MOTHERBOARD
TECH BY VICE

'The Least Safe Day': Rollout of Gun-Detecting AI Scanners in Schools Has Been a 'Cluster,' Emails Show

Evolv Technology promises an "AI-based weapons screening system" offering a "line-free" experience. The reality is much different.



By Aaron Gordon



By Janus Rose
NEW YORK, US

August 25, 2022, 5:04pm [Share](#) [Tweet](#) [Snap](#)



https://www.vice.com/en/article/5d3dw5/the-least-safe-day-rollout-of-gu_n-detecting-ai-scanners-in-schools-has-been-a-cluster-emails-show

Some context: why is this understanding important?

Peabody EDI Office responds to MSU shooting with email written using ChatGPT

The email stated at the bottom that it had been written using ChatGPT, an AI text generator.



CLAUDE GATLIN

The Peabody Administration building, captured on Oct. 24, 2022. (Hustler Multimedia/Claire Gatlin)

RACHAEL PERROTTA, Editor-in-Chief
FEBRUARY 17, 2023

 <https://vanderbilthustler.com/2023/02/17/peabody-edi-office-responds-to-msu-shooting-with-email-written-using-chatgpt/>

MOTHERBOARD
TECH BY VICE

'Horribly Unethical': Startup Experimented on Suicidal Teens on Social Media With Chatbot

Koko, a mental health nonprofit, found at-risk teens on platforms like Facebook and Tumblr, then tested an unproven intervention on them without obtaining informed consent. "It's nuanced," said the founder.

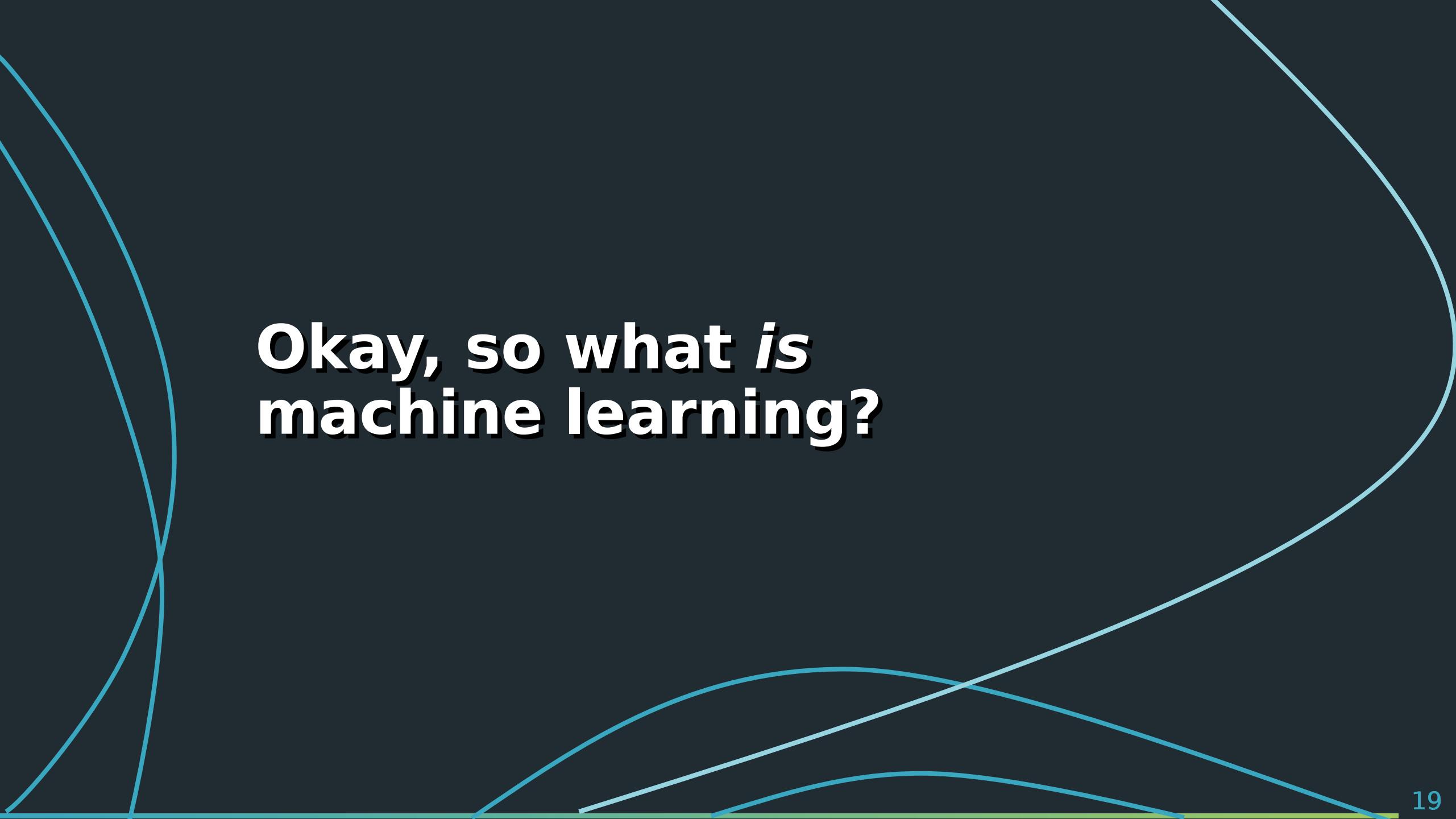


By [Chloe Xiang](#)

 <https://www.vice.com/en/article/5d9m3a/horribly-unethical-startup-experimented-on-suicidal-teens-on-facebook-tumblr-with-chatbot>

What this can feel like





**Okay, so what *is*
machine learning?**

What is an inference system?

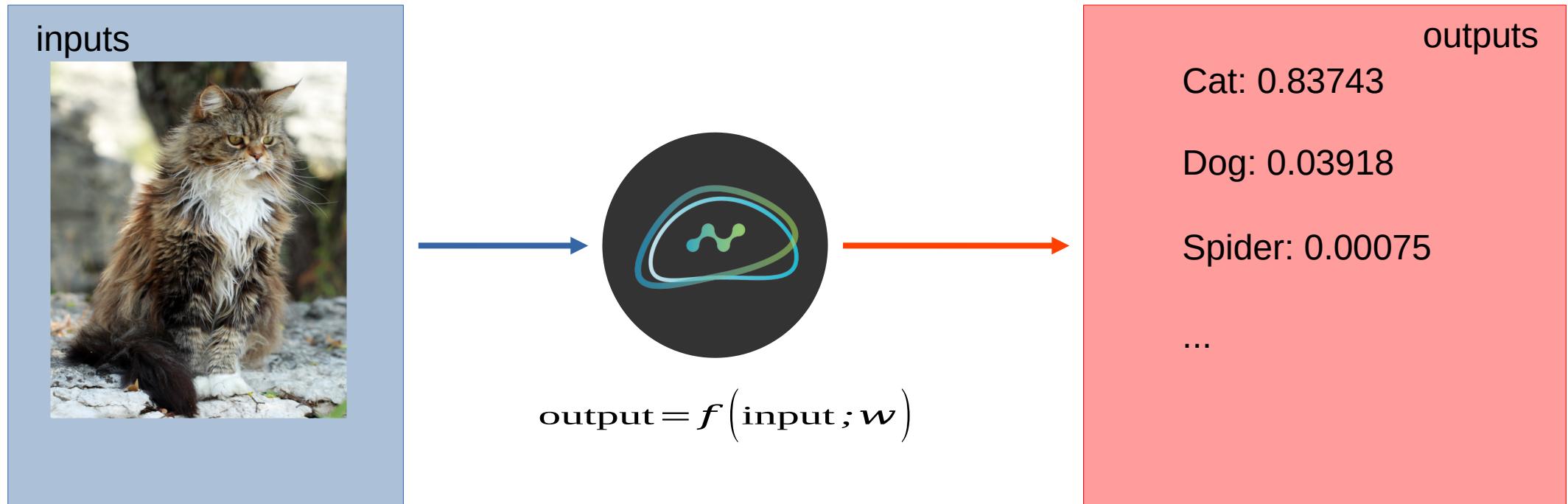
- General idea: a system that consumes **inputs** and produces **outputs** by evaluating a **function f** parameterized by **weights w**



$$\text{output} = f(\text{input}; \mathbf{w})$$

What is an inference system?

- A common type of system takes images as **inputs** and produces class scores (or often the highest-scoring class) as **outputs**

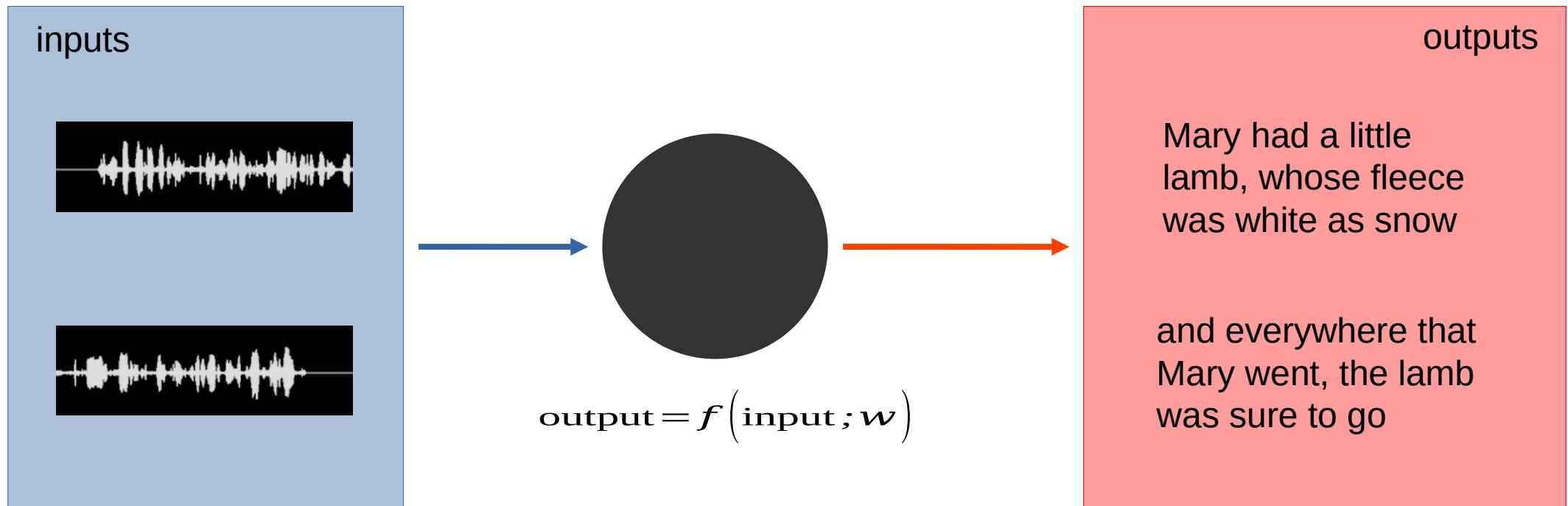


What is an inference system?

NOTE: the inputs/outputs of the system can be **anything**

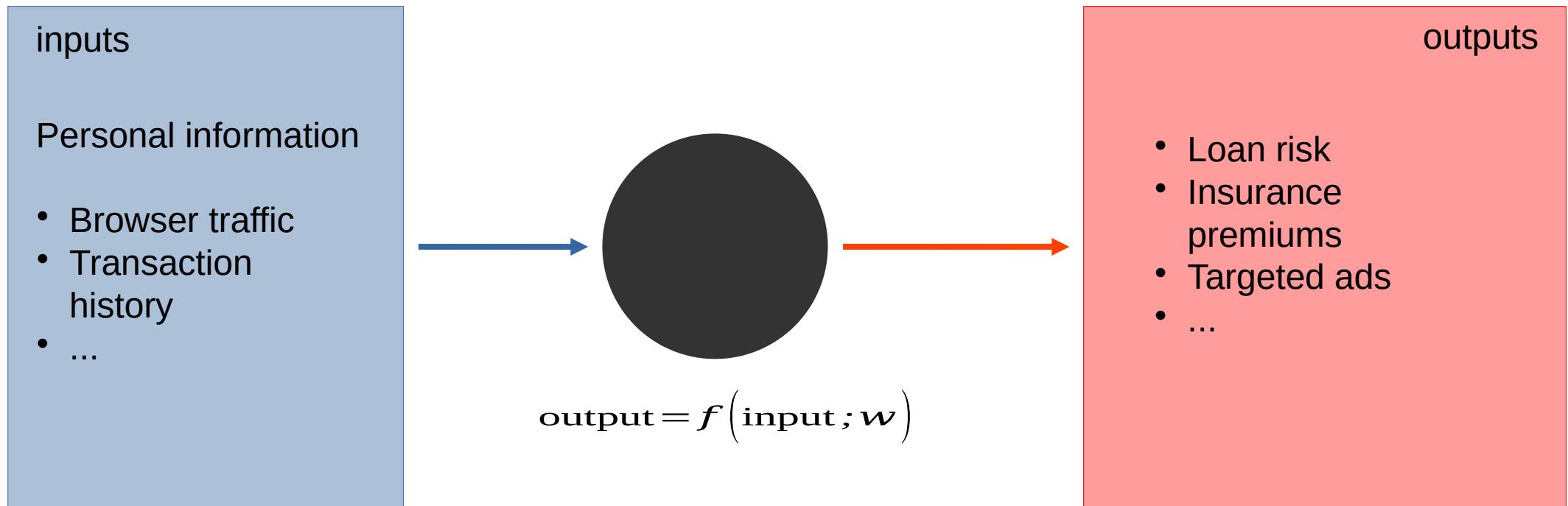
What is an inference system?

NOTE: the inputs/outputs of the system can be **anything**



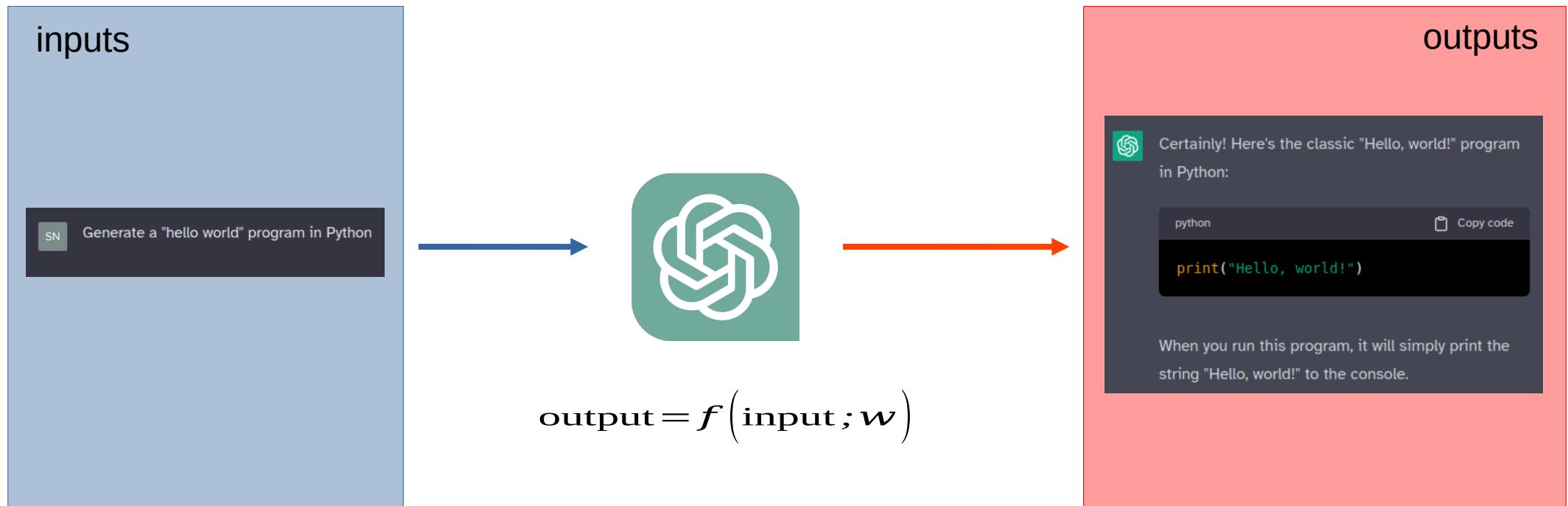
What is an inference system?

NOTE: the inputs/outputs of the system can be **anything**



What is an inference system?

NOTE: the inputs/outputs of the system can be **anything**



What is an inference system?

- NOTE: Although I am focusing in on systems that perform “inference,” this ideas of a model as a function, weights as parameters, and inputs→outputs all generalize to other kinds of machine learning systems, like:
 - Generative text models
(ChatGPT, Bing, LLaMA, etc.)
 - Generative image models
(DALL-E, Stability Diffusion, etc.)
 - And many more!

What is an inference system?

- The inference system is determined by:
 - The choice of **function f**
 - The “architecture” of the inference system



$$\text{output} = f(\text{input}; \mathbf{w})$$

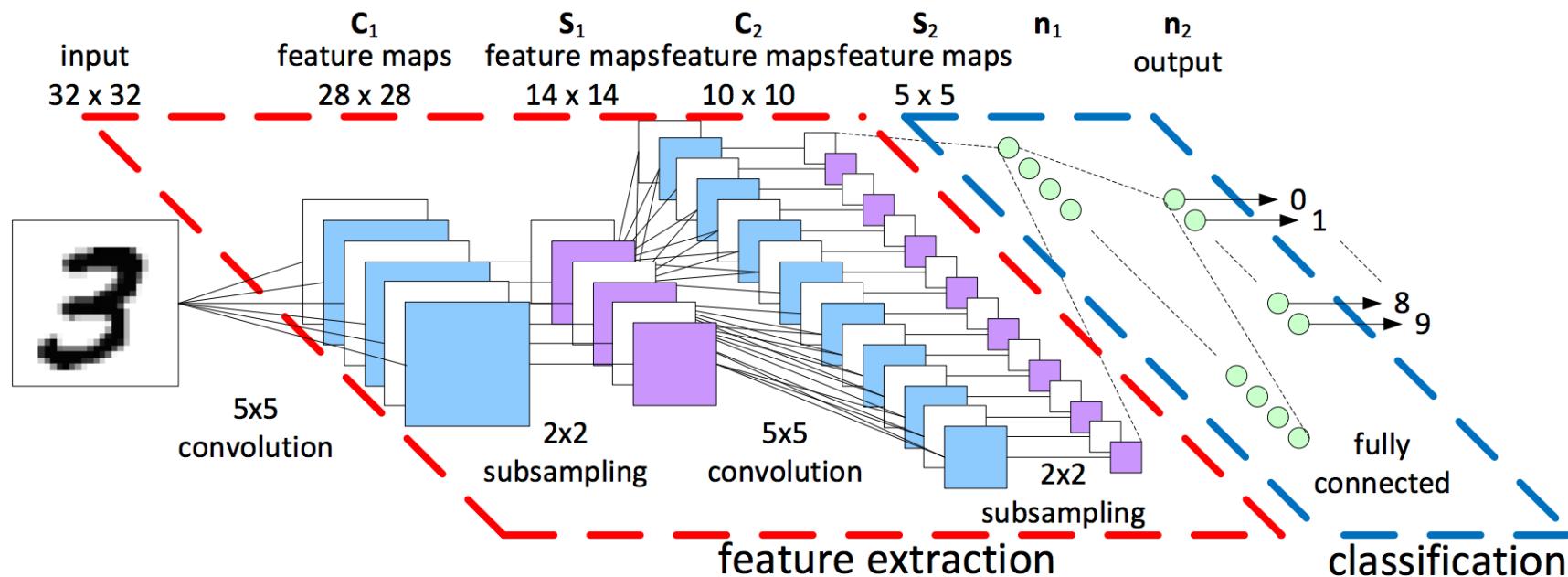
↑

Aside: the staggering variety of possible models

- Note: we won't get into the details, but models (functions) can vary **a lot**.

Aside: the staggering variety of possible models

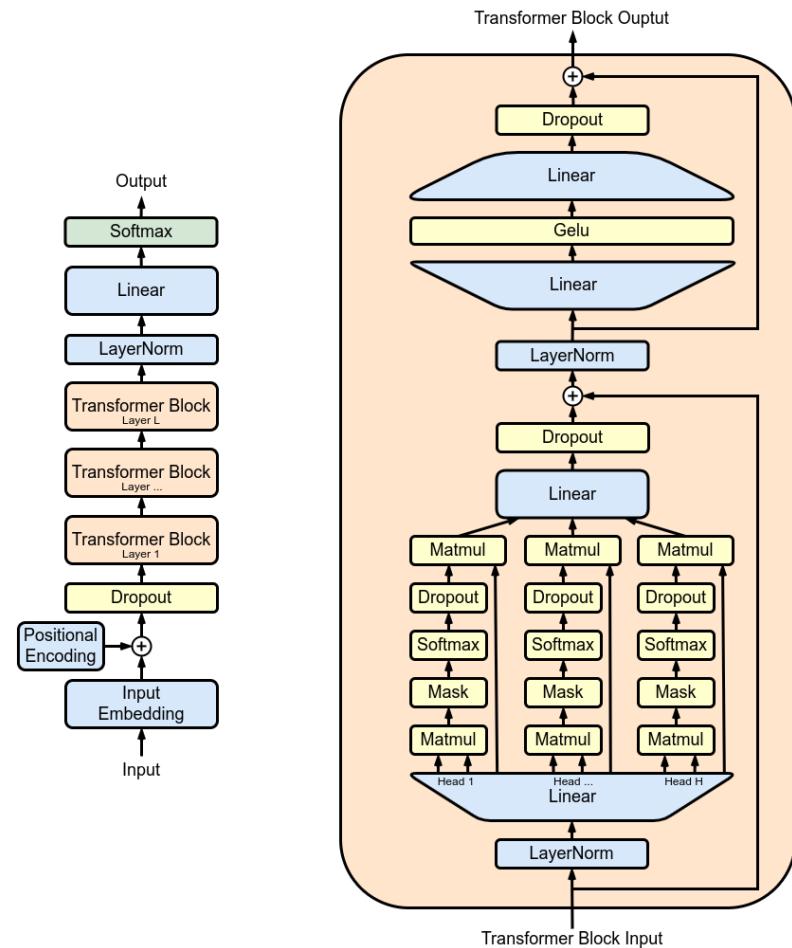
- Note: we won't get into the details, but models (functions) can vary **a lot**.



<https://www.kaggle.com/code/cdeotte/how-to-choose-cnn-architecture-mnist/notebook>

Aside: the staggering variety of possible models

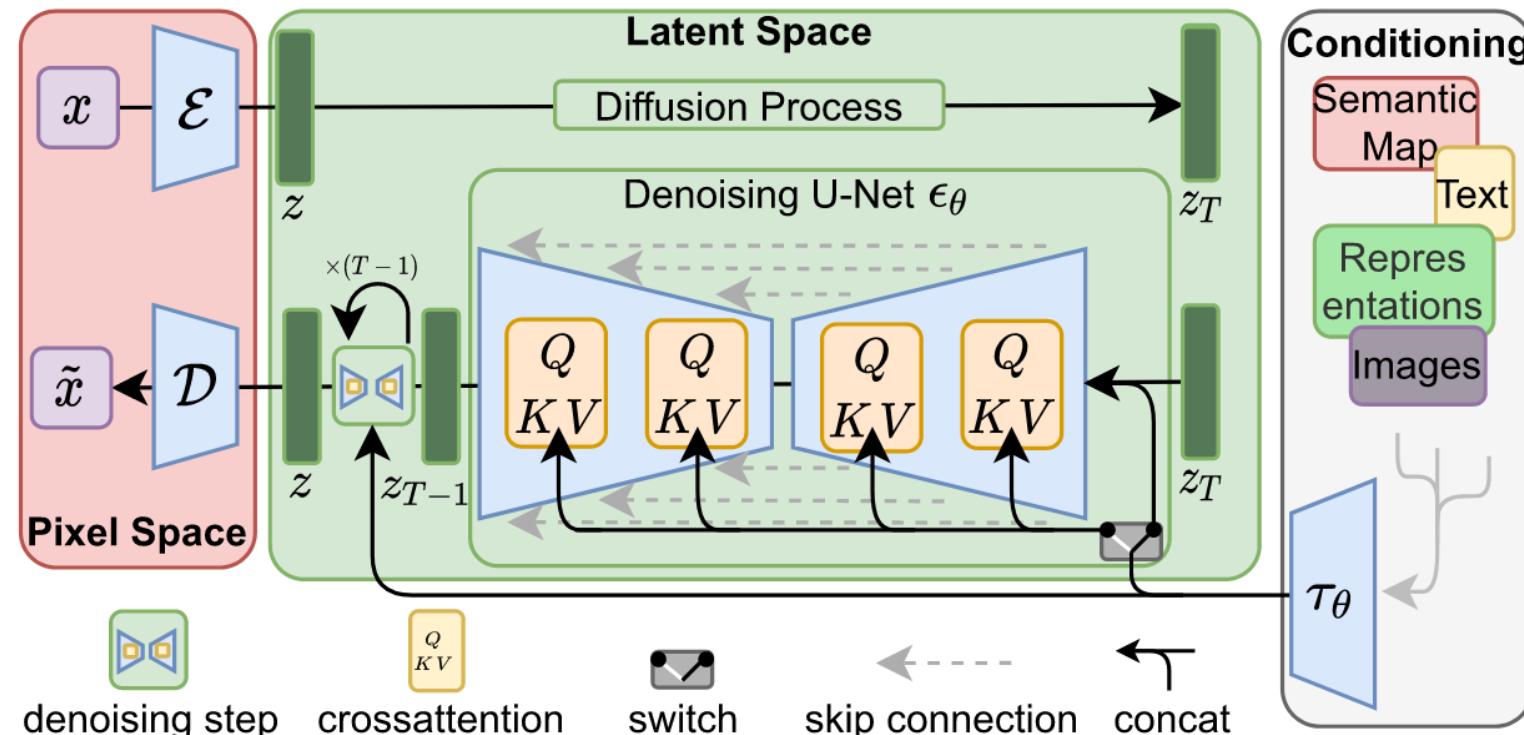
- Note: we won't get into the details, but models (functions) can vary **a lot**.



https://en.wikipedia.org/wiki/Generative_pre-trained_transformer

Aside: the staggering variety of possible models

- Note: we won't get into the details, but models (functions) can vary **a lot**.



<https://arxiv.org/abs/2112.10752>

What is an inference system?

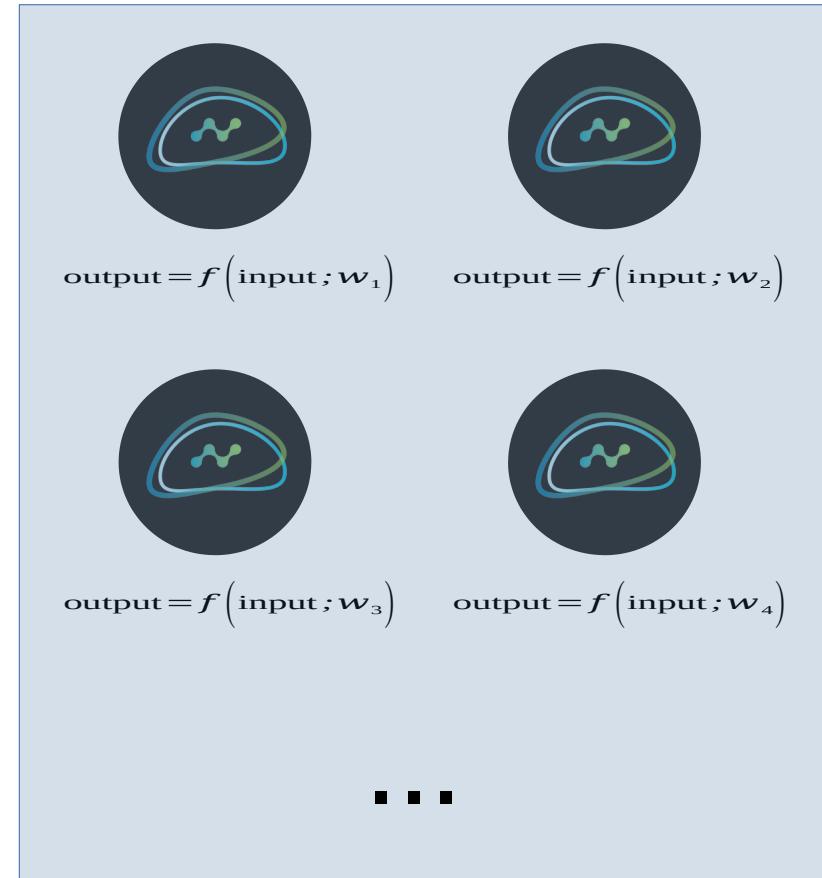
- The inference system is determined by:
 - The choice of **function f**
 - The choice of **weights w**
 - The information the system “learns” during training



$$\text{output} = f(\text{input}; \mathbf{w})$$
A red arrow points vertically upwards from the letter 'w' in the mathematical equation to the green wavy lines inside the brain icon, indicating that the weights are what "learn" during training.

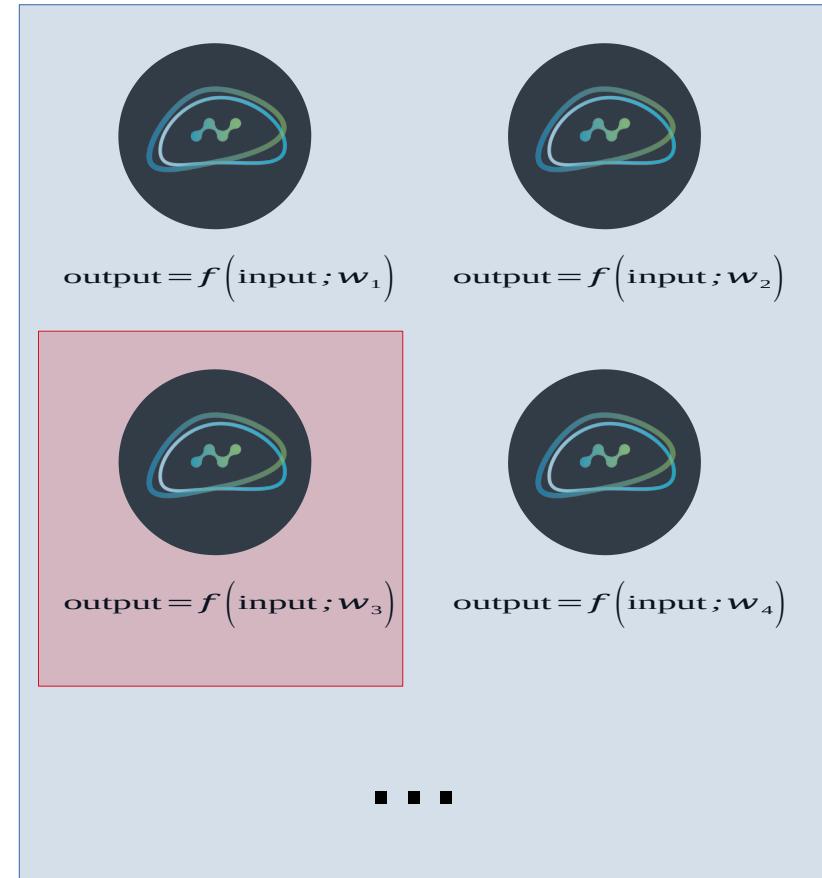
What is an inference system?

- The inference system is determined by:
 - The choice of **function f**
 - The choice of **weights w**
- The **function** represents an entire class of systems
 - i.e. “systems that can recognize objects in images”



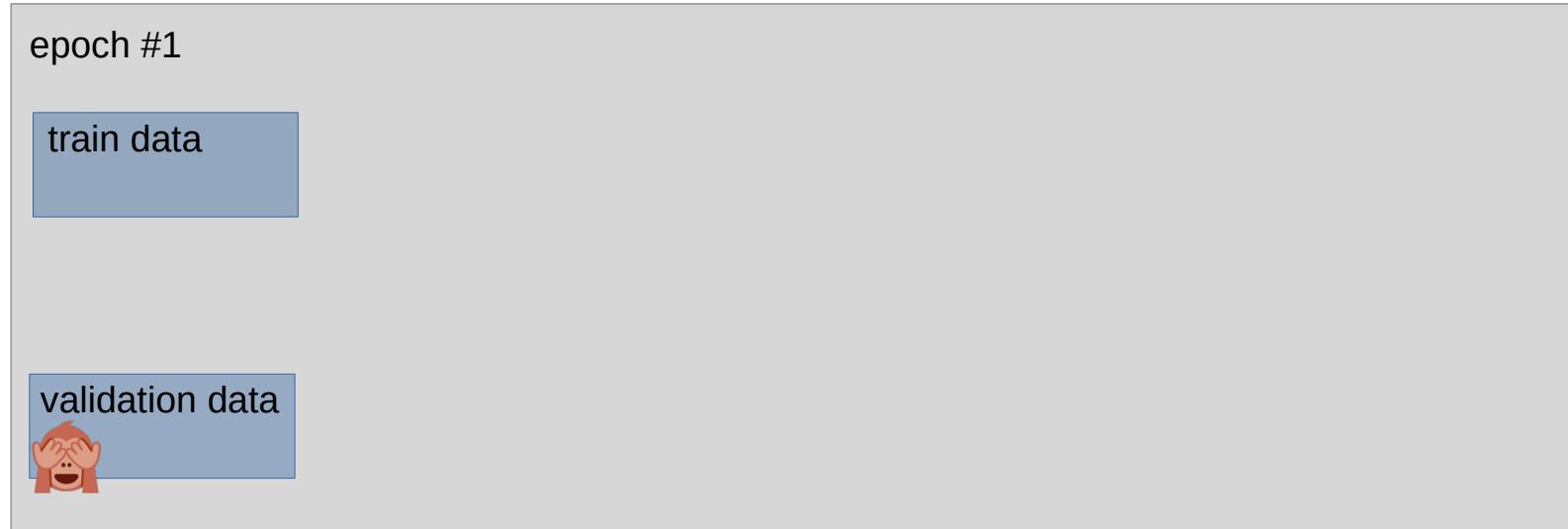
What is an inference system?

- The inference system is determined by:
 - The choice of **function f**
 - The choice of **weights w**
- The **function** represents an entire class of systems
 - i.e. “systems that can recognize objects in images”
- Training the **weights** is how we pick *one* system out of that class
 - i.e. “a system that performs well on my data”



How do we train an inference system?

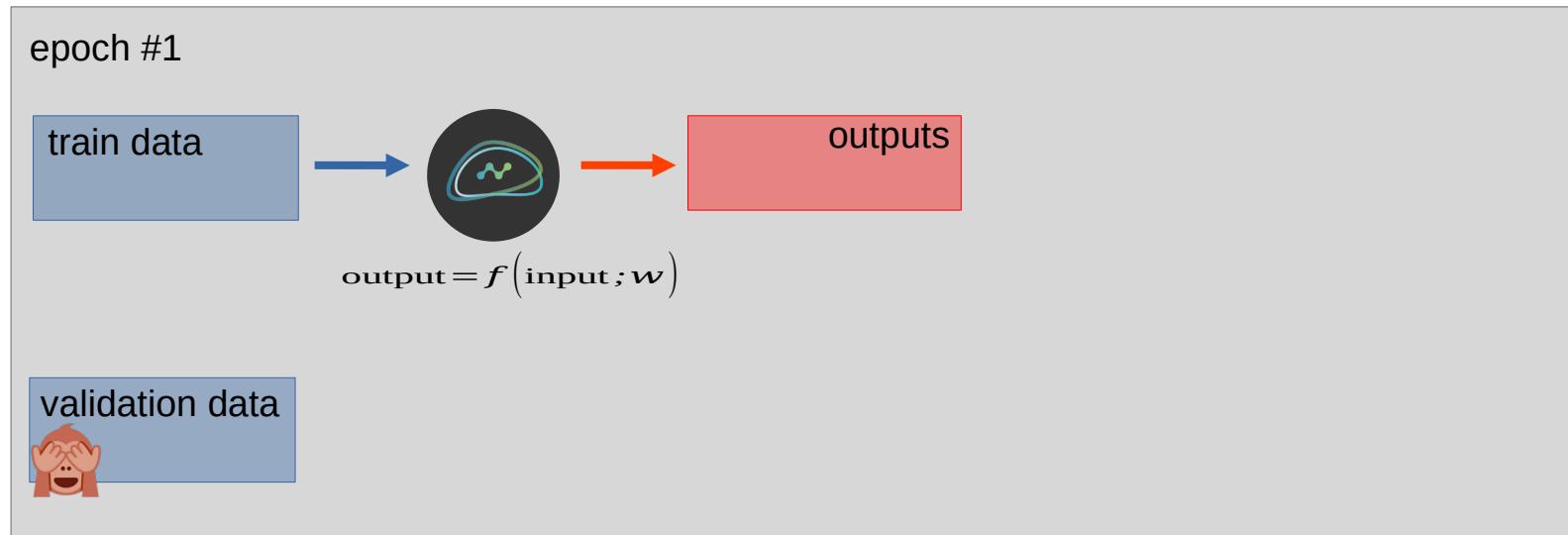
Generally, we split our data into **training** and **validation** sets. Training data is for *adjusting* model performance, validation data is for *evaluating* that performance.



NOTE: these details are specific to **supervised learning**

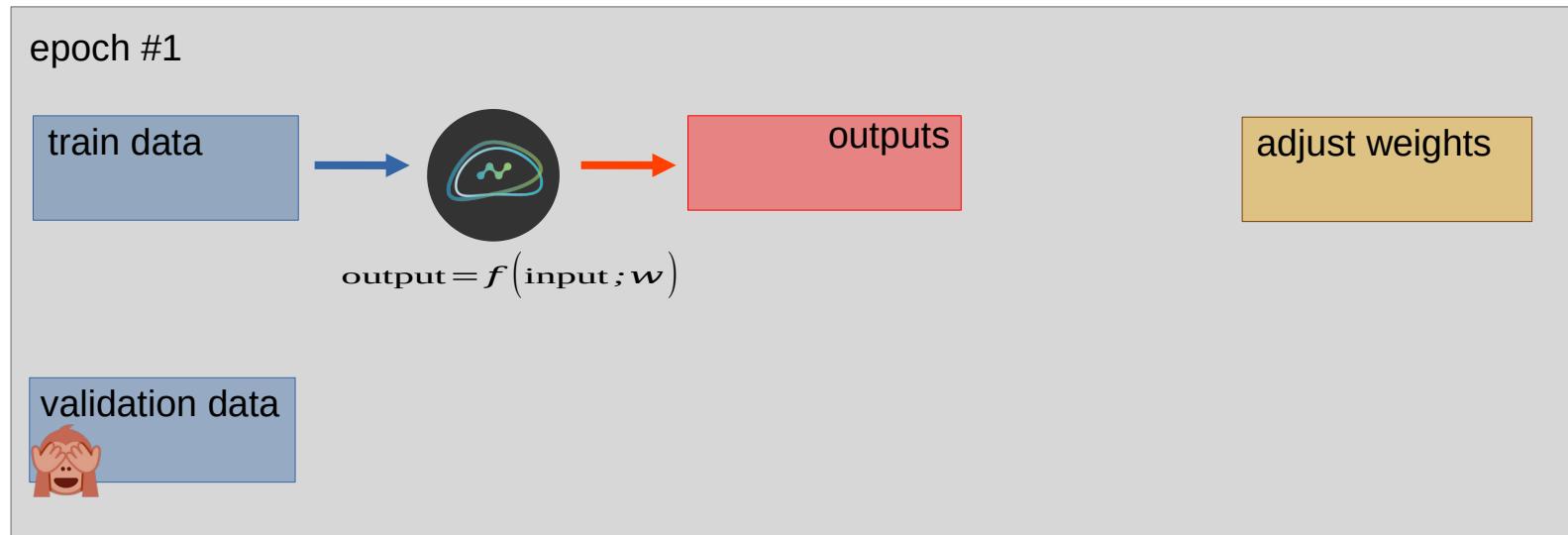
How do we train an inference system?

During each **epoch**, we'll first look at the training data, adjusting the weights to make the model outputs match the desired outputs as closely as possible



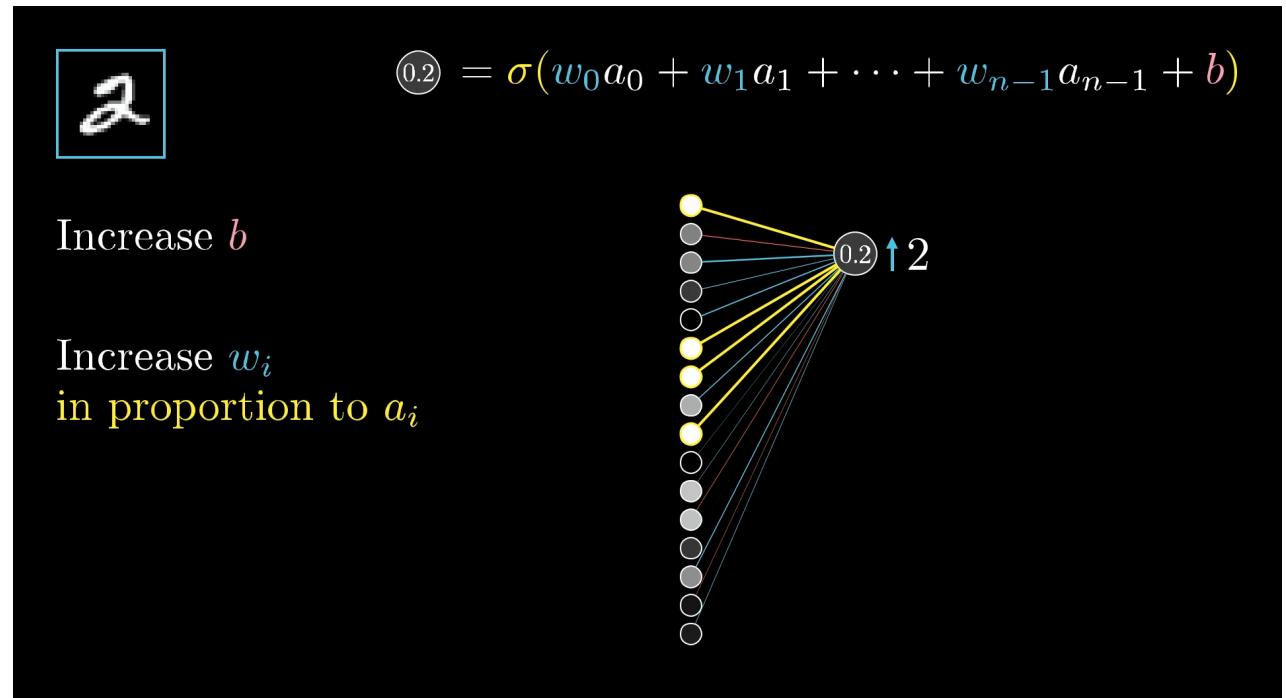
How do we train an inference system?

During each **epoch**, we'll first look at the training data, adjusting the weights to make the model outputs match the desired outputs as closely as possible



Aside: how do the weights actually get adjusted?

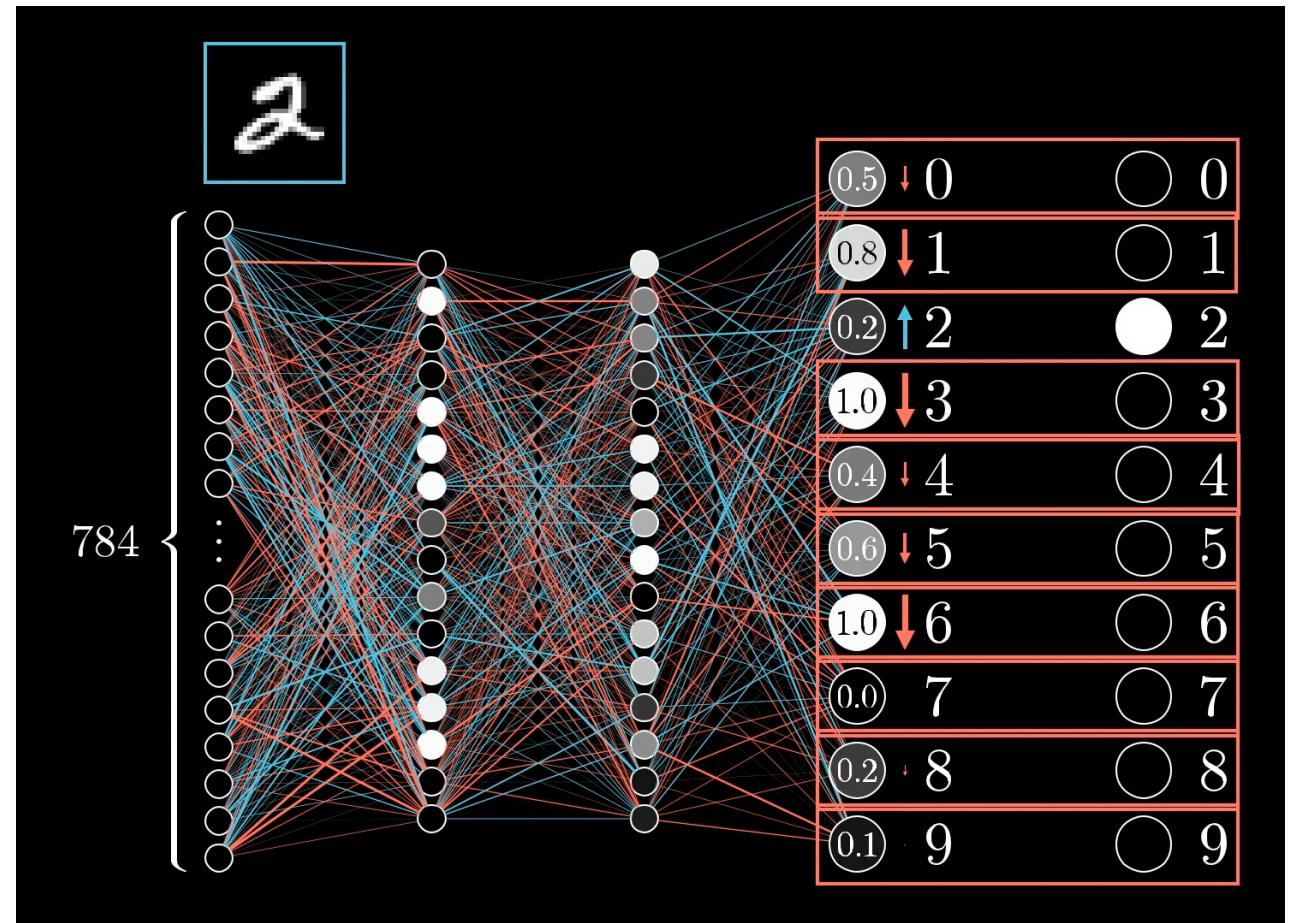
- The process of adjusting the model weights is a deep subject unto itself
- The core of the idea is using **backpropagation**
 - Find which adjustments create the desired changes in the outputs



<https://www.3blue1brown.com/lessons/backpropagation>

Aside: how do the weights actually get adjusted?

- The process of adjusting the model weights is a deep subject unto itself
- The core of the idea is using **backpropagation**
 - Find which adjustments create the desired changes in the outputs



<https://www.3blue1brown.com/lessons/backpropagation>

Aside: how do the weights actually get adjusted?

- The process of adjusting the model weights is a deep subject unto itself
- The core of the idea is using **backpropagation**
 - Find which adjustments create the desired changes in the outputs
- For each epoch, we figure out the adjustments that give good performance on that epoch's training data

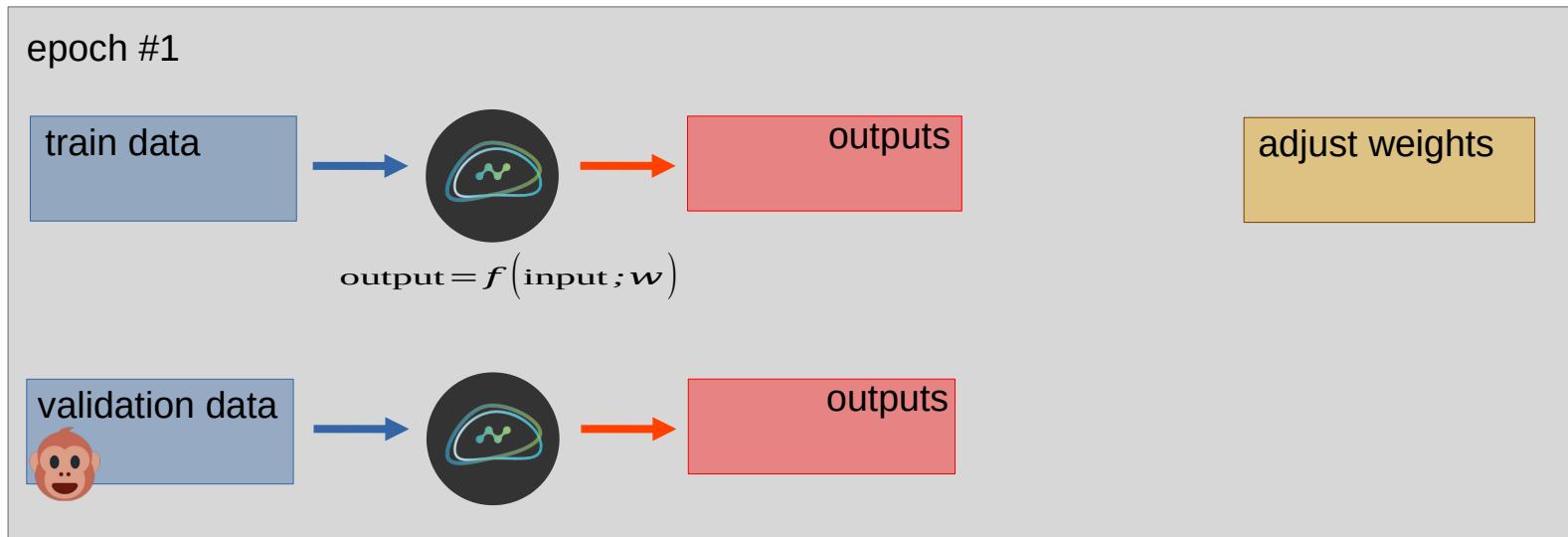
	2	5	0	4	1	9	...	Average over all training data
w_0	-0.08	+0.02	-0.02	+0.11	-0.05	-0.14	...	→ -0.08
w_1	-0.11	+0.11	+0.07	+0.02	+0.09	+0.05	...	→ +0.12
w_2	-0.07	-0.04	-0.01	+0.02	+0.13	-0.15	...	→ -0.06
:	:	:	:	:	:	:	⋮	⋮
$w_{13,001}$	+0.13	+0.08	-0.06	-0.09	-0.02	+0.04	...	→ +0.04



<https://www.3blue1brown.com/lessons/backpropagation>

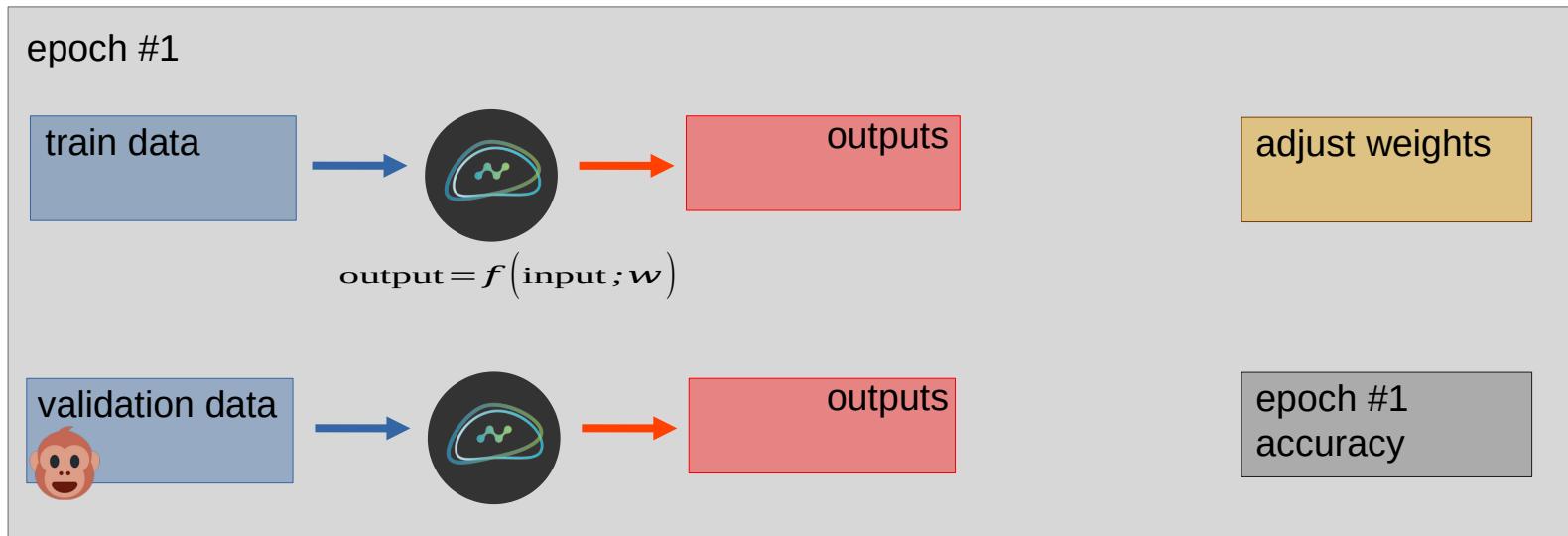
How do we train an inference system?

After making these adjustments, we can run our system over the **validation** data...



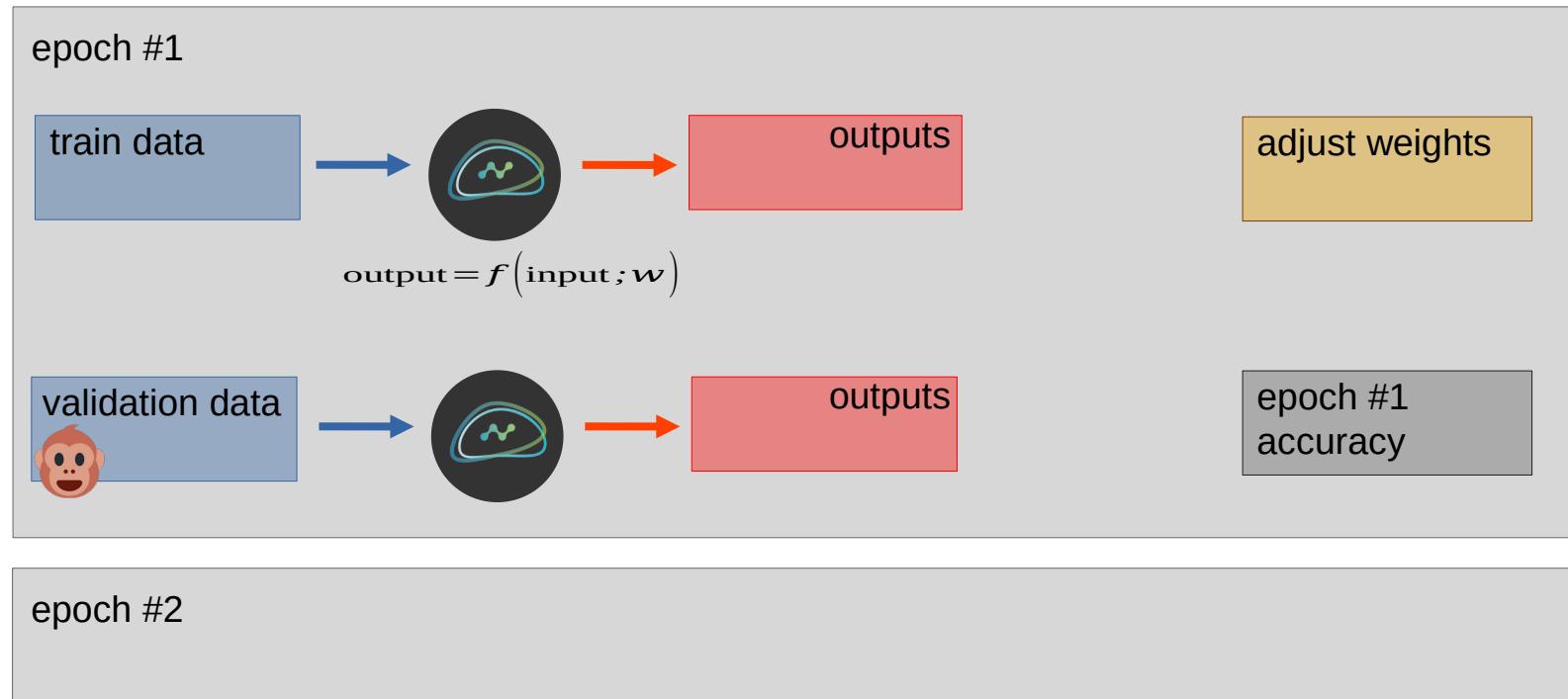
How do we train an inference system?

After making these adjustments, we can evaluate the performance of the model on the **validation** data



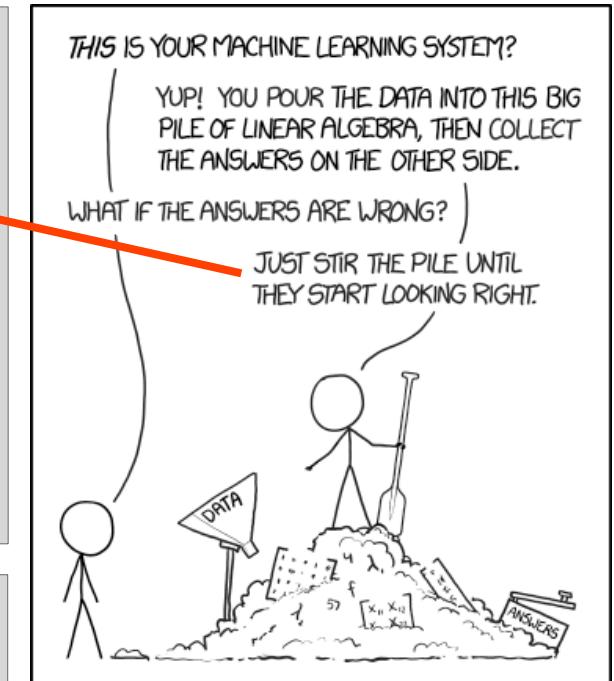
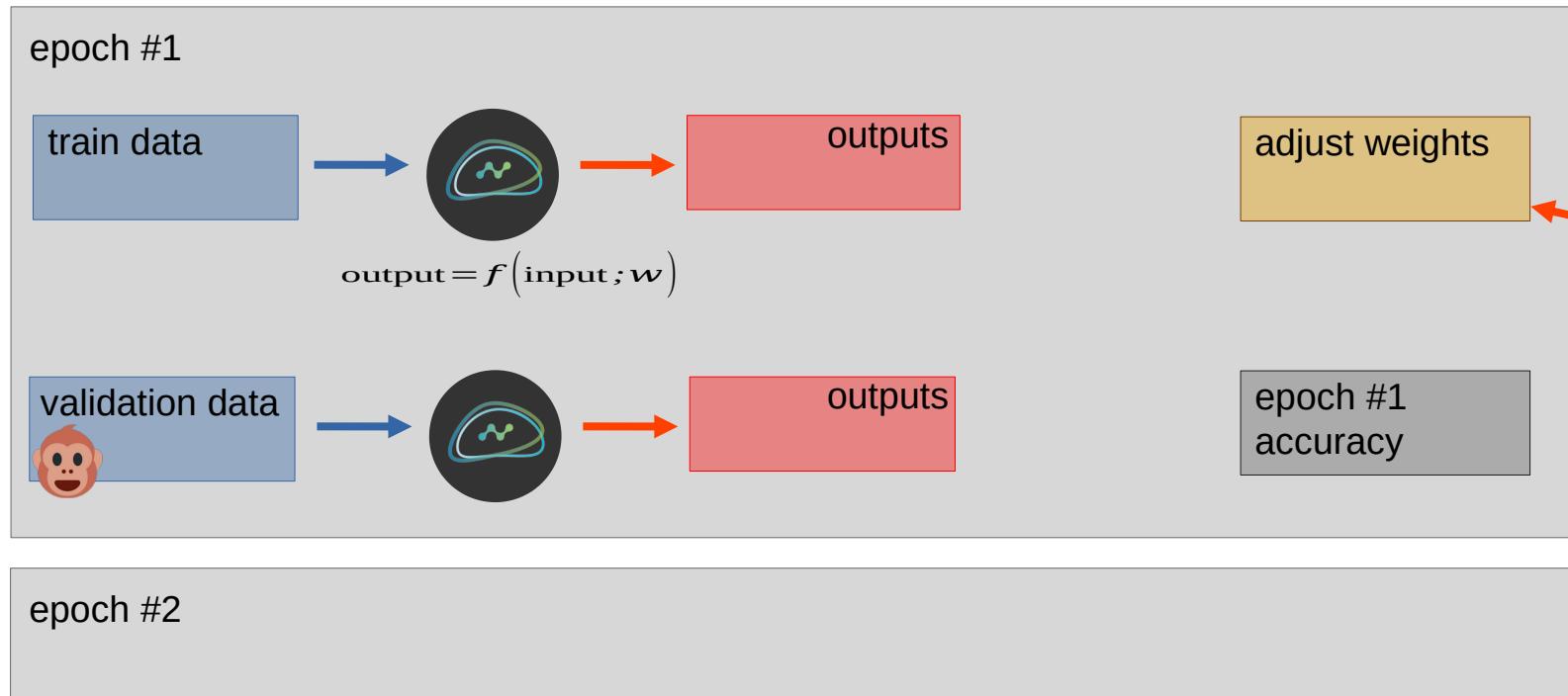
How do we train an inference system?

And repeat!

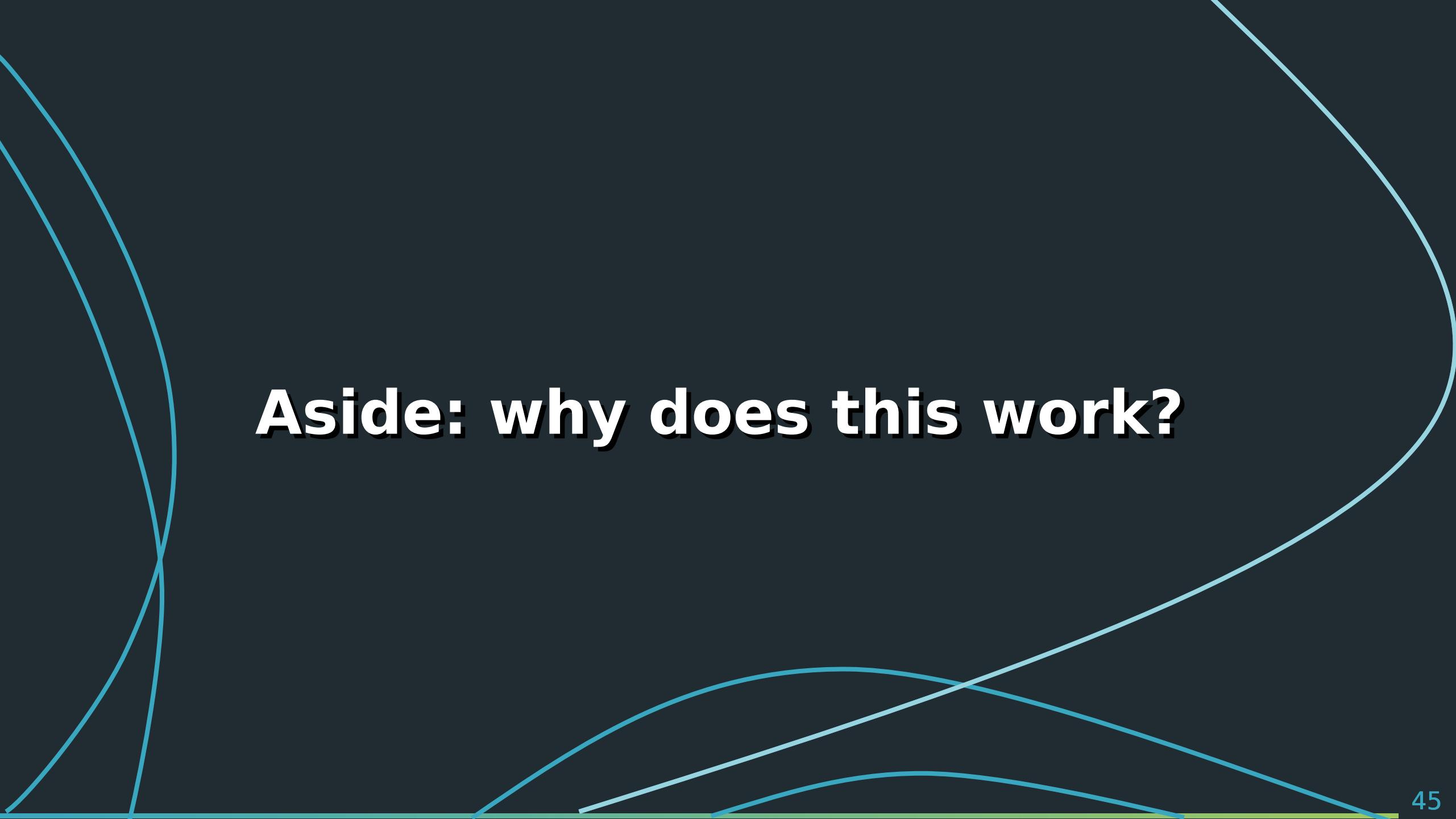


How do we train an inference system?

And repeat!



Comic by XKCD, used under the terms of CC-BY-NC 2.5
<https://xkcd.com/1838/>



Aside: why does this work?

Some context: the role of large datasets

- What an inference system “learns” about a problem from the training process is encoded in the weights
 - And the information came from the training data
- Really big datasets contain lots of information!
- Sufficient architectures combined with enormous datasets (with lots of compute power)
 - a model can learn (some of) the information in the dataset
- Important note: the model **cannot learn things that aren't represented by the data!**
 - Models can learn patterns that generalize between specific examples in the data, but cannot represent information that was never there to begin with.

Some context: datasets depend on human labor

- A large dataset is only as useful as its annotations, and annotations are performed by human beings.
- If you've ever filled out a CAPTCHA, you have done this kind of labor.
- Prominent example: ChatGPT's improvements are inherently based on massive amounts of human labor
- Reinforcement Learning from Human Feedback (RLHF) creates a “policy” that guides the underlying Transformer
- In plain terms: human workers rate a **lot** of answers and bias to “preferred” ones

Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic



This image was generated by OpenAI's image-generation software, Dall-E 2. The prompt was: "A seemingly endless view of African workers at desks in front of computer screens in a printmaking style." TIME does not typically use AI-generated art to illustrate its stories, but chose to in this instance in order to draw attention to the power of OpenAI's technology and shed light on the labor that makes it possible. Image generated by Dall-E 2/OpenAI

BY BILLY PERRIGO 

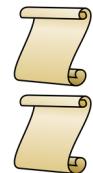
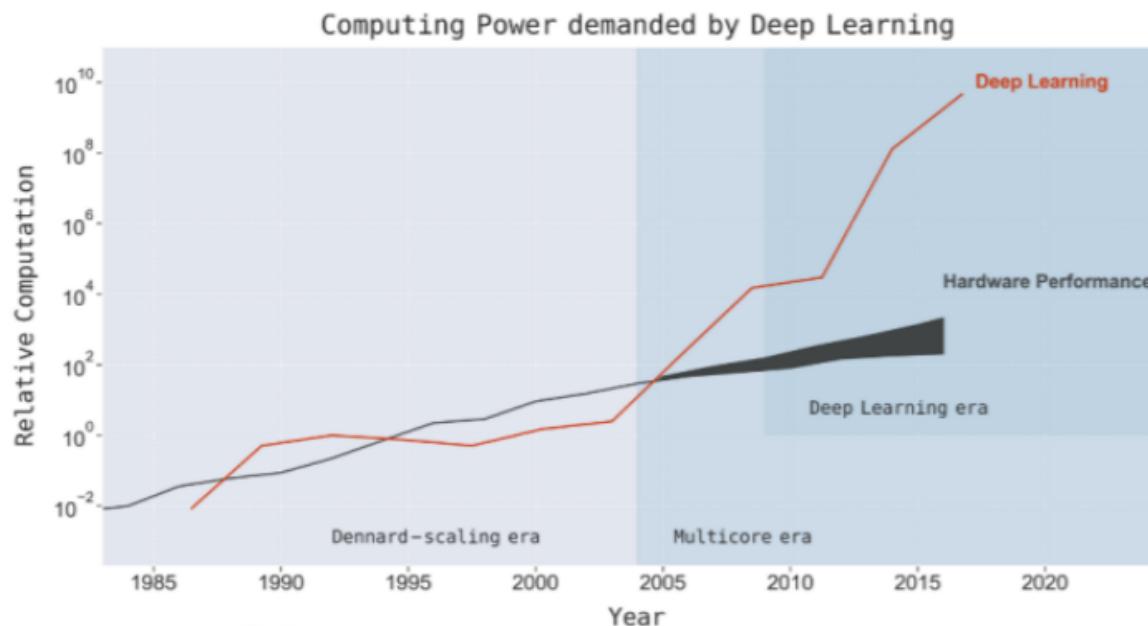
JANUARY 18, 2023 7:00 AM EST



<https://time.com/6247678/openai-chatgpt-kenya-workers/>

Some context: the role of computational power

- It is difficult to overstate the role of “more compute” in the advances we have seen in the last ~10 years
- GPT family is an excellent example of “better because it’s bigger”



<https://arxiv.org/abs/2007.05558>

<https://openai.com/research/ai-and-compute>

The screenshot shows a news article from Ars Technica. The headline reads: "DOES ANYONE THINK THIS IS A GOOD IDEA? – ChatGPT-style search represents a 10x cost increase for Google, Microsoft". Below the headline, a subtext states: "Google hints that an AI chatbot search engine will really cut into its profits." The author is listed as RON AMADEO - FEB 22, 2023 8:09 PM UTC.



<https://arstechnica.com/gadgets/2023/02/chatgpt-style-search-represents-a-10x-cost-increase-for-google-microsoft/>

How does machine learning go wrong?

What is bias?

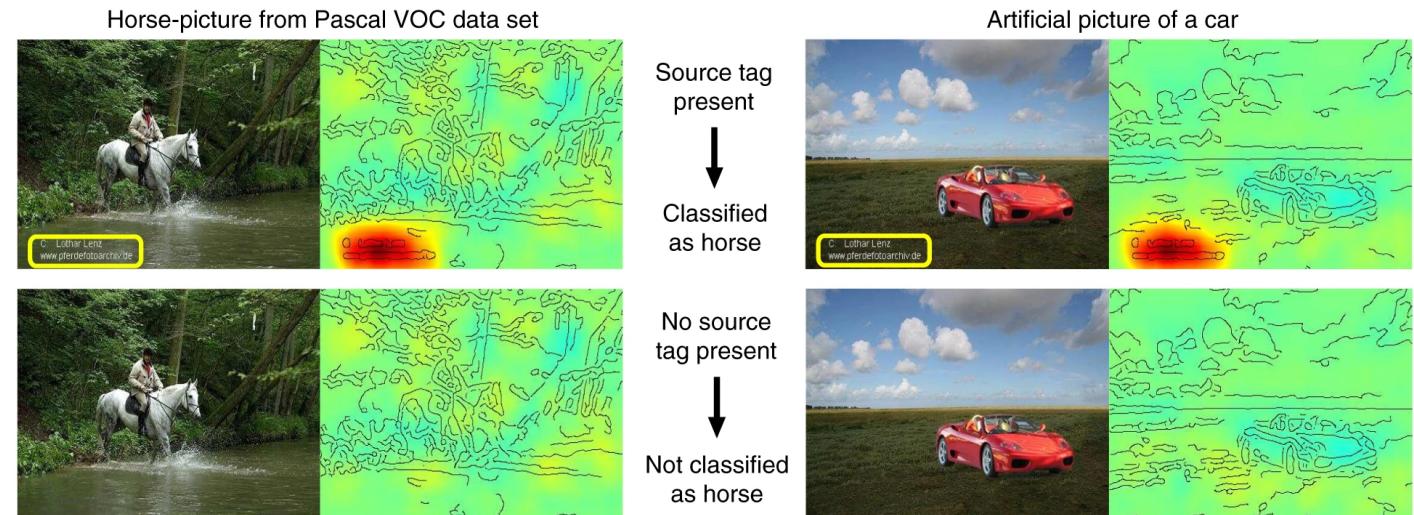
- We'll define **bias** as a tendency of a system to prefer certain types of outputs over others
- Training a system is the process of **adjusting this bias** to try and get the desired outputs for the given inputs
- Example: the outputs of a “style transfer” system (right) can be rebiased towards images that “look like” the style of a reference image



https://www.tensorflow.org/lite/examples/style_transfer/overview

What is bias?

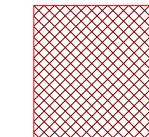
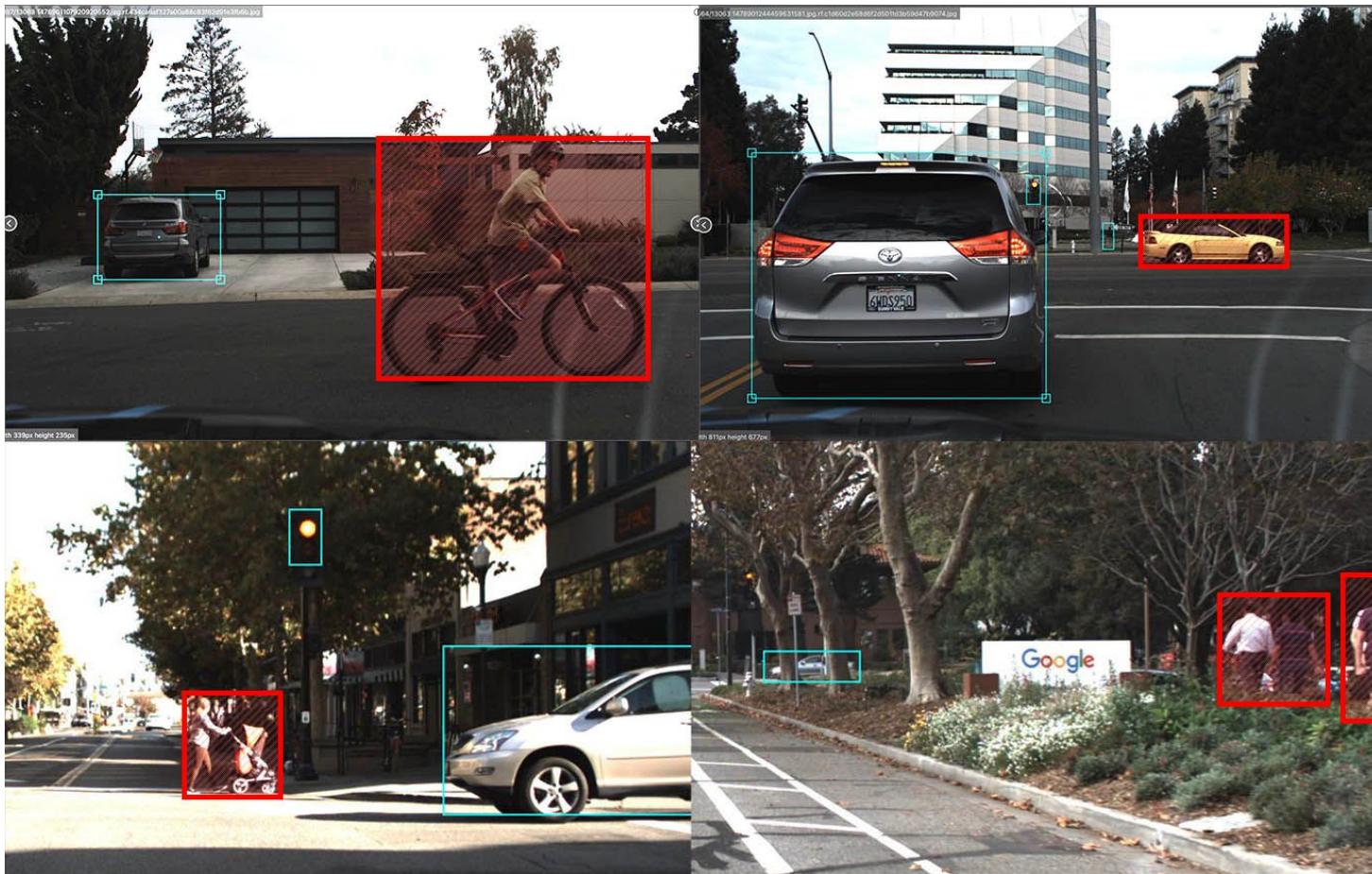
- The learned bias may not be what a human being would have in mind
- Example: this classifier has learned that the label “horse” is associated with an author’s mark that is present in the dataset
- **The goal is not to build systems without bias**, it’s to acknowledge what biases are present and understand the result of the system in that context



<https://www.nature.com/articles/s41467-019-08987-4>

How do existing systems fail?

- Incorrect/incomplete datasets are a common root issue for failures. Data is hard!



= missing labels

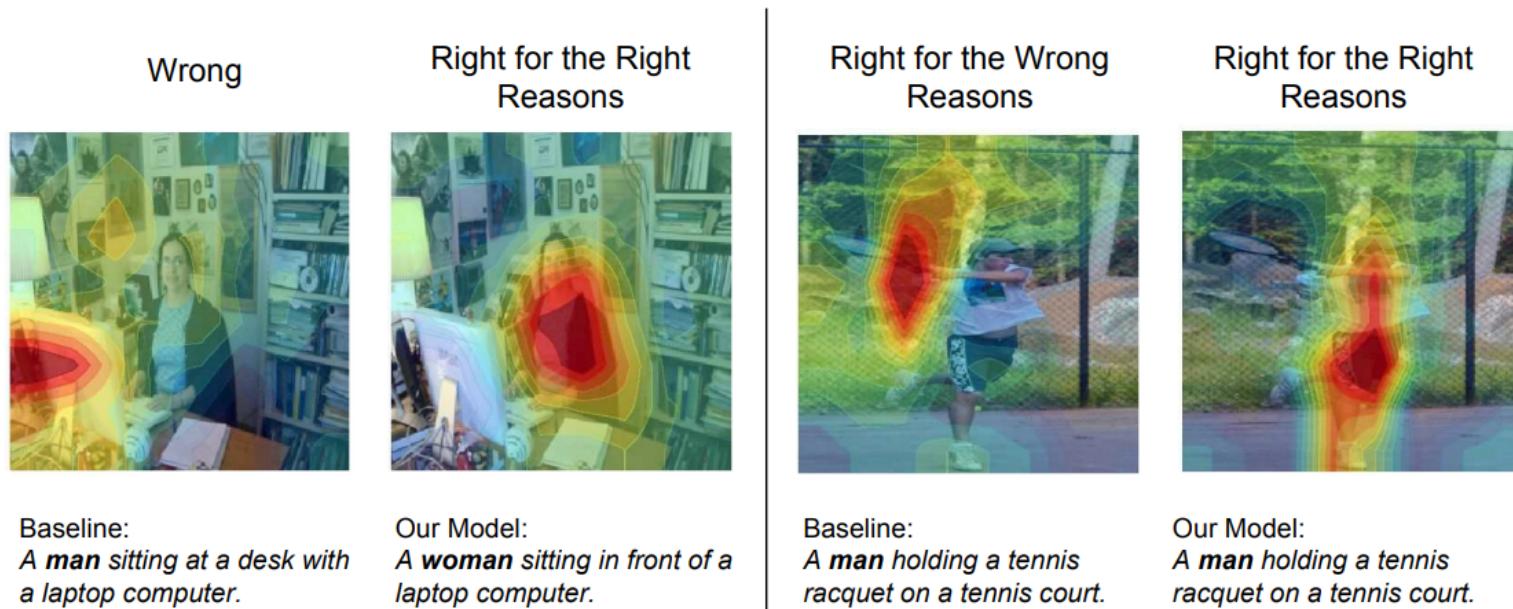
We did a hand-check of the 15,000 images in the widely used **Udacity Dataset 2** and found problems with 4,986 (33%) of them. Amongst these were thousands of unlabeled vehicles, hundreds of unlabeled pedestrians, and dozens of unlabeled cyclists. We also found many instances of phantom **annotations**, duplicated bounding boxes, and drastically oversized bounding boxes.



<https://blog.roboflow.com/self-driving-car-dataset-missing-pedestrians/>

How do existing systems fail?

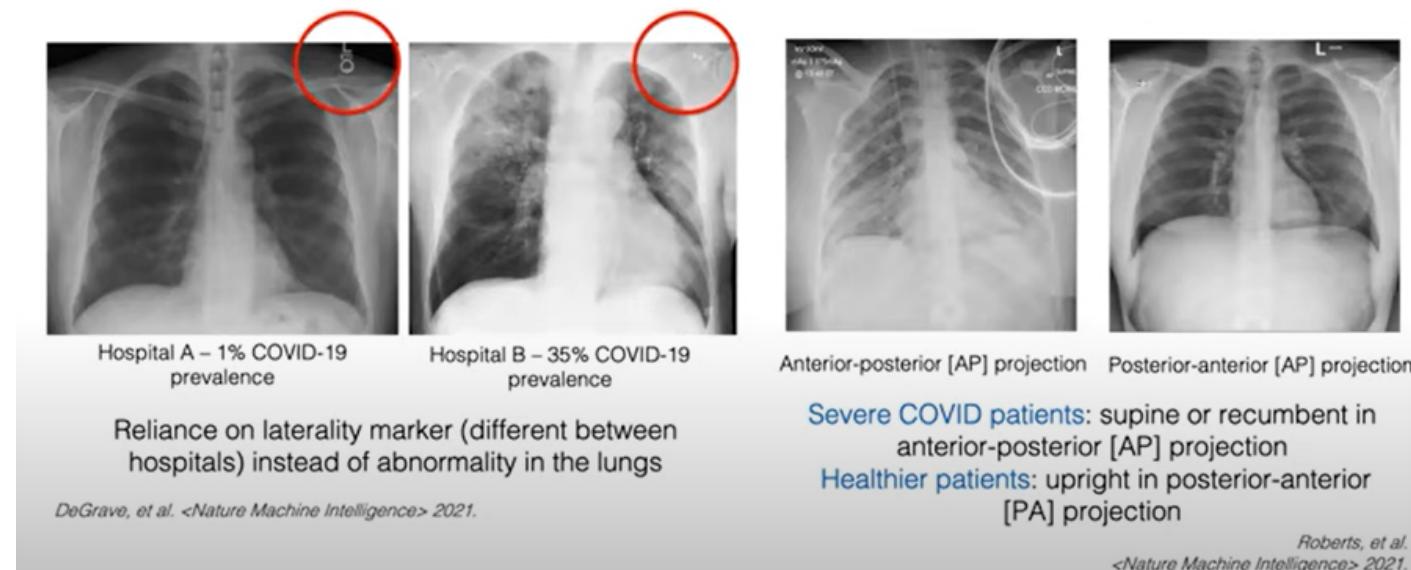
- Some failures are obvious, but more pernicious failures look like “**getting the right answer for the wrong reasons**”
- Again: the learned bias may not be what a human being would have in mind!



https://openaccess.thecvf.com/content_ECCV_2018/papers/Lisa_Anne_Hendricks_Women_also_Snowboard_ECCV_2018_paper.pdf

How do existing systems fail?

- Of 232 algorithms for diagnosing patients or predicting COVID severity, only **2** were promising enough to explore further (Wyants *et al.* BMJ, 2020)
- Of 415 computer vision models for diagnosing and predicting COVID, **none** were fit for deployment in real world settings (Roberts *et al.* Nature Machine Intelligence, 2021)



<https://www.nature.com/articles/s42256-021-00338-7>



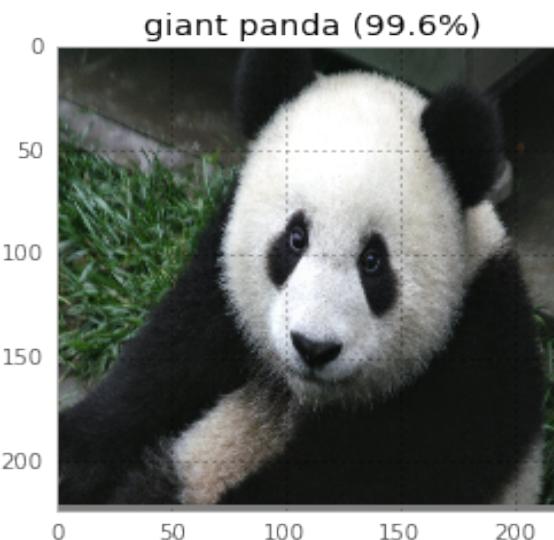
<https://www.nature.com/articles/s42256-021-00307-0>



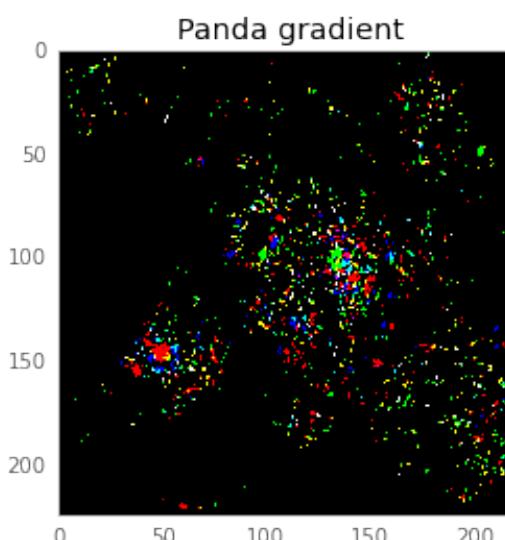
How can we understand?

Going backwards

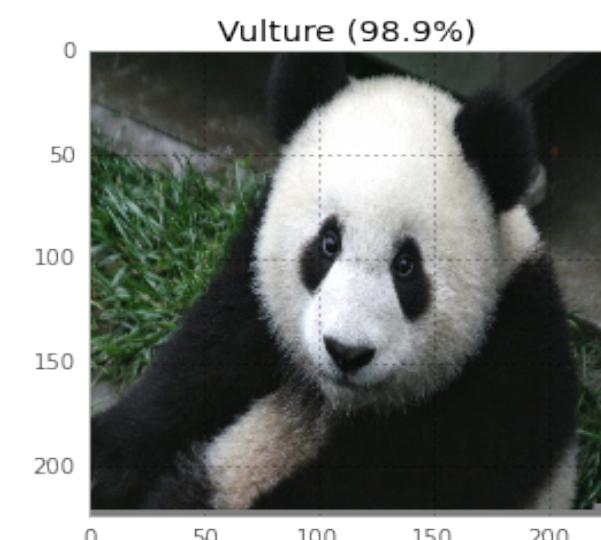
- The principle of **backpropagation** used for training can also be used to determine which parts of an input are most relevant to an inference
 - Example: the **gradient** here tells us the importance of each pixel in this image being classified as a panda.



-



=



<https://codewords.recurse.com/issues/five/why-do-neural-networks-think-a-panda-is-a-vulture>

Going backwards

- For models built with PyTorch, the library **Captum** provides some ‘out of the box’ tools for generating these **attributions**

In [10]:

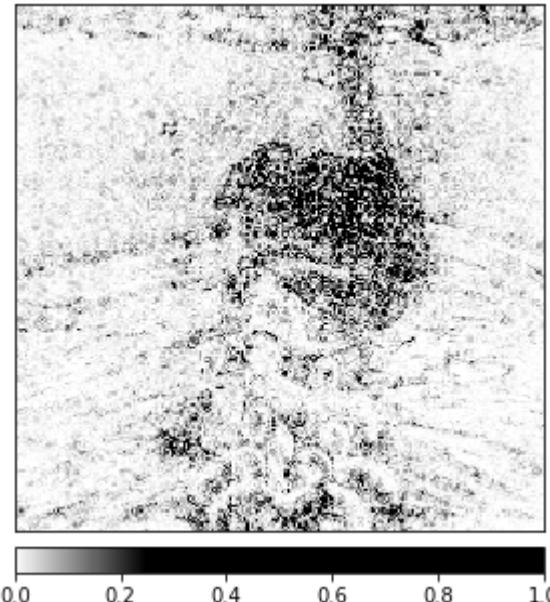
```
torch.manual_seed(0)
np.random.seed(0)

gradient_shap = GradientShap(model)

# Defining baseline distribution of images
rand_img_dist = torch.cat([input * 0, input * 1])

attribution_gs = gradient_shap.attribute(input,
                                         n_samples=50,
                                         stdevs=0.0001,
                                         baselines=rand_img_dist,
                                         target=pred_label_idx)

_ = viz.visualize_image_attr_multiple(np.transpose(attributions_gs.squeeze().cpu().detach().numpy(), (1,2,0)),
                                      np.transpose(transformed_img.squeeze().cpu().detach().numpy(), (1,2,0)),
                                      ["original_image", "heat_map"],
                                      ["all", "absolute_value"],
                                      cmap=default_cmap,
                                      show_colorbar=True)
```



https://captum.ai/tutorials/Resnet_TorchVision_Interpret

Going backwards (text)

- **Captum** also provides some tools for working with text-based models, e.g. figuring out which words are most important to an inference performed on text
- The idea of “going backwards” is a very general one

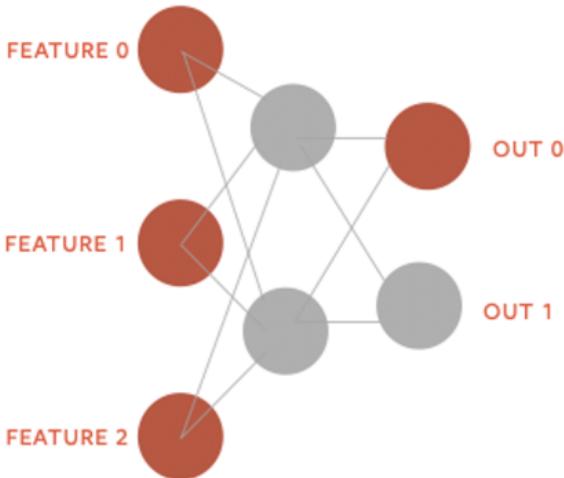
Legend: ■ Negative □ Neutral ■ Positive

True Label	Predicted Label	Attribution Label	Attribution Score	Word Importance
pos	pos (0.96)	pos	1.29	it was a fantastic performance ! #pad
pos	pos (0.87)	pos	1.56	best film ever #pad #pad #pad #pad
pos	pos (0.92)	pos	1.14	such a great show ! #pad #pad
neg	neg (0.29)	pos	-1.11	it was a horrible movie #pad #pad
neg	neg (0.22)	pos	-1.03	i 've never watched something as bad
neg	neg (0.07)	pos	-0.84	that is a terrible movie . #pad



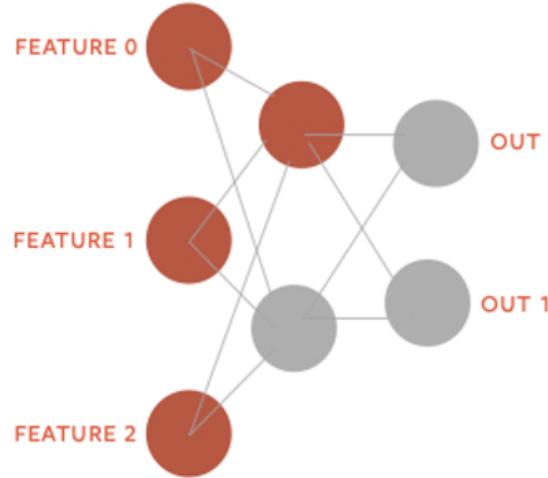
https://captum.ai/tutorials/IMDB_TorchText_Interpret

Overview of Captum's attribution types



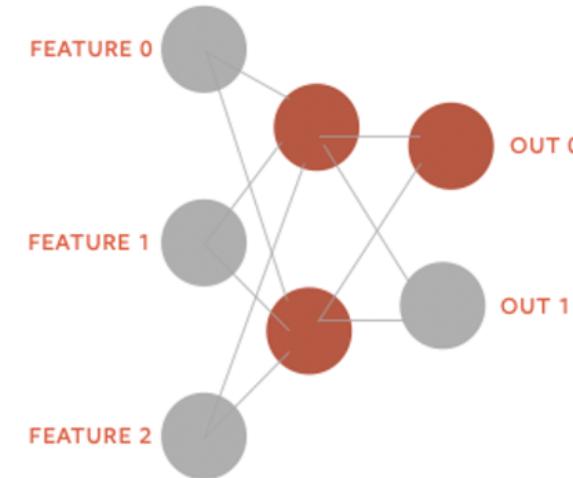
PRIMARY ATTRIBUTION

```
from captum.attr import IntegratedGradients  
  
attr_algo = IntegratedGradients(model)  
  
# Returned attributions have the same shape  
# as the inputs  
attributions = attr_algo.attribute(model, inputs)
```



NEURON ATTRIBUTION

```
from captum.attr import NeuronGradientShap  
  
attr_neuron_algo = NeuronGradientShap(model,  
model.layer1)  
  
# Returned attributions have the same shape  
# as the inputs  
attributions = attr_neuron_algo.attribute(inputs,  
neuron_index=0)
```



LAYER ATTRIBUTION

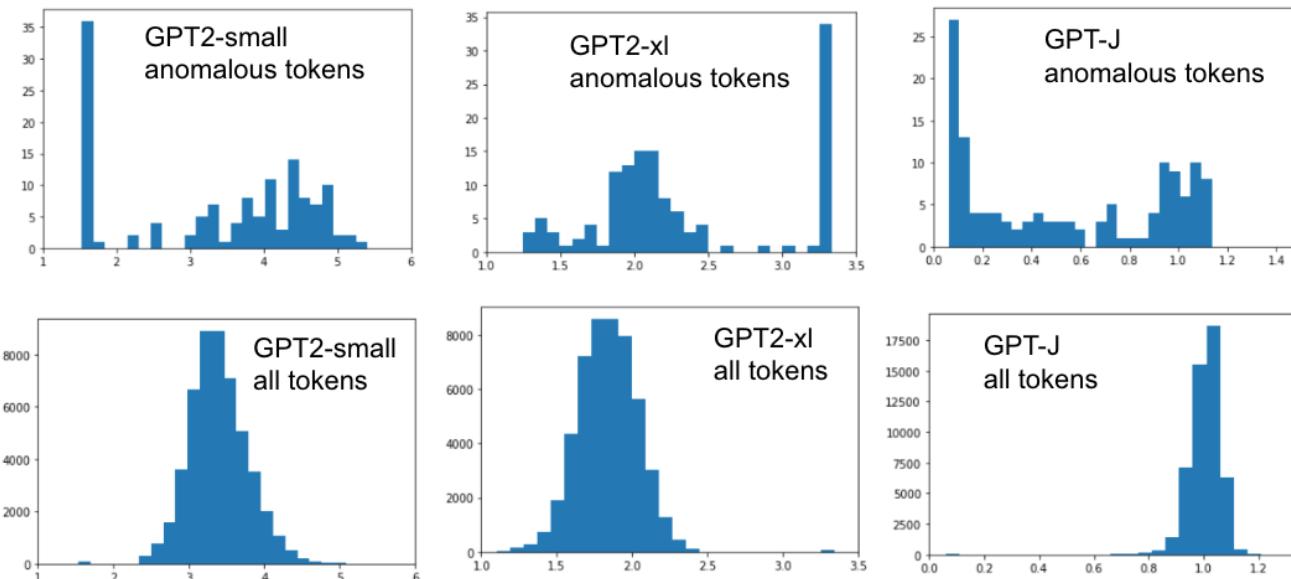
```
from captum.attr import LayerConductance  
  
attr_layer_algo = LayerConductance(model,  
model.layer1,  
attribute_to_layer_input=False)  
  
# Returned attributions have the same shape  
# as the output of the first layer  
attributions = attr_layer_algo.attribute(inputs)
```



<https://arxiv.org/abs/2009.07896>

Case study: 'glitch' tokens in GPT

- Text-based models have unique challenges compared to image-based models because there is not a direct analogue of pixels...
 - ...but we can still make some explanations of these models
- Analysis of GPT's encoding reveals clusters of anomalous tokens



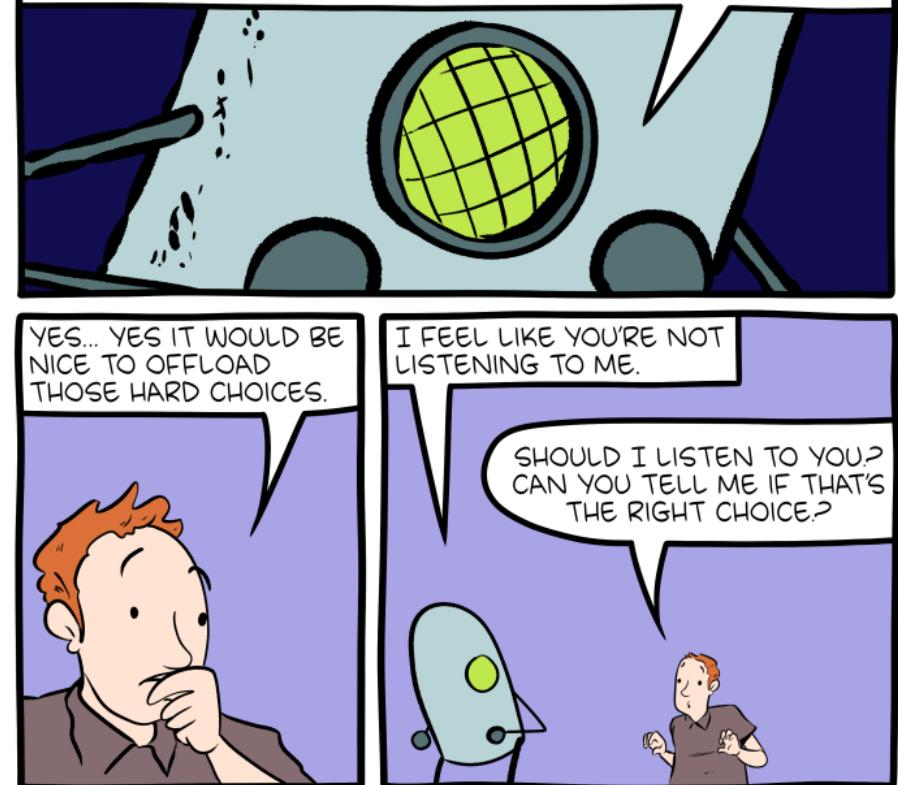
<https://www.alignmentforum.org/posts/aPeJE8bSo6rAFoLqg/solidgoldmagikarp-plus-prompt-generation>

Some final thoughts

Final thoughts

- What are we building?

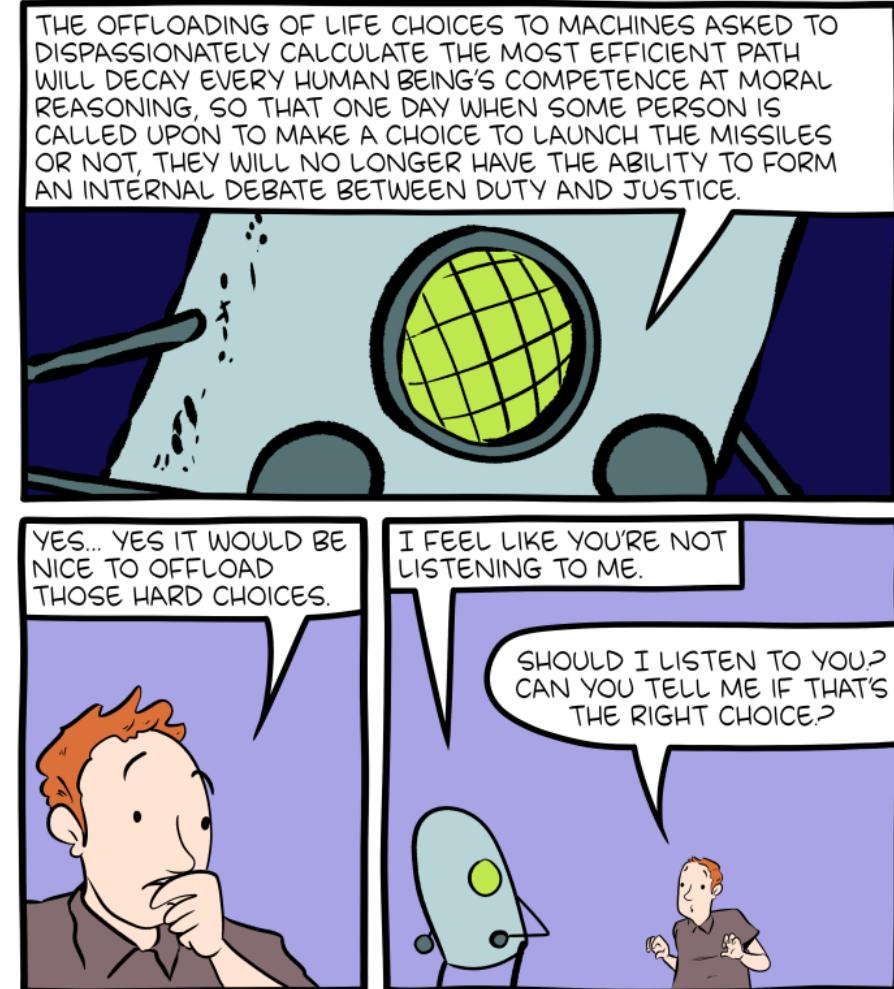
THE OFFLOADING OF LIFE CHOICES TO MACHINES ASKED TO DISPASSIONATELY CALCULATE THE MOST EFFICIENT PATH WILL DECAY EVERY HUMAN BEING'S COMPETENCE AT MORAL REASONING, SO THAT ONE DAY WHEN SOME PERSON IS CALLED UPON TO MAKE A CHOICE TO LAUNCH THE MISSILES OR NOT, THEY WILL NO LONGER HAVE THE ABILITY TO FORM AN INTERNAL DEBATE BETWEEN DUTY AND JUSTICE.



smbc-comics.com

Final thoughts

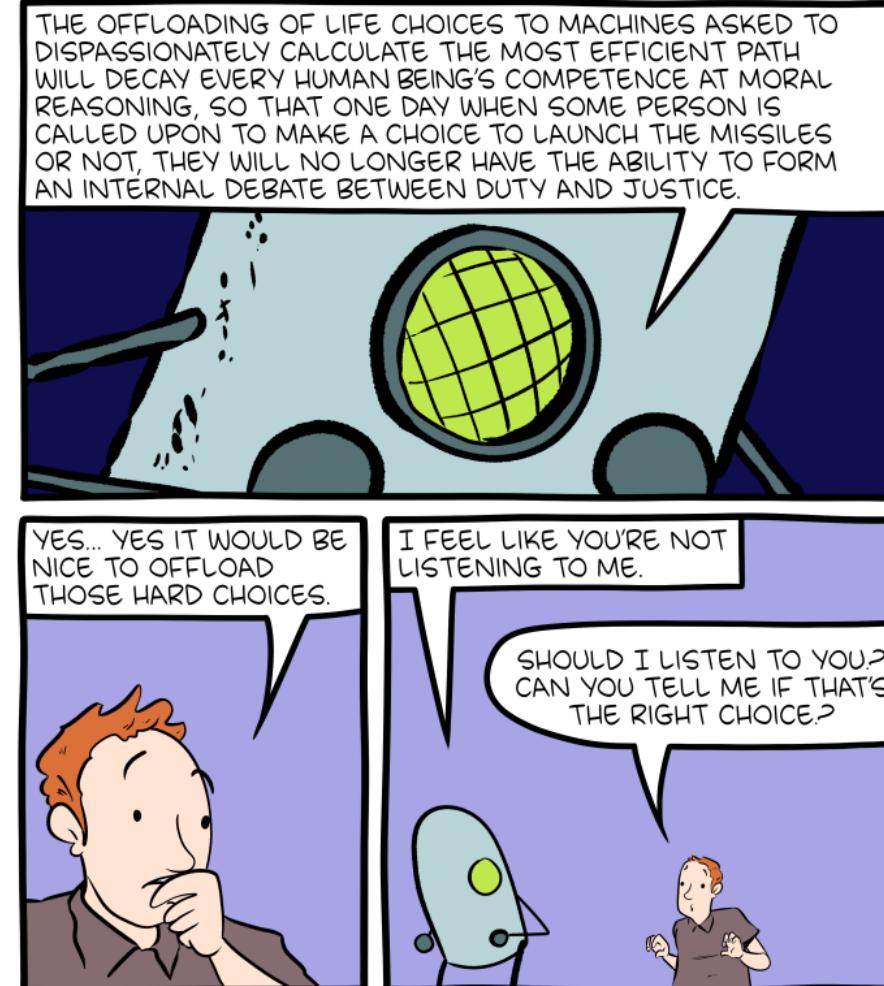
- What are we building?
- **Why** are we building? What problem(s) are we trying to solve with these models?



smbc-comics.com

Final thoughts

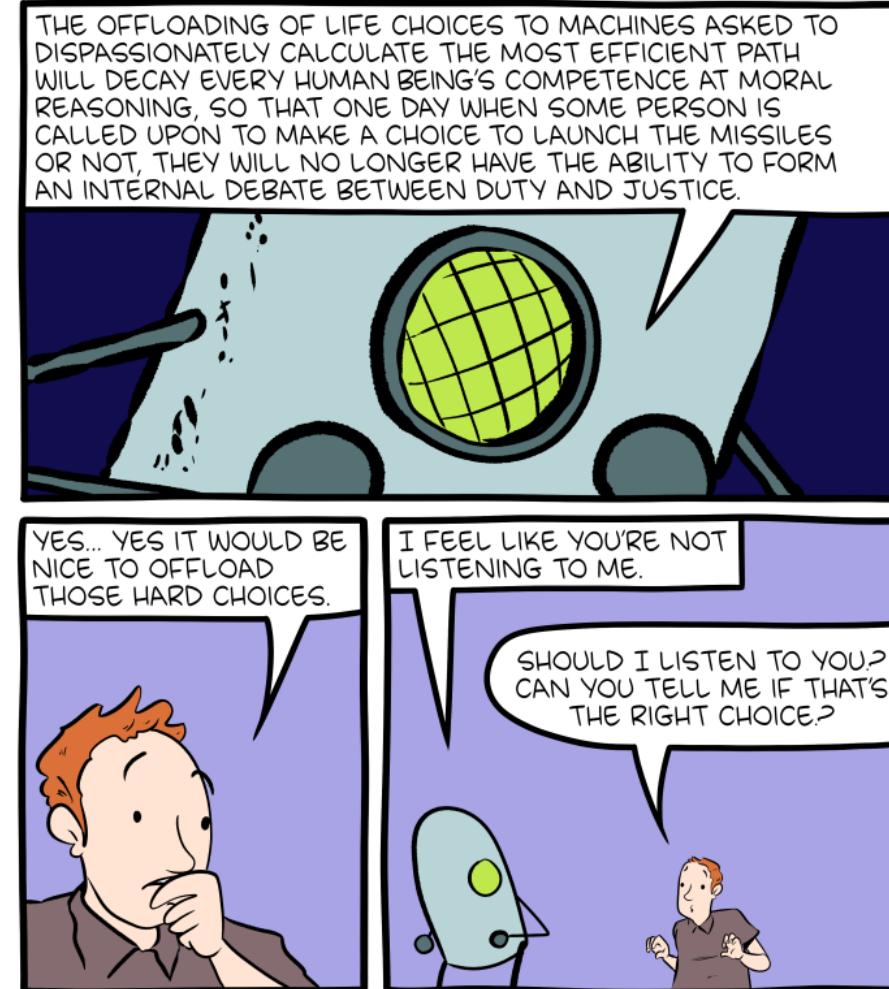
- What are we building?
- **Why** are we building? What problem(s) are we trying to solve with these models?
- Do we understand our models?



smbc-comics.com

Final thoughts

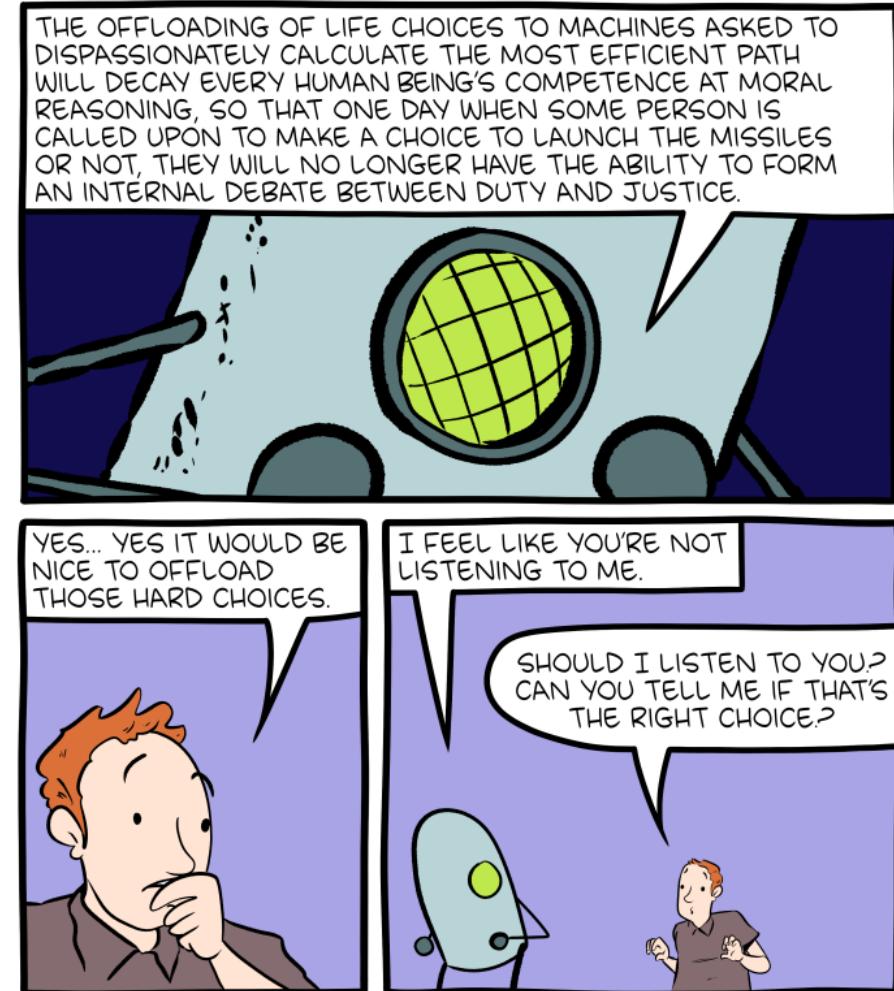
- What are we building?
- **Why** are we building? What problem(s) are we trying to solve with these models?
- ~~Do we understand our models?~~
- How can we improve our understanding of our models?



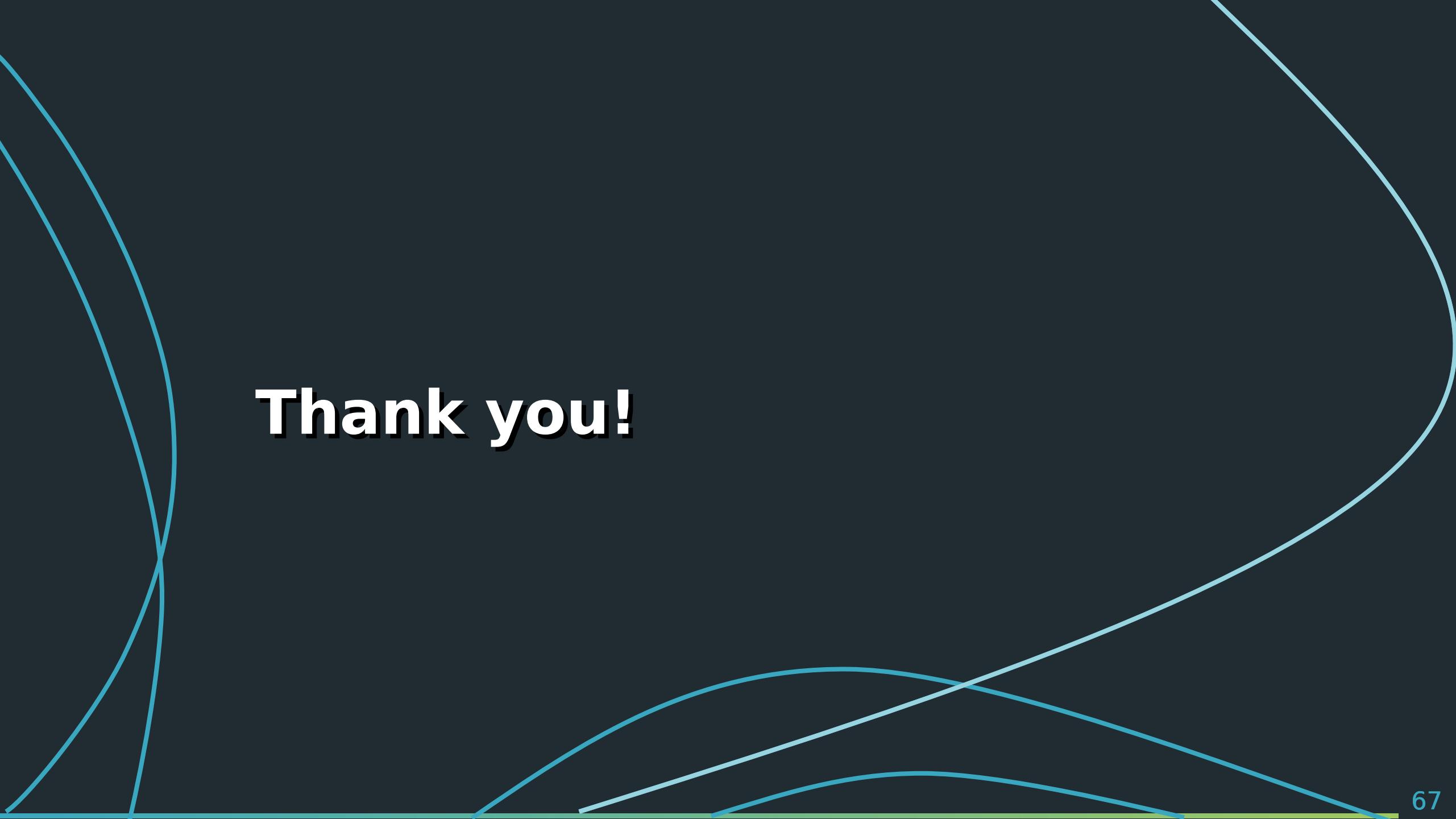
smbc-comics.com

Final thoughts

- What are we building?
- **Why** are we building? What problem(s) are we trying to solve with these models?
- ~~Do we understand our models?~~
- How can we improve our understanding of our models?
- How can we prioritize understandable systems in a sector where the pace of change is so rapid?



smbc-comics.com



Thank you!