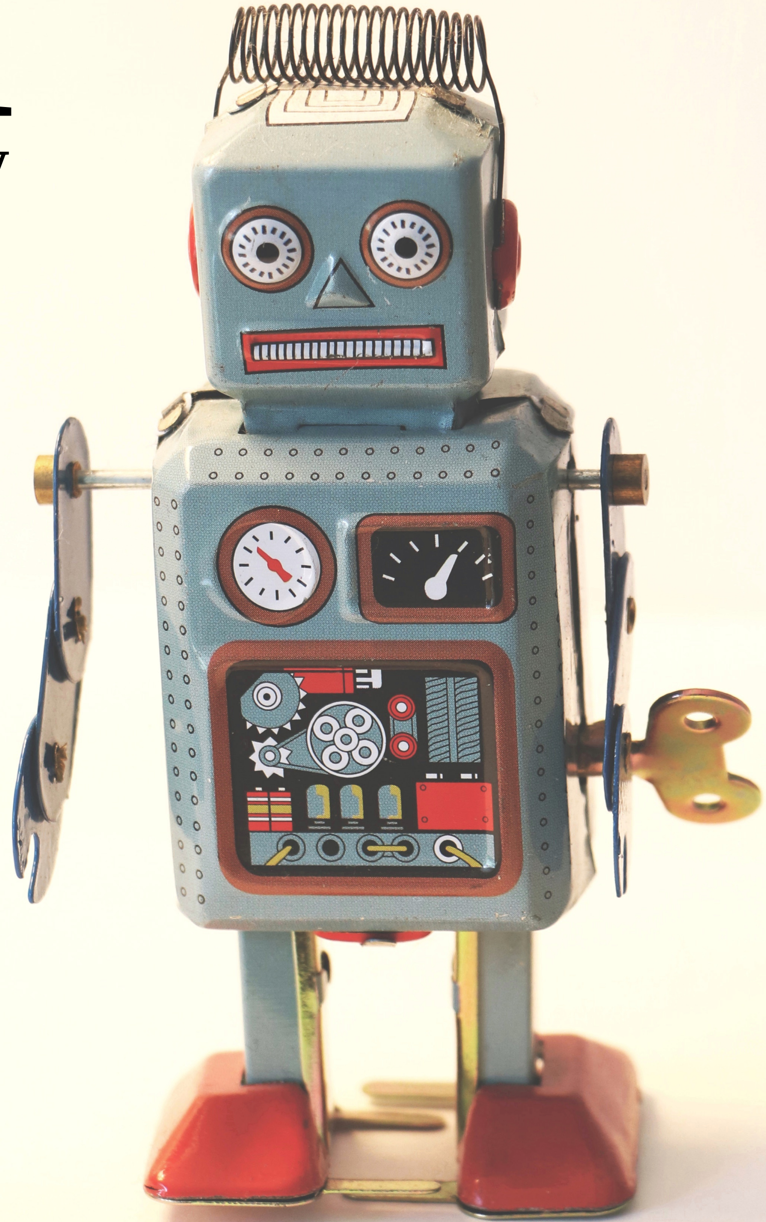
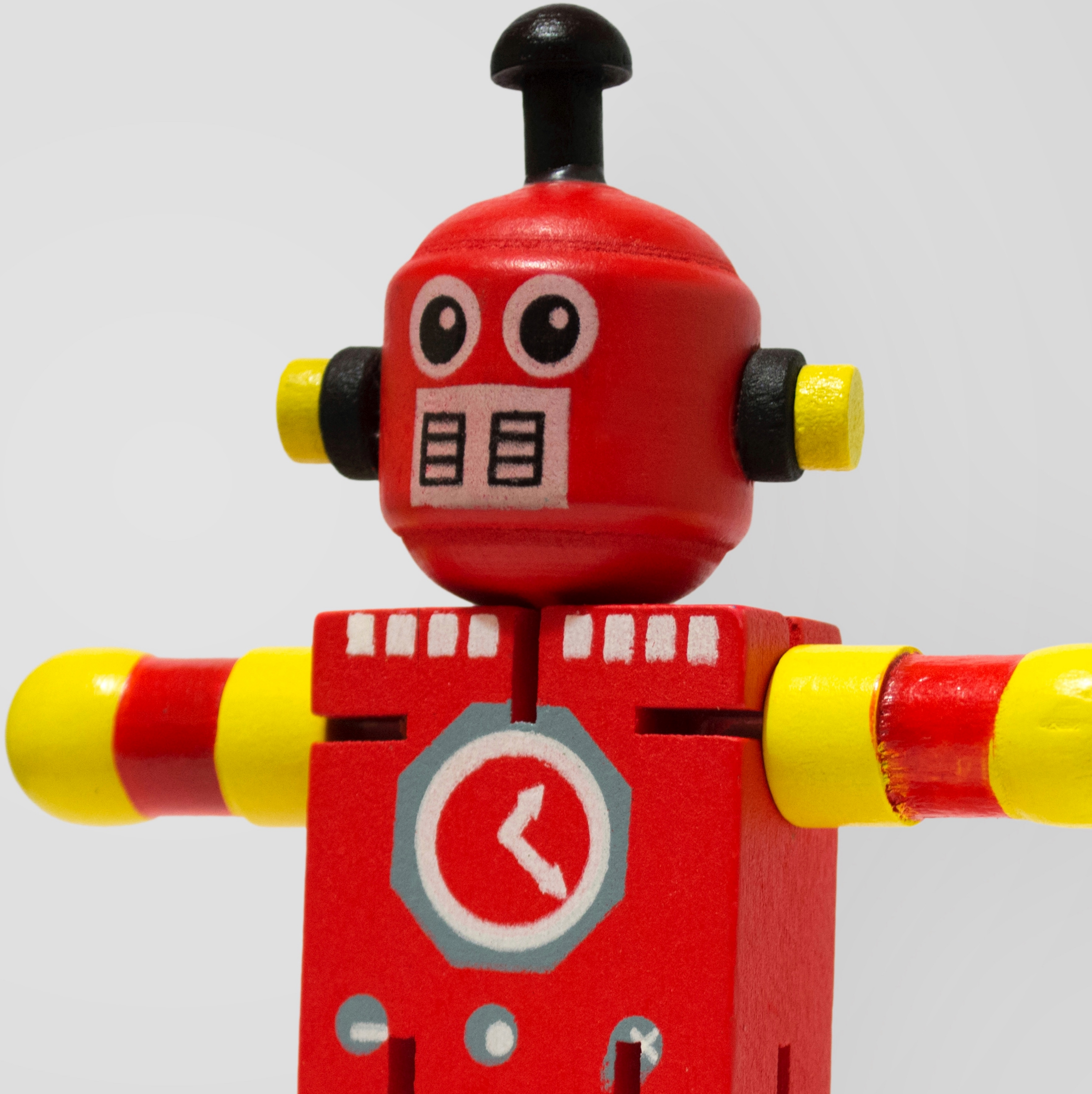
AI Is Broken

Sophie Searcy



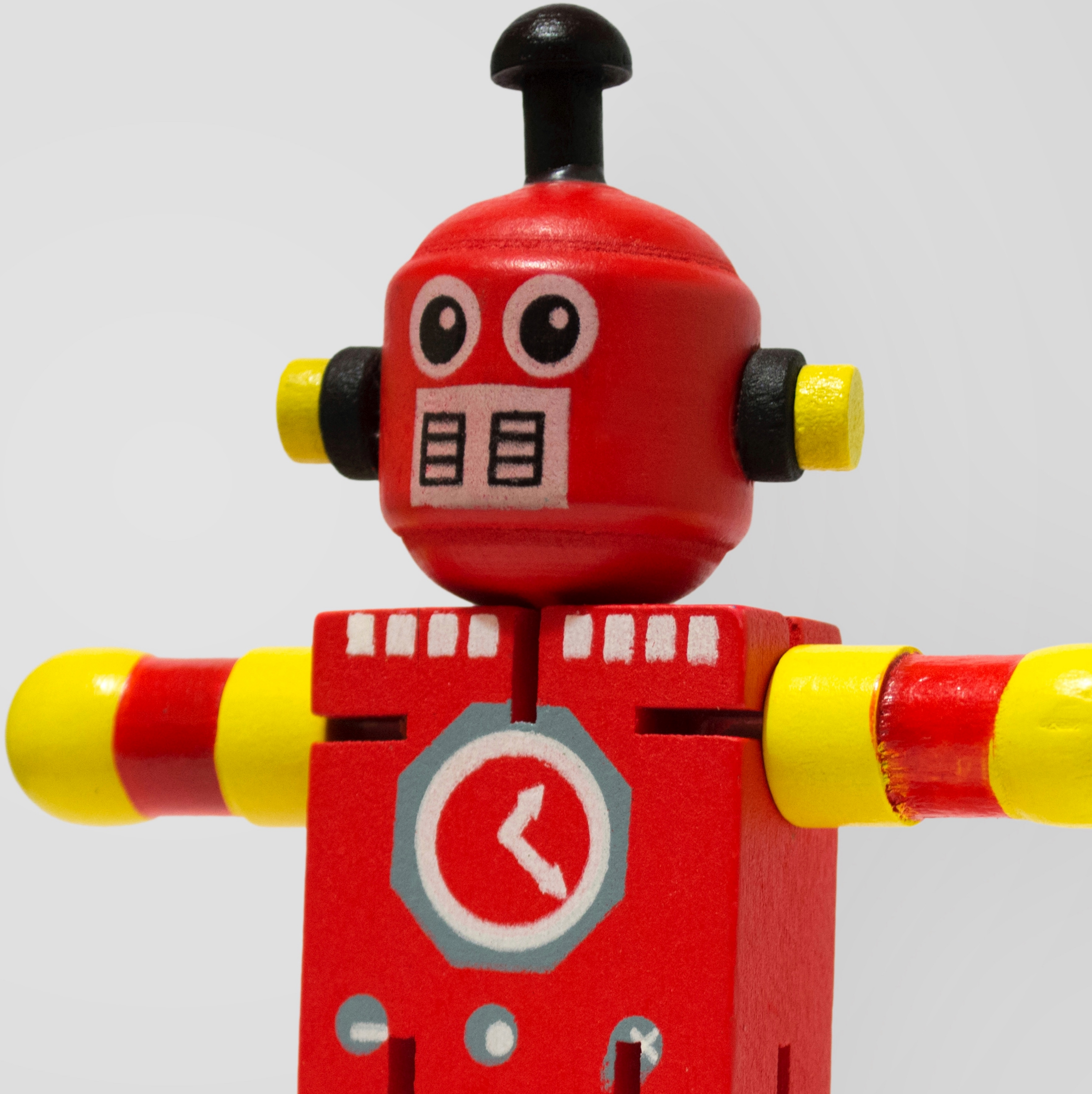
What is AI?



Caveats

- AI
  - lumping together Data Science, Artificial Intelligence, Machine Learning, Data Mining, etc.
- Audience
  - Conversant in AI topics. Not necessarily experts or practicioners.

What is AI?



Learn verb

\’lern\

to process past experience and update a model such that the the model is more useful for future experience

Learn verb

\’lern\

to process past experience and update a model such that the the model is useful for future experience

Learn verb

\’lern\

to process past experience and update a model such that the the model is useful for future experience

Learn verb

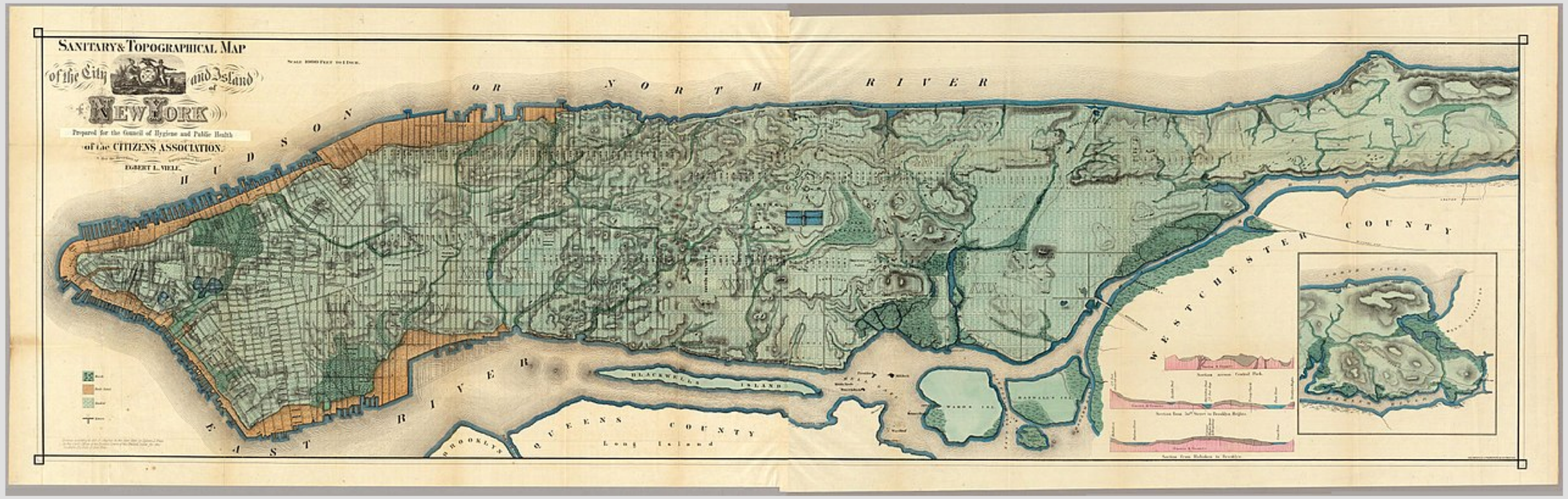
\’lern\

to process past experience and update a model

such that the the model is useful for future experience

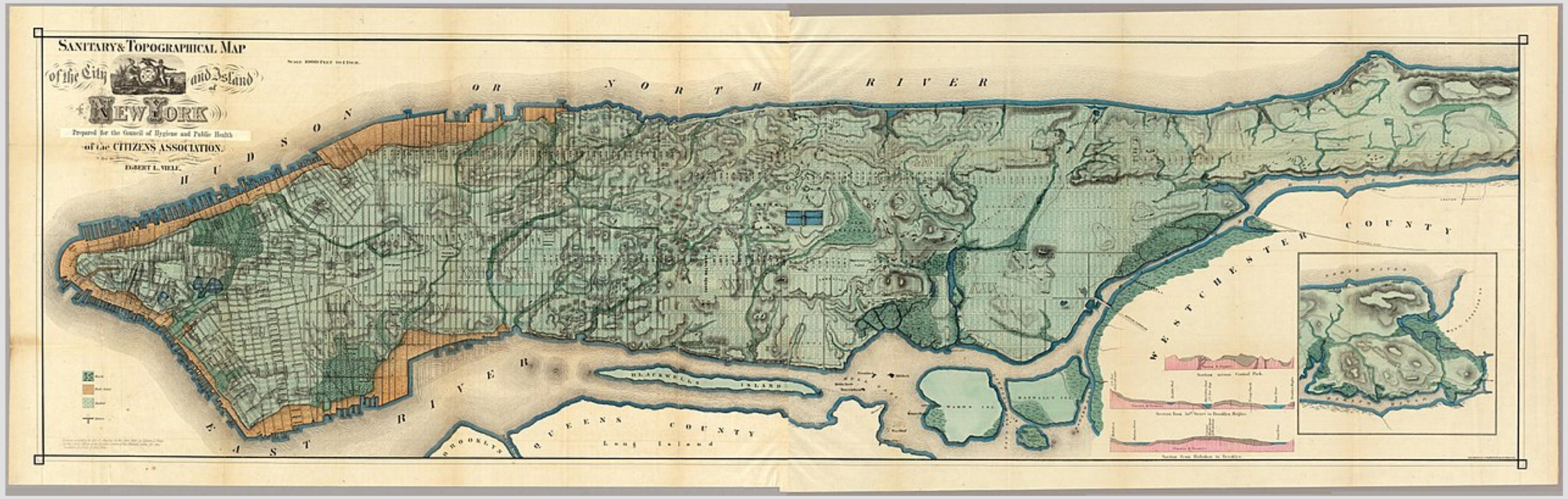
Model: a learning algorithm

- A model is a small thing that captures a larger thing.
- A good model omits unimportant details while retaining what's important.



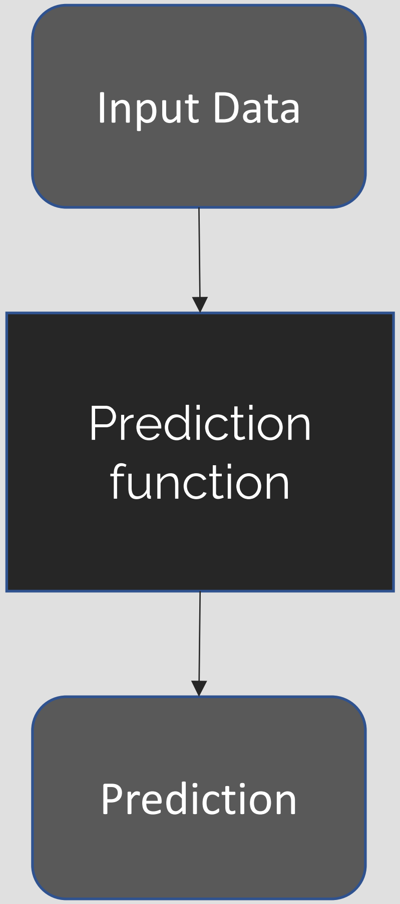
Model: a learning algorithm

- Industry sometimes uses "algorithm" and "model" interchangeably.
  - Words are complicated (ask anyone who works in NLP)



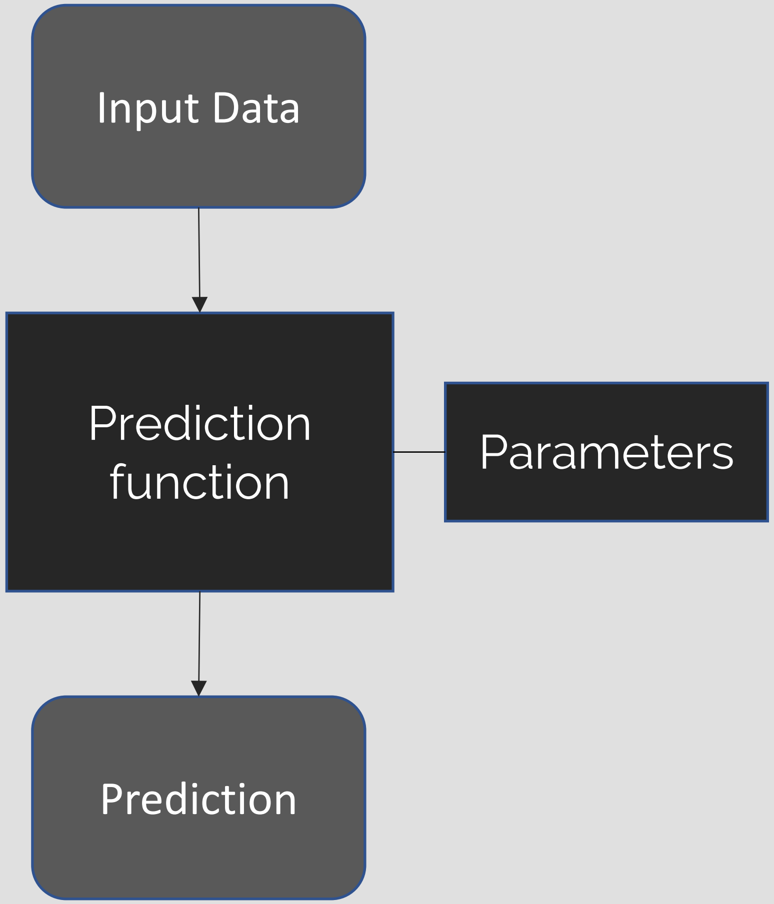
Model: a learning algorithm

- All models contain a prediction function

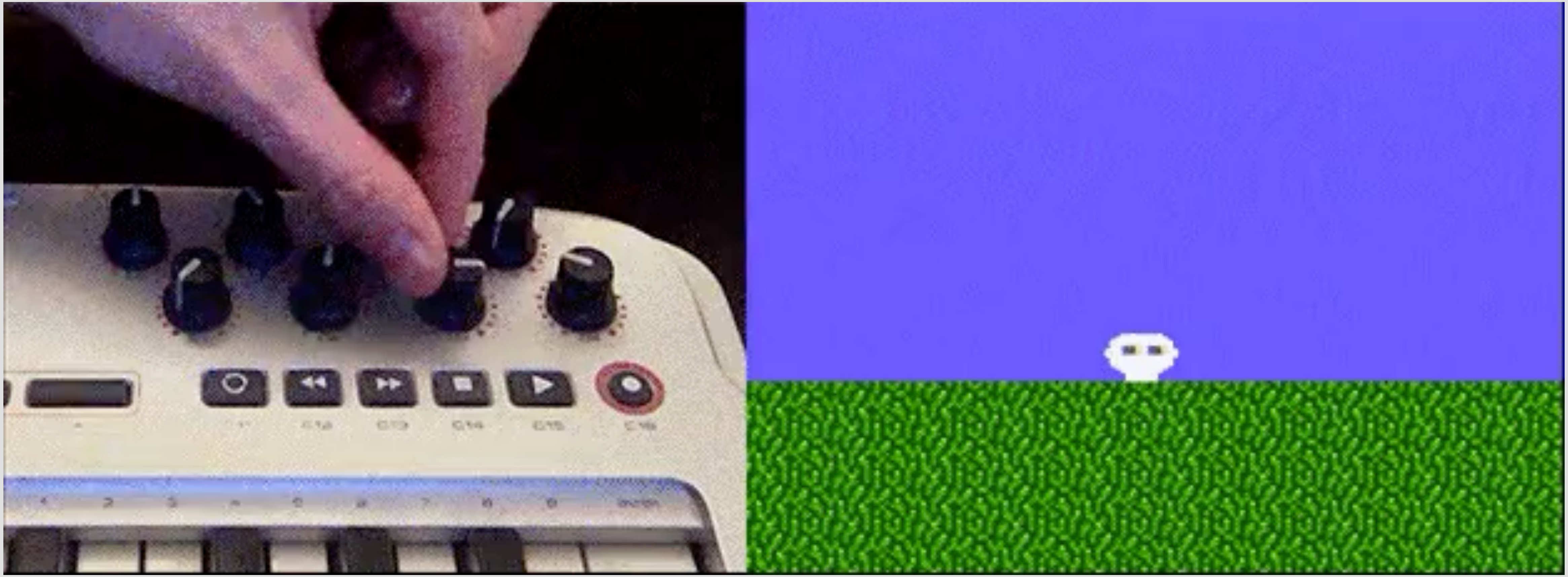


Model: a learning algorithm

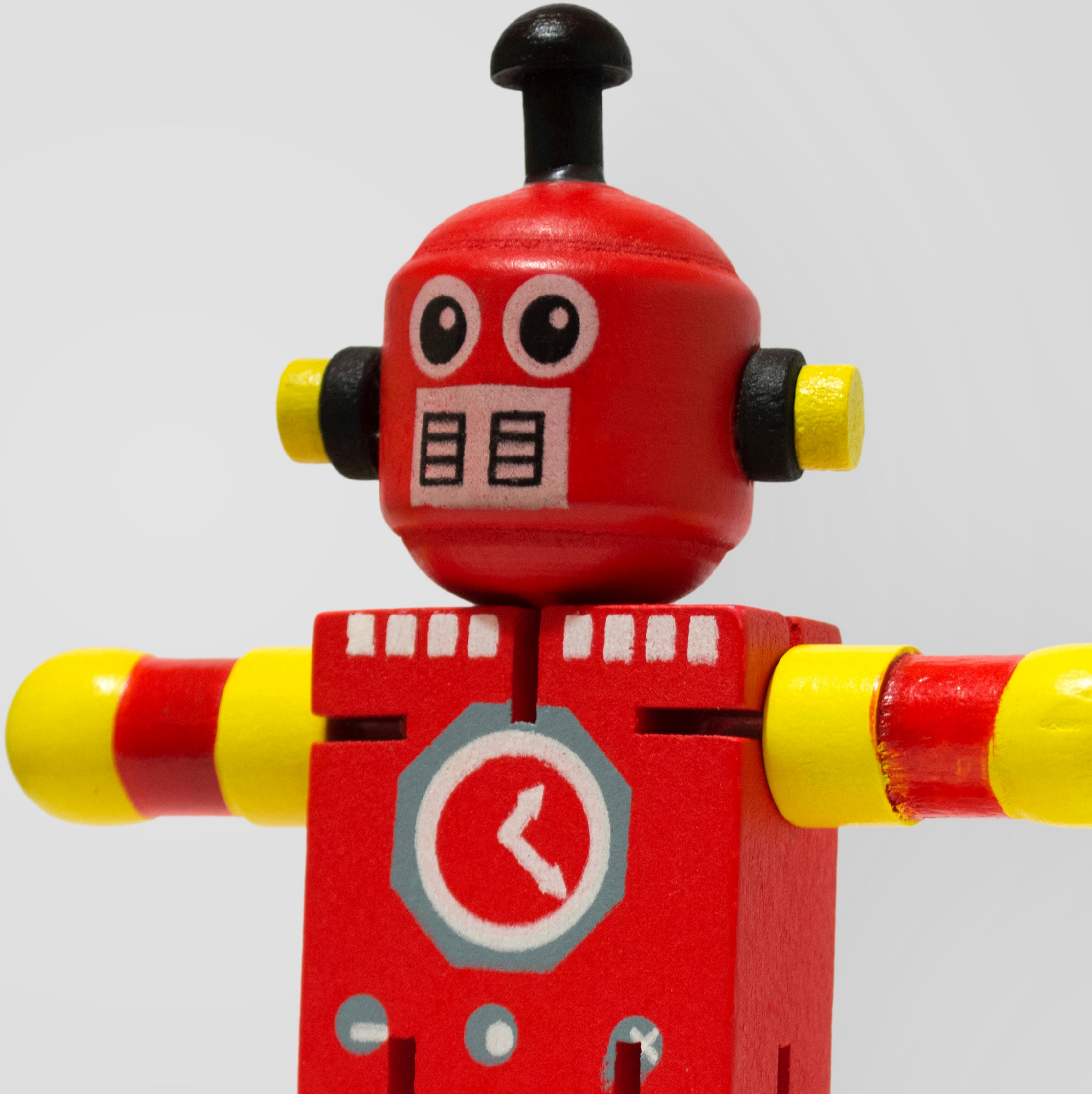
- Parameters
  - Determine model output
  - Learned from data



Model: a learning algorithm



Models are data hungry



Models are data hungry

Models

- Learn from a limited set of training data
- Apply what was learned to *production*
    - “Production” is data science lingo for the entire world

Models are data hungry

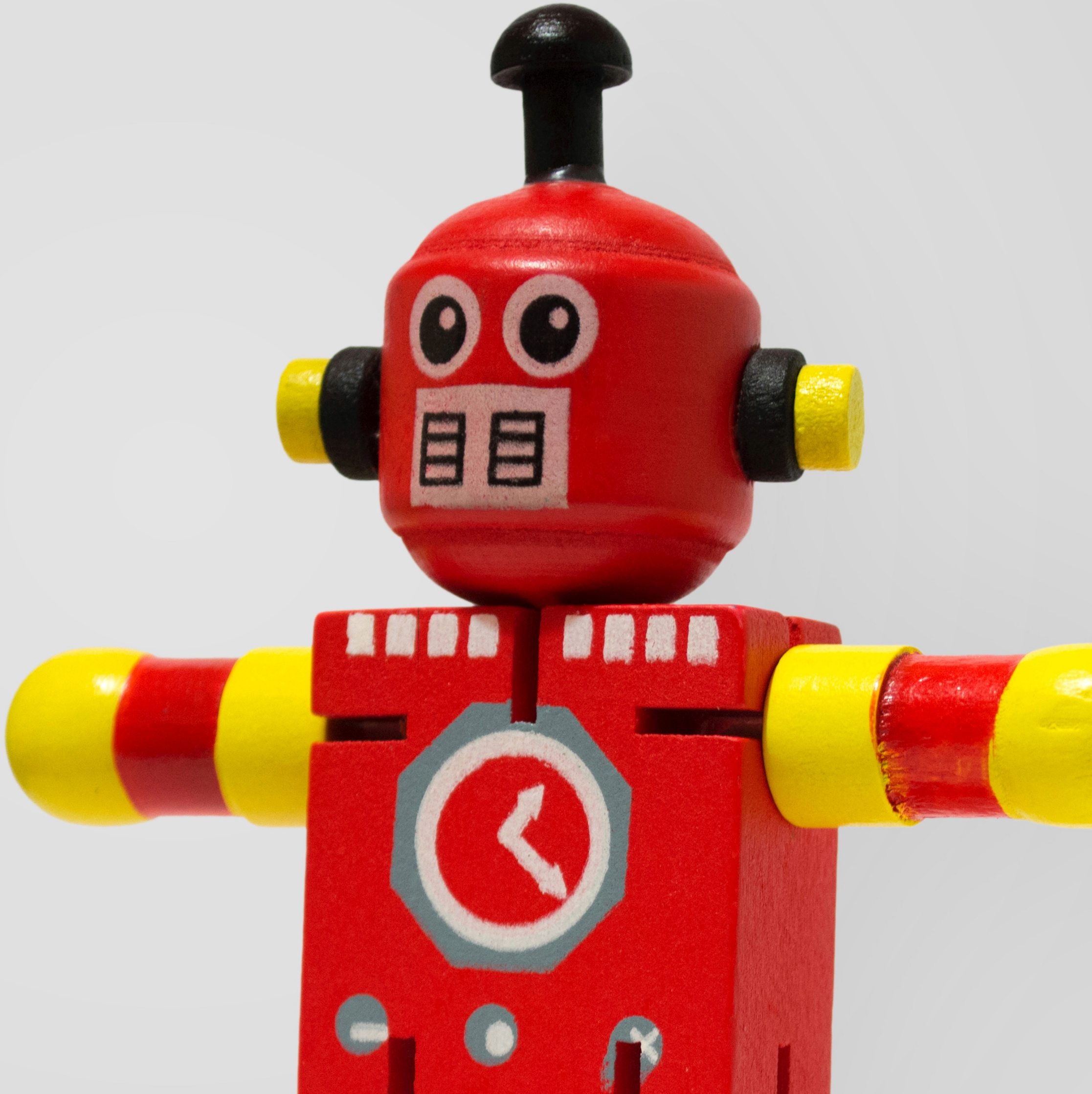
Models

- Learn from a limited set of training data
- Apply what was learned to *production*
  - "Production" is data science lingo for the entire world

One of the most difficult tasks in AI:

- use training data (data you have) to judge how a model will perform in production (data you don't have) .

Speed limits for data



Speed limits for data

“Traditional” models (Support Vector Machines, Linear Models, Random Forests, K Nearest Neighbors)

- Batch data: look at the entire dataset at once.
- Training time increases with dataset size.

Speed limits for data

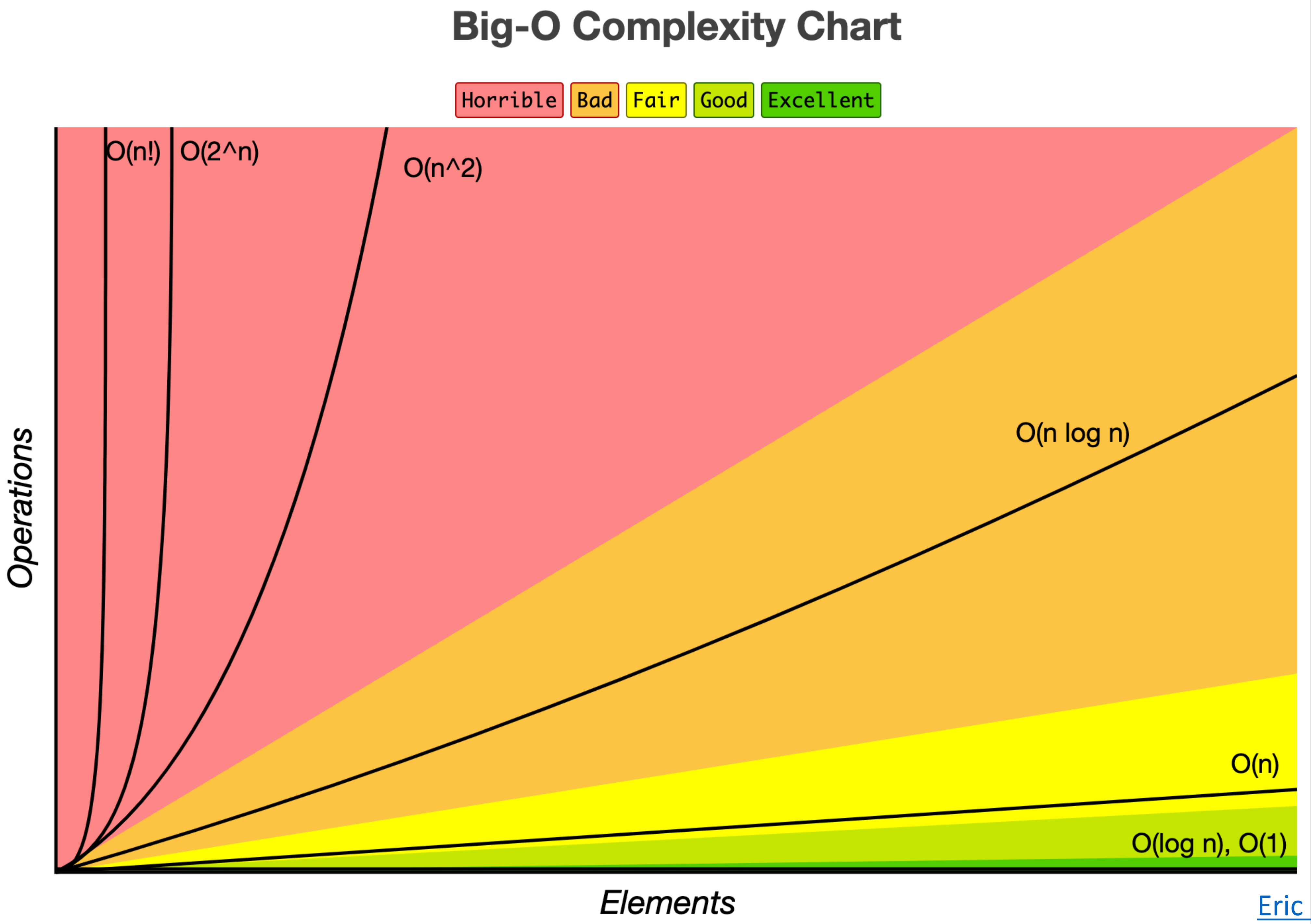
Data Set Size	Time to train
	
	
	

Speed limits for data

Data Set Size	Time to train
	
	
	

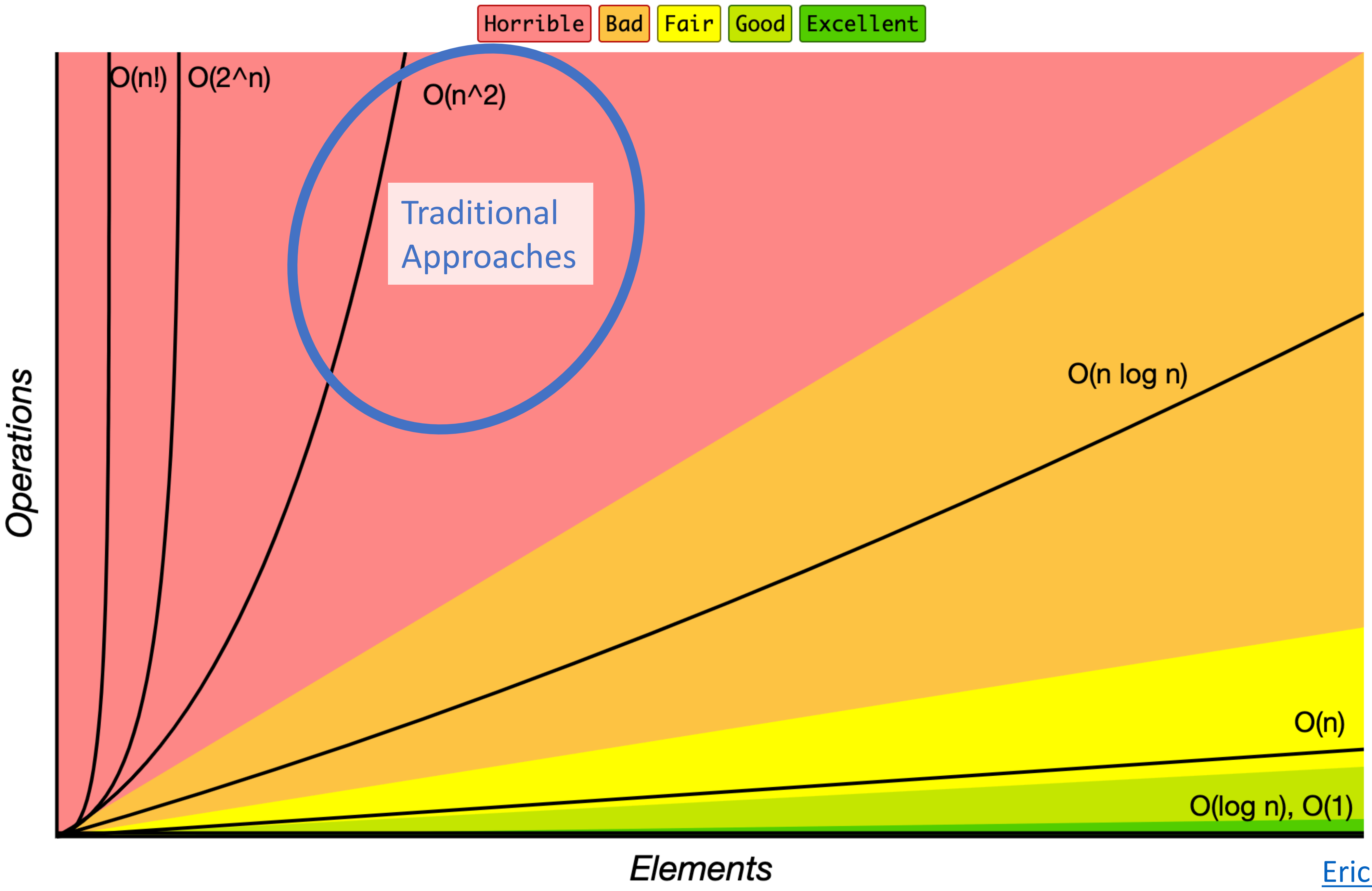
Speed limits for data

Data Set Size	Time to train
	
	
	



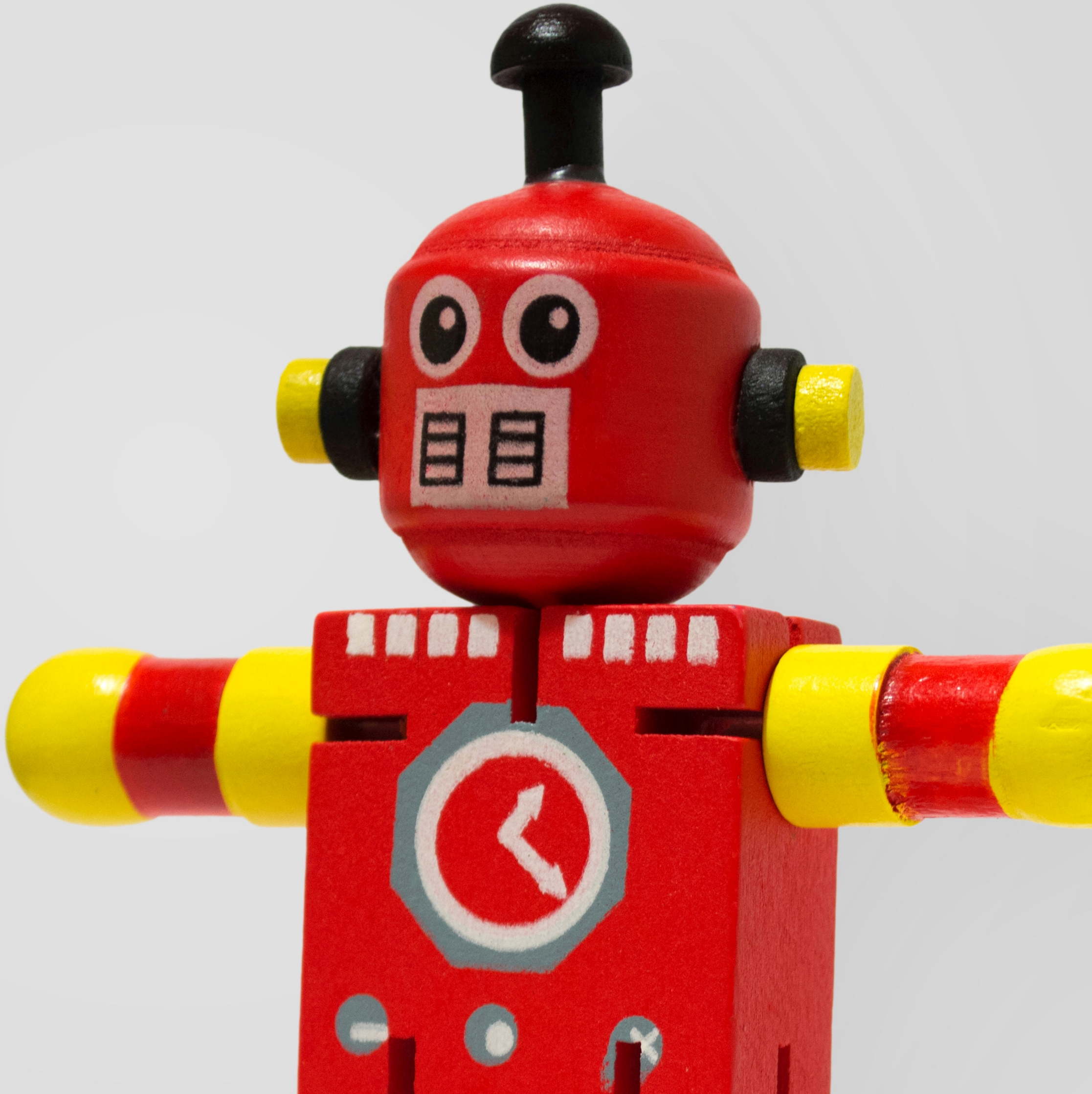
Eric Drowel et al.

Big-O Complexity Chart



Eric Drowel et al.

Modern AI removes the speed limit



Enter Stochastic Gradient Descent

- In the last two decades, AI has shifted to approaches that strongly incentivize large datasets
- SGD powers Deep Learning models
- Traditional AI models have been modified to take advantage of SGD

How does SGD work?

Gradient descent (*not stochastic*)

1. Put a number on your model's performance. (Loss function)
2. Determine which direction decreases the loss function. (Find the Gradient).
3. Turn the knob in that direction. (Backpropagation)

(Wash, rinse, repeat for every parameter)

How does SGD work?

Stochastic Gradient Descent:

- Use a small subset of your dataset to estimate the loss for the entire dataset (Minibatch)

The Tradeoffs of Large Scale Learning

Léon Bottou
NEC laboratories of America
Princeton, NJ 08540, USA
leon@bottou.org

Olivier Bousquet
Google Zürich
8002 Zurich, Switzerland
olivier.bousquet@m4x.org

- For SGD-based models, the amount of time it takes to fit a model **does not depend on the size of the dataset**.

Stochastic Gradient Descent

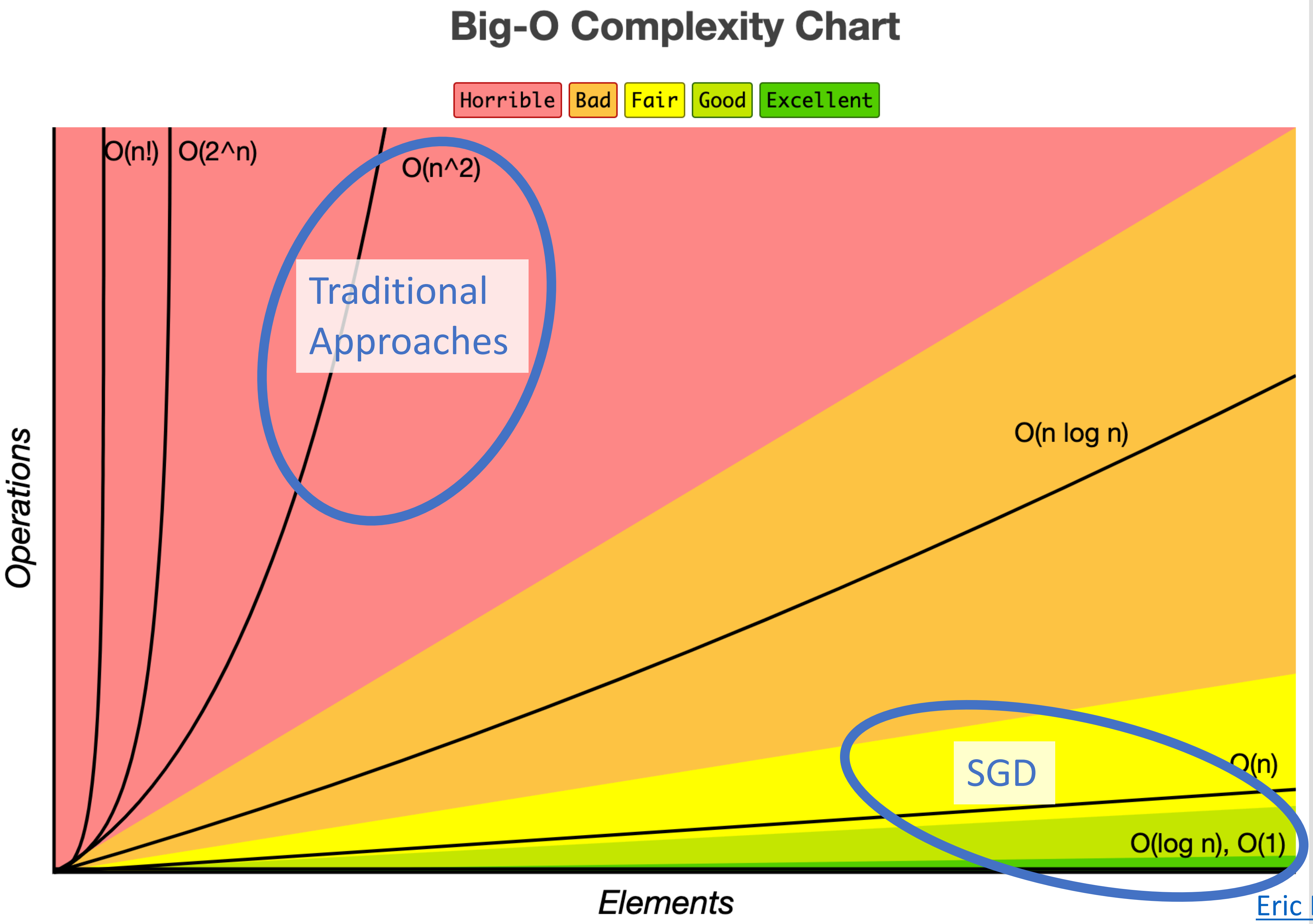
Data Set Size	Time to train
	
	
	

Stochastic Gradient Descent

Data Set Size	Time to train
	
	
	

Stochastic Gradient Descent

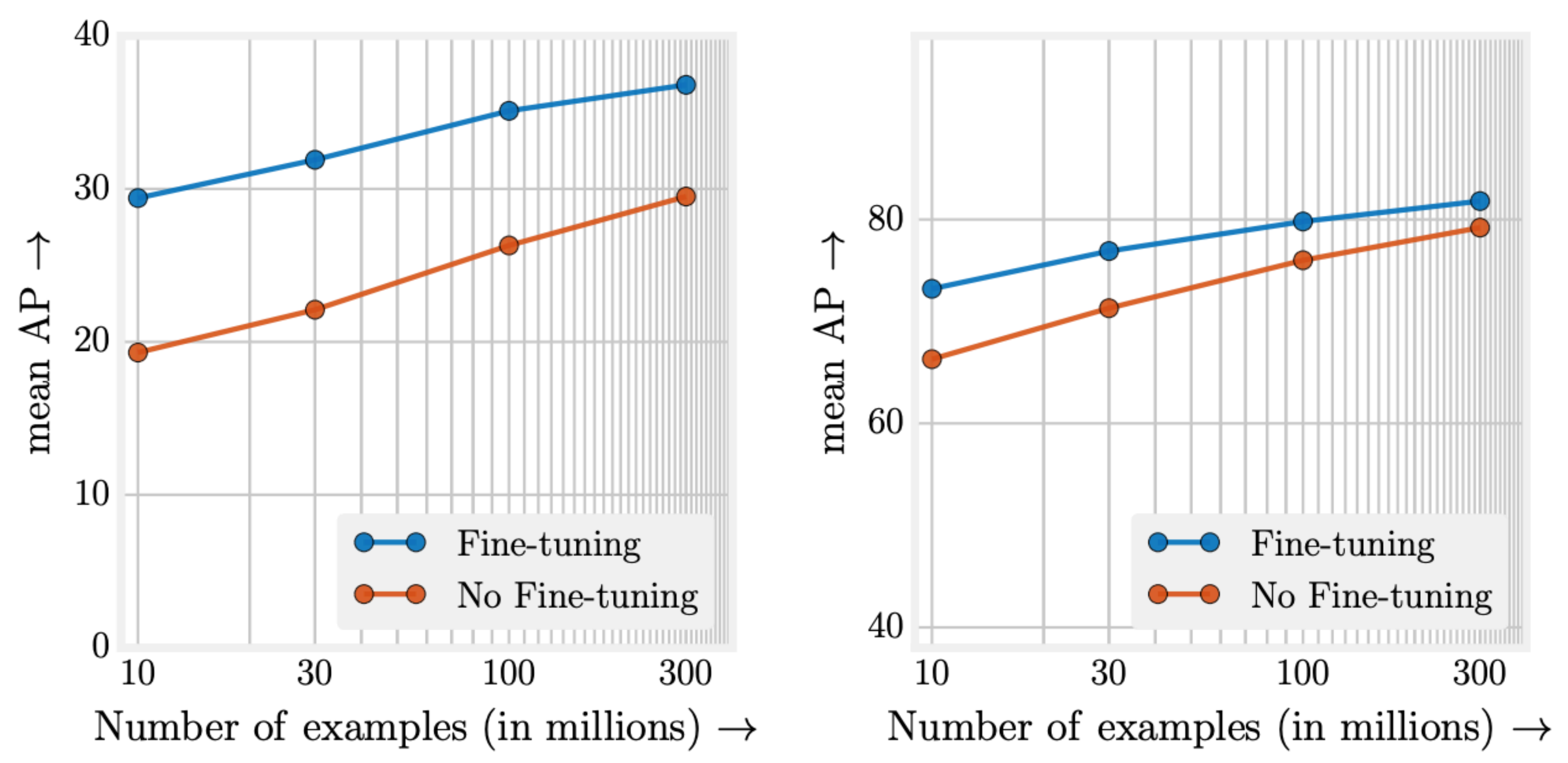
Data Set Size	Time to train
	
	
	

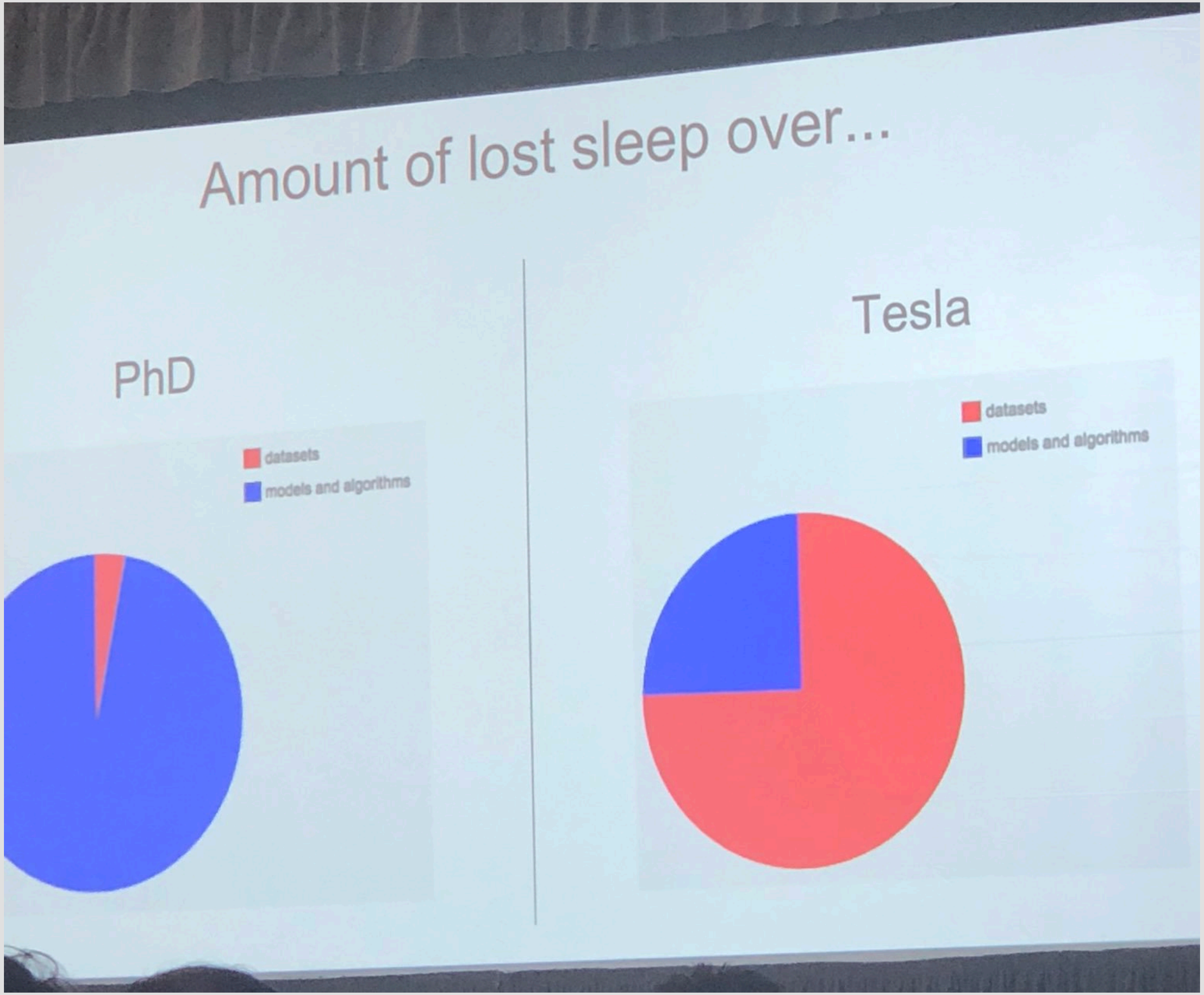


Eric Drowel et al.

Revisiting Unreasonable Effectiveness of Data in Deep Learning Era

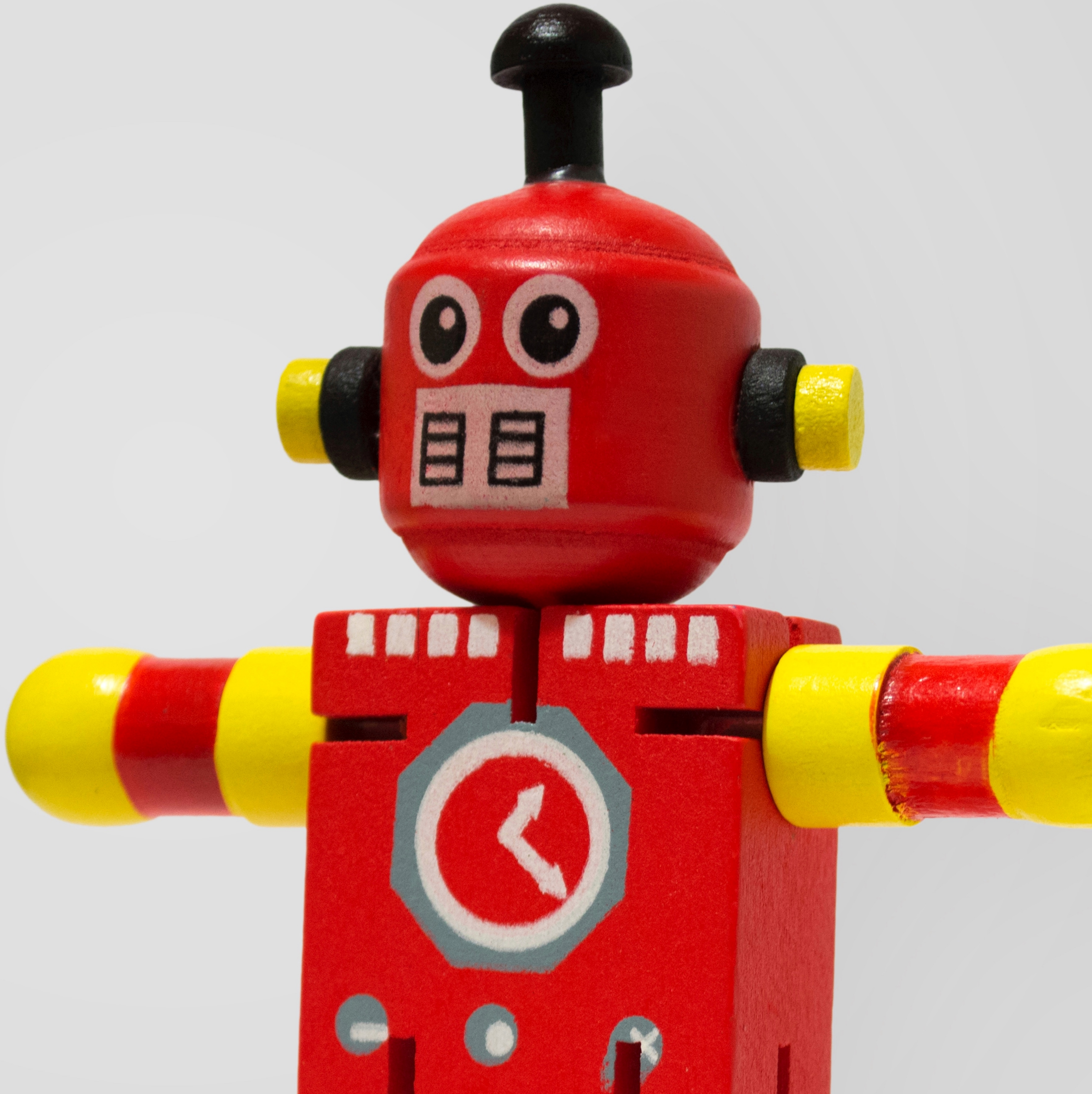
Chen Sun[1], Abhinav Shrivastava[1,2], Saurabh Singh[1], and Abhinav Gupta[1,2]





slide: Andrej Karpathy; photo: *Lisha Li*

Scale is bad



Scale is bad

AI models either

- Replace labor humans would do
- Make new forms of labor possible

Both of these are most profitable at scale!

Scale is bad

- Cathy O'Neil: “the three elements of a WMD: Opacity, **Scale**, and Damage”

Scale is bad

For AI companies bigger means

- Better performing models
- Monopolies on data/content
- Monopsonies on AI developers
- Leverage over regulators

Scale is bad

For AI companies bigger means

- Better performing models
- Monopolies on data/content
- Monopsonies on AI developers
- Leverage over regulators

BAD!

Scale is bad

For AI companies bigger means

- Better performing models
- Monopolies on data/content
- Monopsonies on AI developers
- Leverage over regulators

BAD!

These incentives have always been present. But now there's no speed limit!

What now?

There is a fundamental incentive for AI to scale

This will not be fixed by:

- Technical advances
- A more diverse industry
- Quantifying or removing bias in models/datasets

What now?

AI as an industry **must** be treated as one with inherent risk.

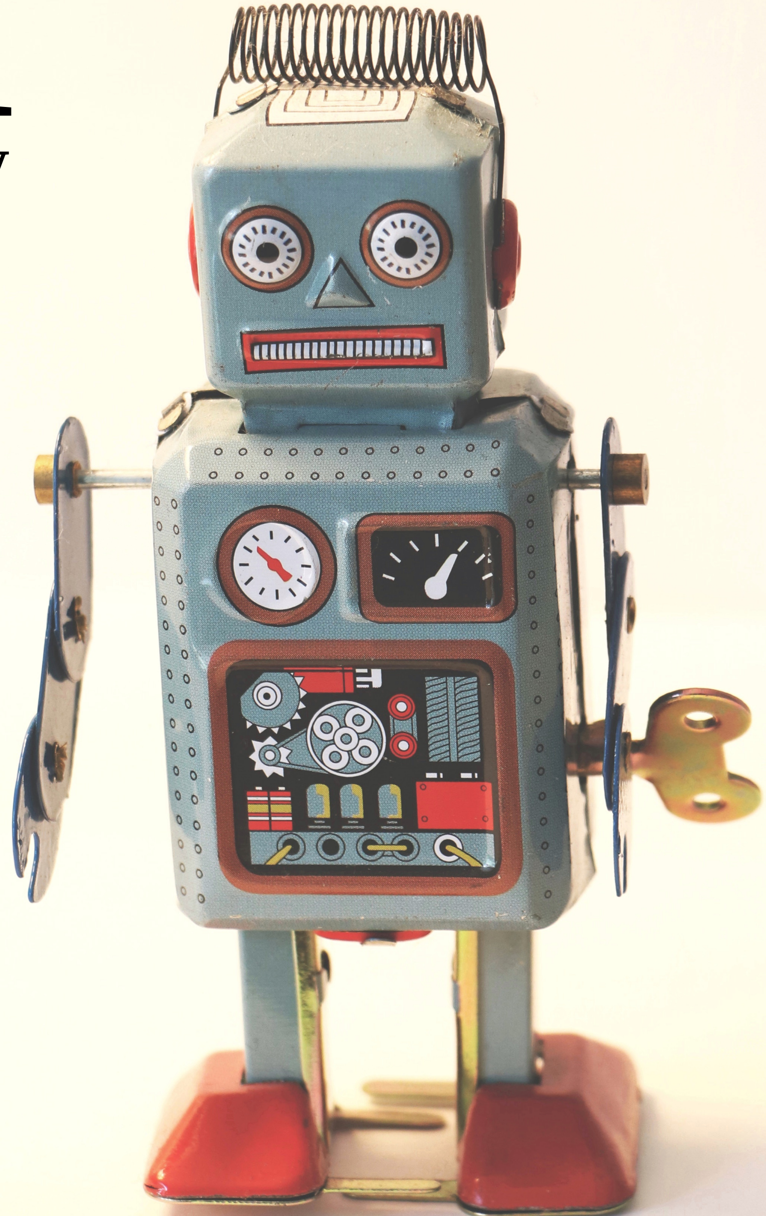
- Regulation with teeth.
- Professional accountability.
- Default presumption of harm.

Examples

- Medicine
- Weapons

AI Is Broken

Sophie Searcy



web: soph.info
github: @artificialsoph
twitter: @artificialsoph