

PROJECT HUB

PROJECT

WEEK 1
MACHINE
LEARNING

ICICS 014

3:30PM | FEB 2 | BIWEEKLY



UBC LAUNCH PAD
SOFTWARE ENGINEERING TEAM



UBC CSSS

UBC Project Hub

- Joint initiative between UBC Launch Pad and CSSS.
- Goal: to create a learning environment at UBC that nurtures a culture of design, innovation, and community amongst the future hackers and entrepreneurs of the tech industry.
- Biweekly meetings with talks & workshops.

Introduction to Machine Learning

Kevin Yap & Sherry Yuan

Today's Agenda

- Talk: **An Overview of Machine Learning** (Kevin)
 - Motivations, successes, and limitations of ML.
- Workshop: **Predicting Credit Card Defaults** (Sherry)
 - Interactive dive into ML with real-world data.

About

- Kevin Yap (@iKevinY)
- 5th Year Honours Computer Science
- Experimented with NLP at Axiom Zen
- Built neural network for nwHacks 2018 project
- Took CPSC 340 two years ago
- Finishing up thesis on machine learning & StarCraft II



Goals for this Talk

- Discuss motivations for machine learning.
- Short overview of the history of the field.
- Briefly touch on various techniques.
- Introduce jargon and other terminology.
- Show that machine learning is approachable!

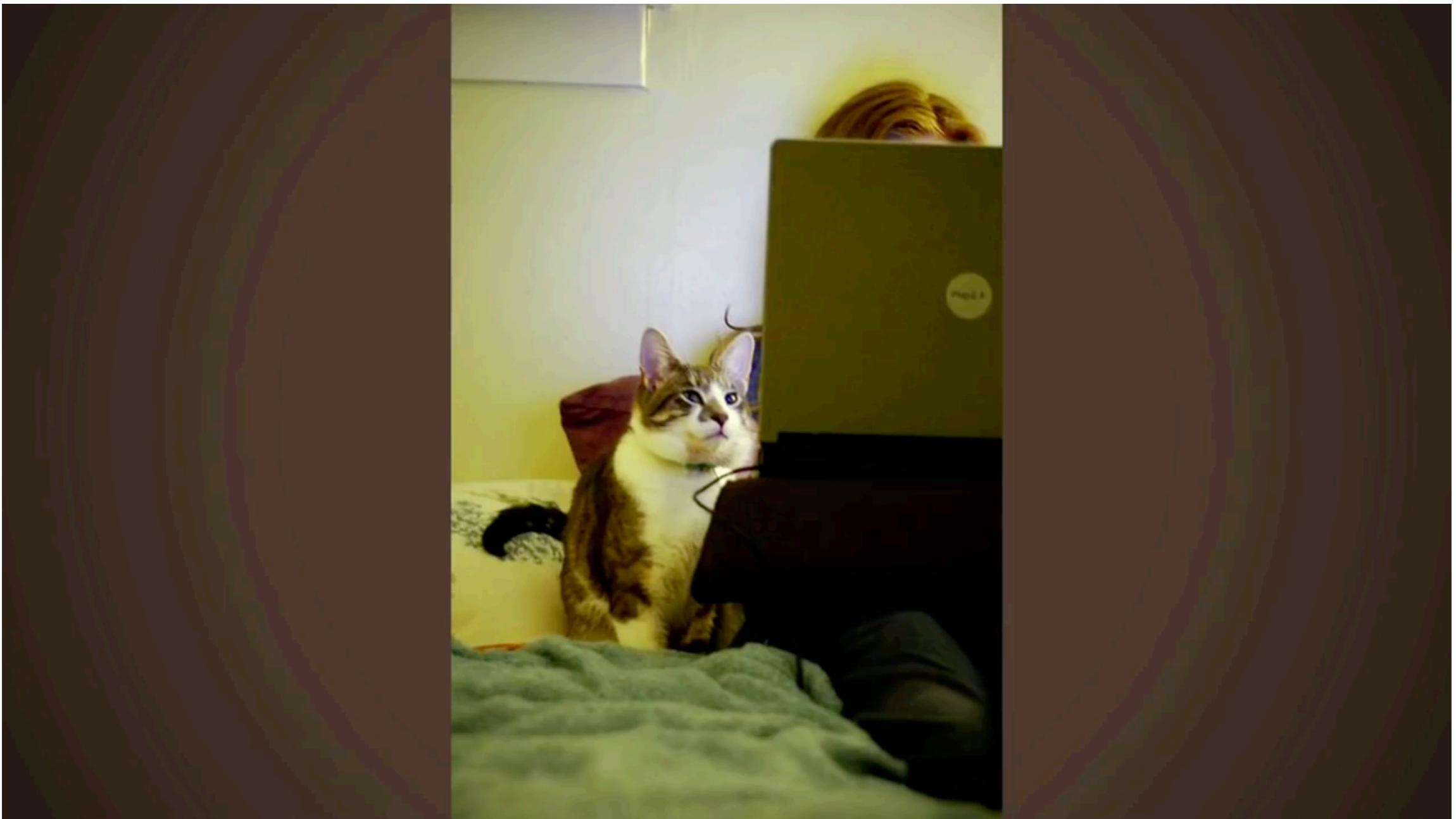
What is "machine learning"?

- Machine learning (ML) is the study using **algorithms** and **statistical models** to allow computer systems to effectively perform a specific task **without using explicit instructions**, relying on **models** and **inference** instead.
- Subfield of AI (artificial intelligence).

Applications of Machine Learning

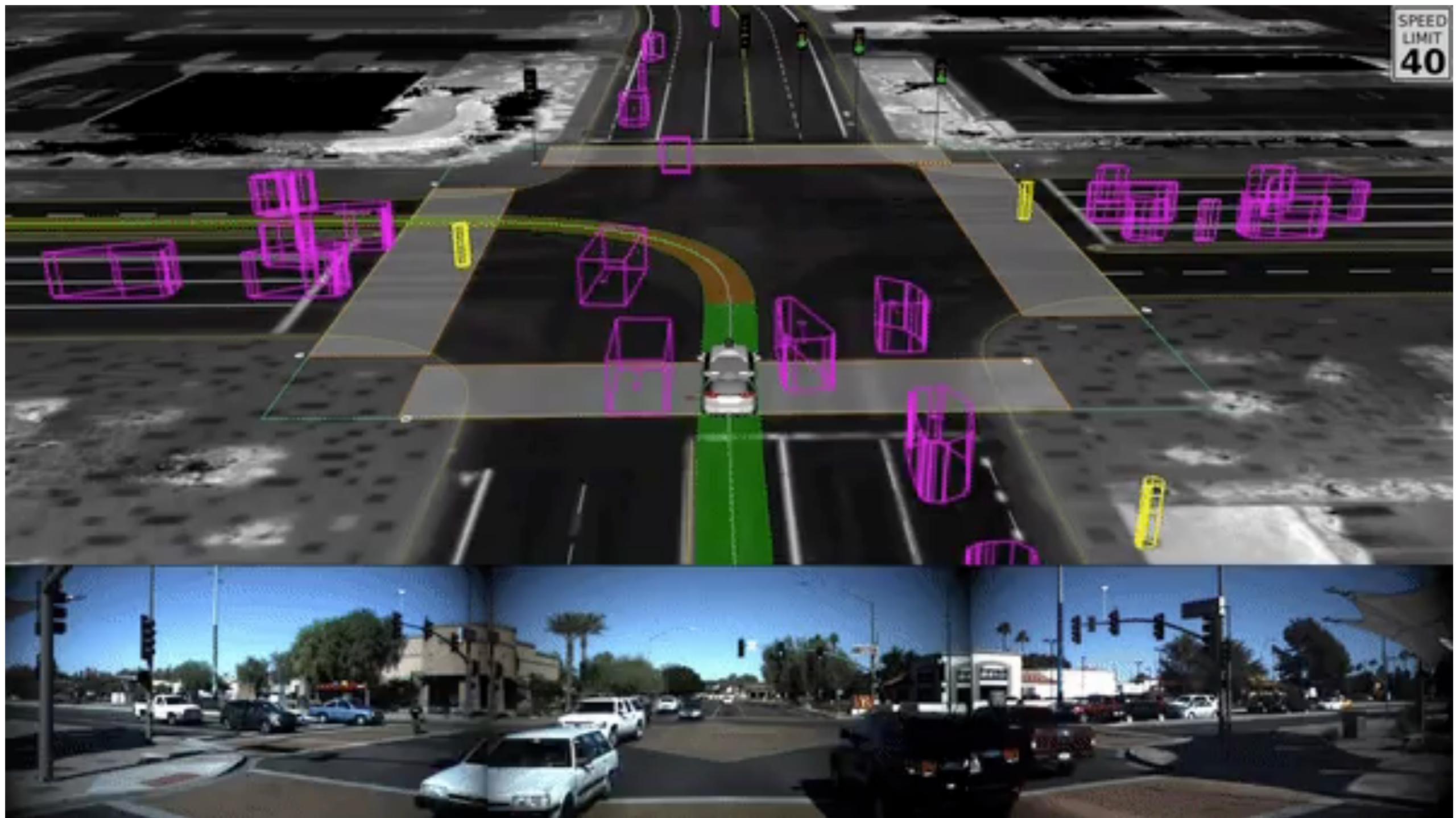
- Artificial Intelligence (game agents)
- Computer Vision (self-driving cars)
- Natural Language Processing (machine translation)
- Recommendation Systems (Netflix/Amazon suggestions)

Computer Vision



TED Talk: How we teach computers to understand pictures (Fei Fei Li)

Waymo's Self-Driving Car



Chihuahua or Muffin



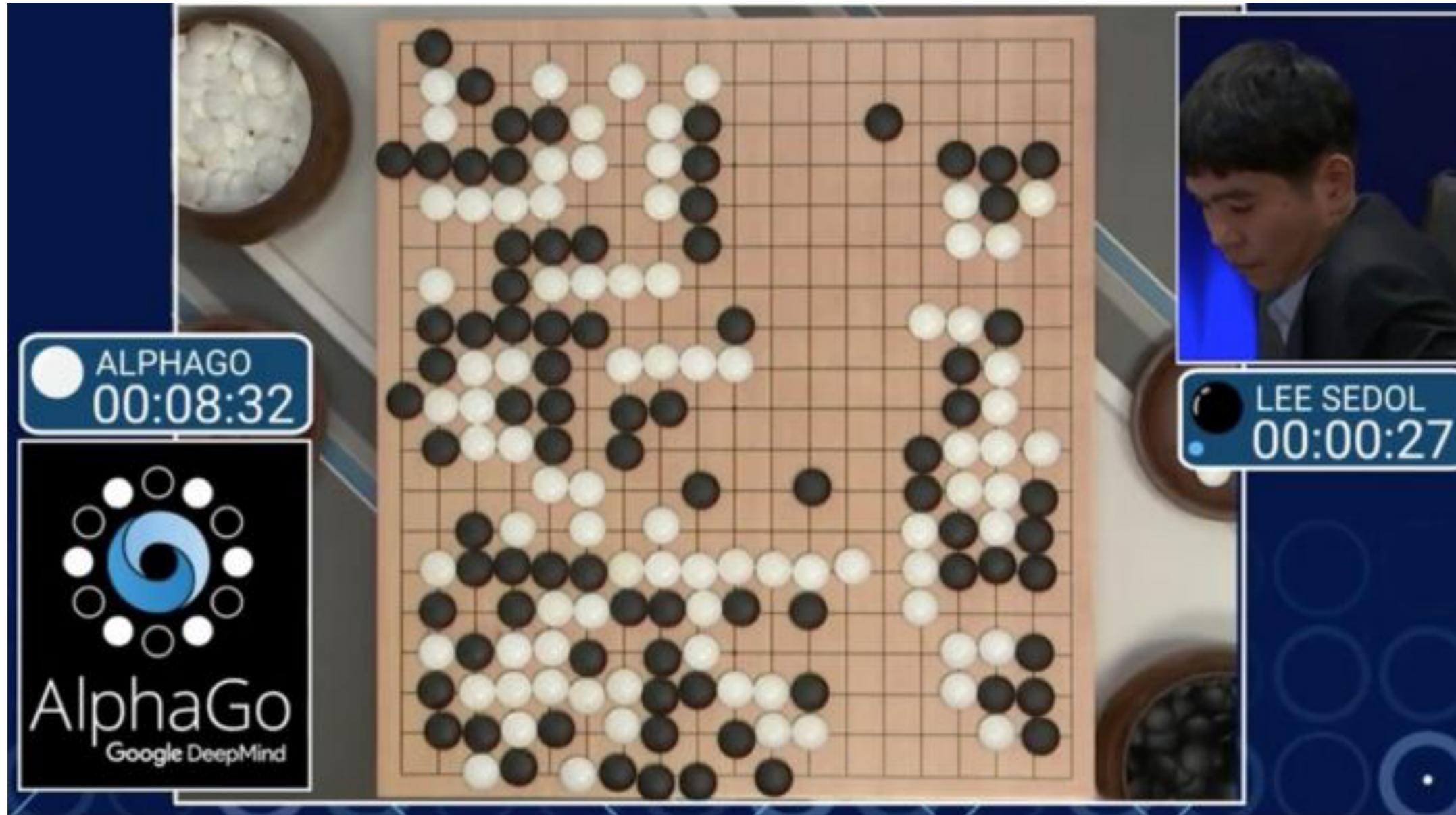
WHEN A USER TAKES A PHOTO,
THE APP SHOULD CHECK WHETHER
THEY'RE IN A NATIONAL PARK...

SURE, EASY GIS LOOKUP.
GIMME A FEW HOURS.



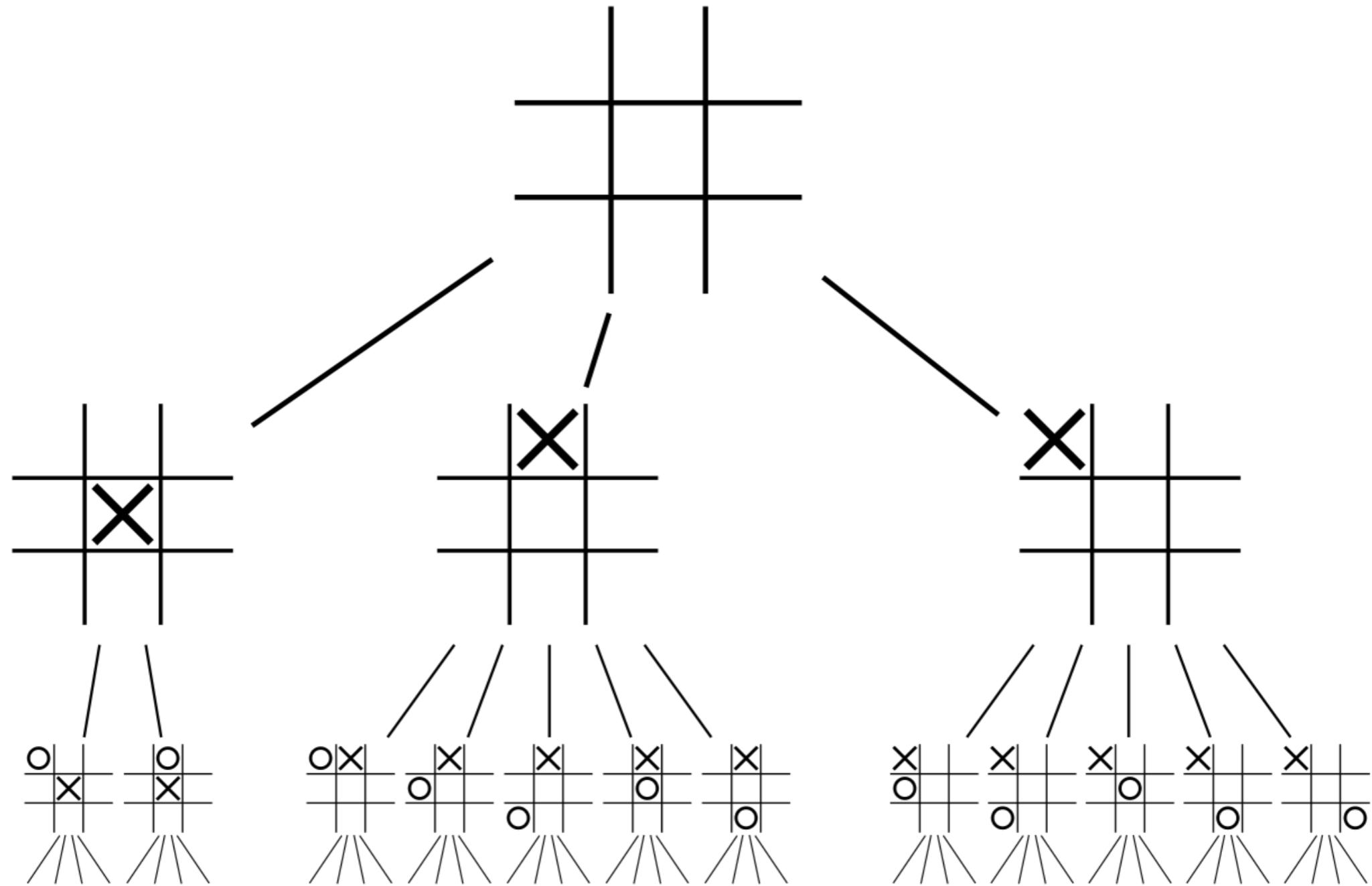


1997: Deep Blue beats Garry Kasparov in chess



2016: AlphaGo beats Lee Se-dol at Go

Solving Chess vs. Go



Solving Chess vs. Go

	Board Size	Pieces	Branching Factor	Space
Tic-Tac-Toe	3×3	9	4	512
Checkers	8×8	24	2.8	$5 \cdot 10^{20}$
Chess	8×8	24	35	10^{120}
Go	19×19	361	250	10^{360}

Solving Chess vs. Go

- Deep Blue: rule-based system, basic tree search
- AlphaGo: tree search + neural network

The Big Data Boom

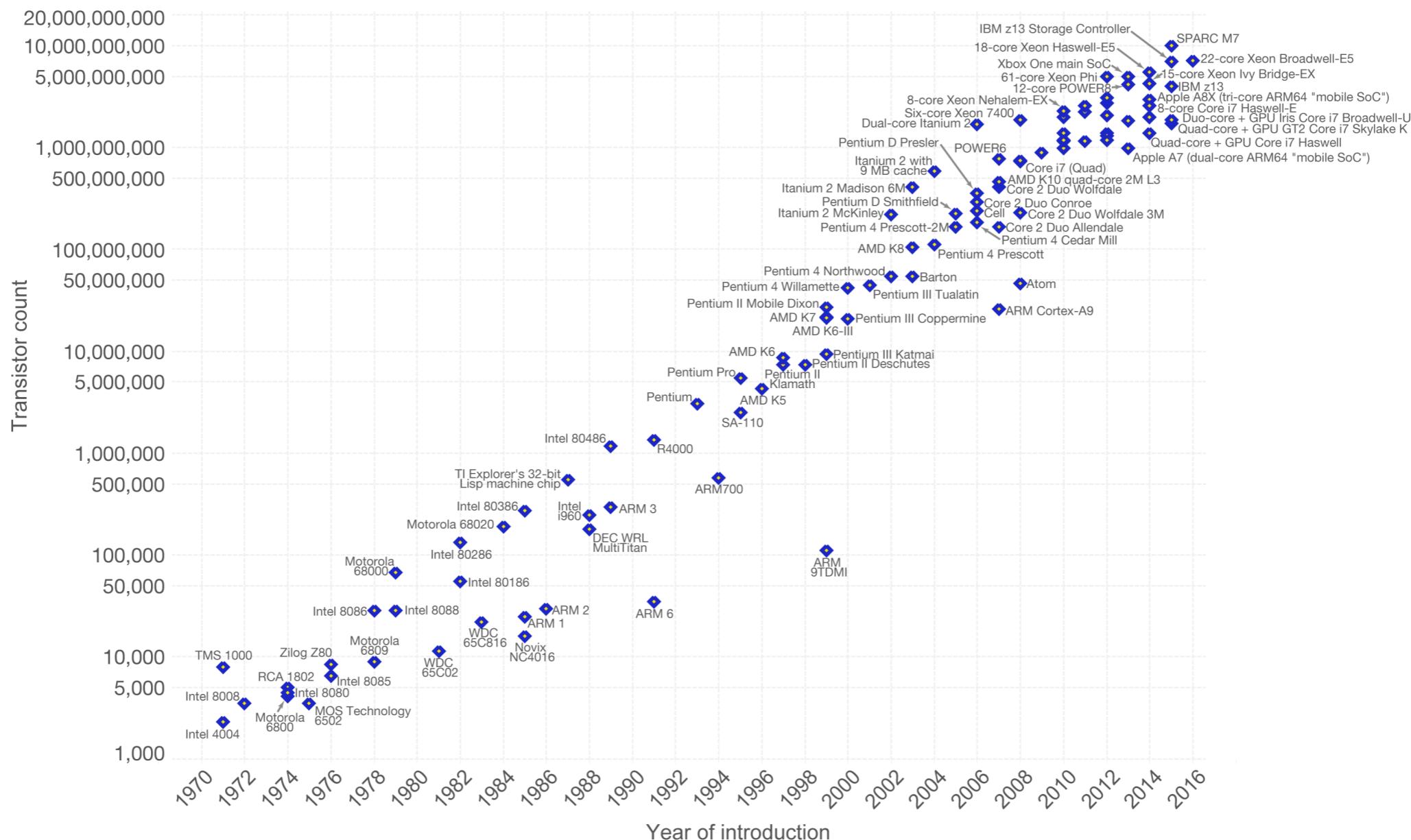


Moore's Law – The number of transistors on integrated circuit chips (1971-2016)

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years.

Our World in Data

This advancement is important as other aspects of technological progress – such as processing speed or the price of electronic products – are strongly linked to Moore's law.



Data source: Wikipedia (https://en.wikipedia.org/wiki/Transistor_count)

The data visualization is available at OurWorldInData.org. There you find more visualizations and research on this topic.

Licensed under CC-BY-SA by the author Max Roser.

Machine Learning Basics

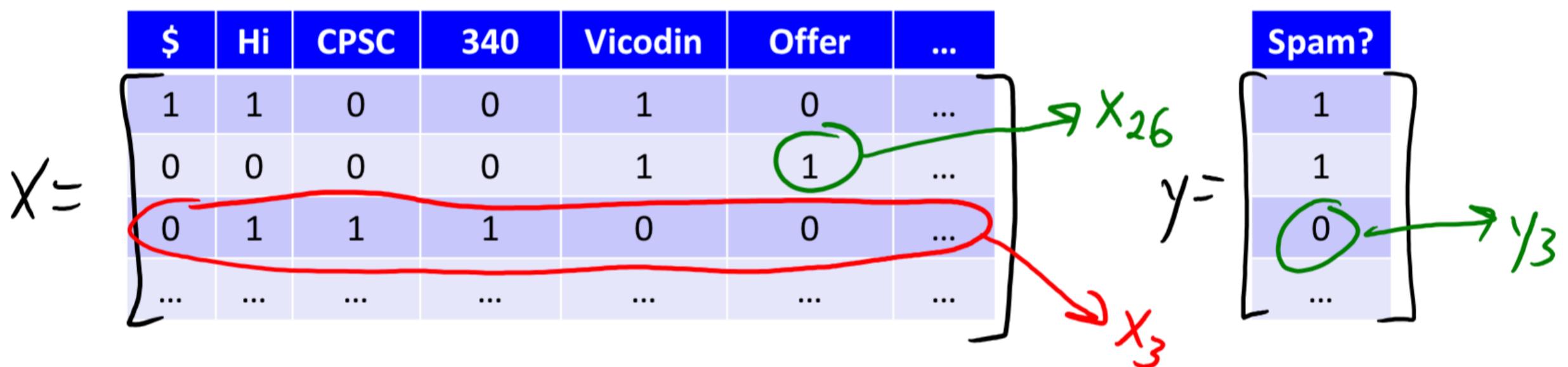
ML in Practice

- Python
- NumPy to interact with data (matrices)
- Uses C bindings under the hood
- We choose "hyperparameters"
- Models learn "parameters"

Predicting y from X

$$y = f(x)$$

↑
output classification input
 function



Supervised Learning



apple

pear

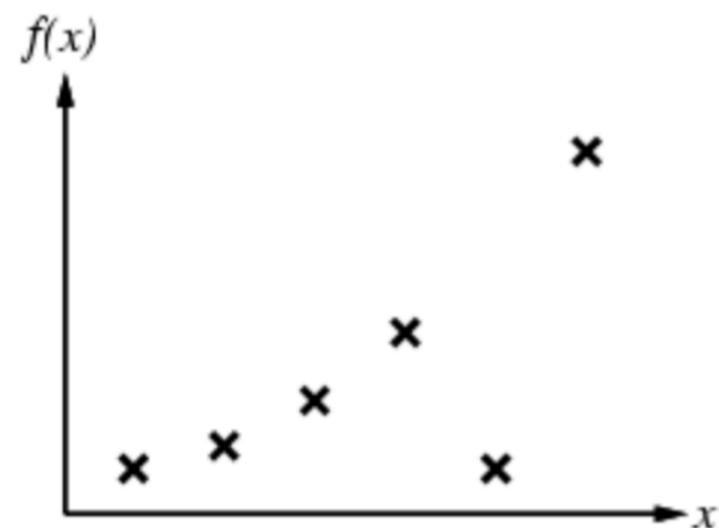
tomato

cow

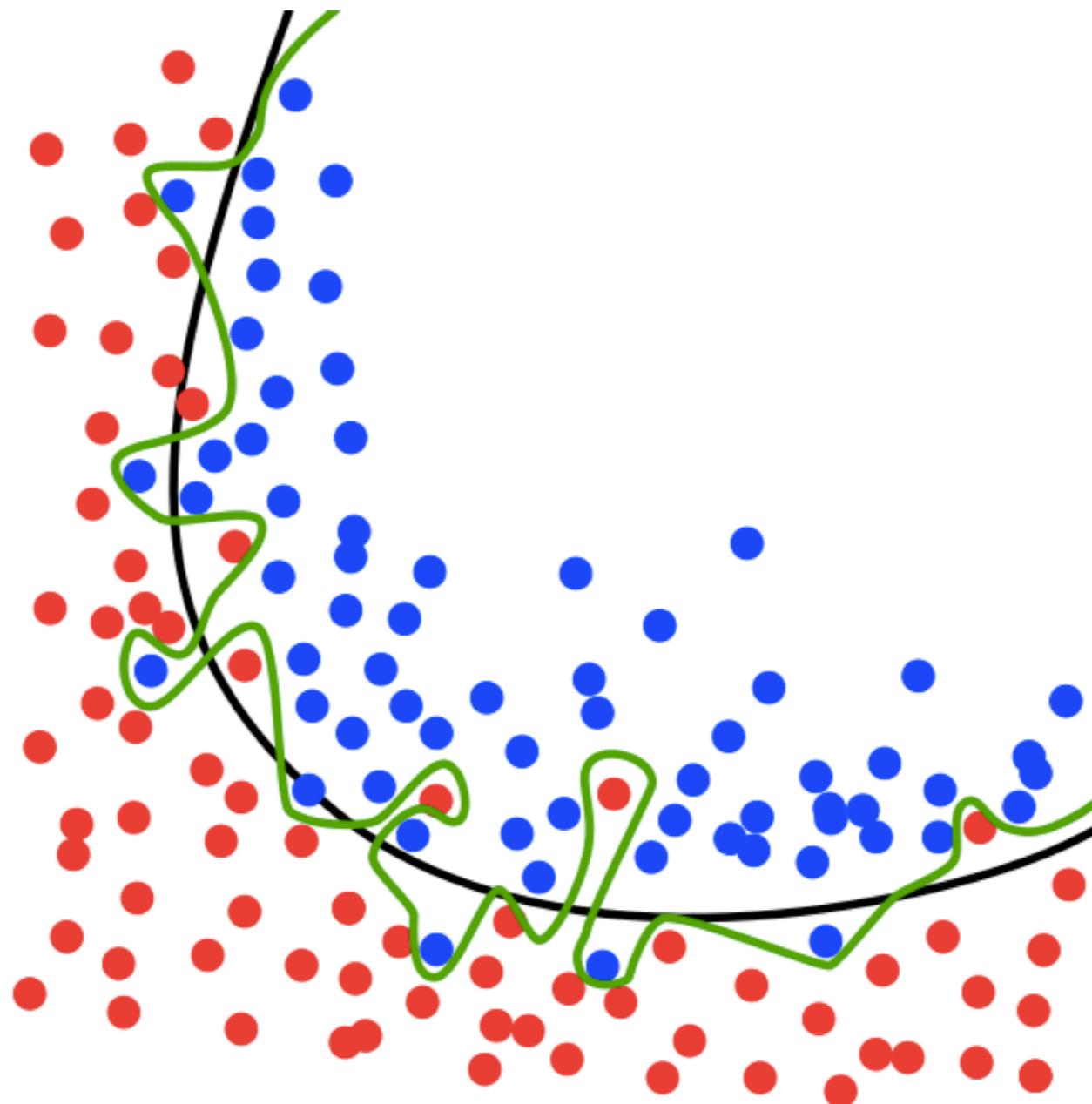
dog

horse

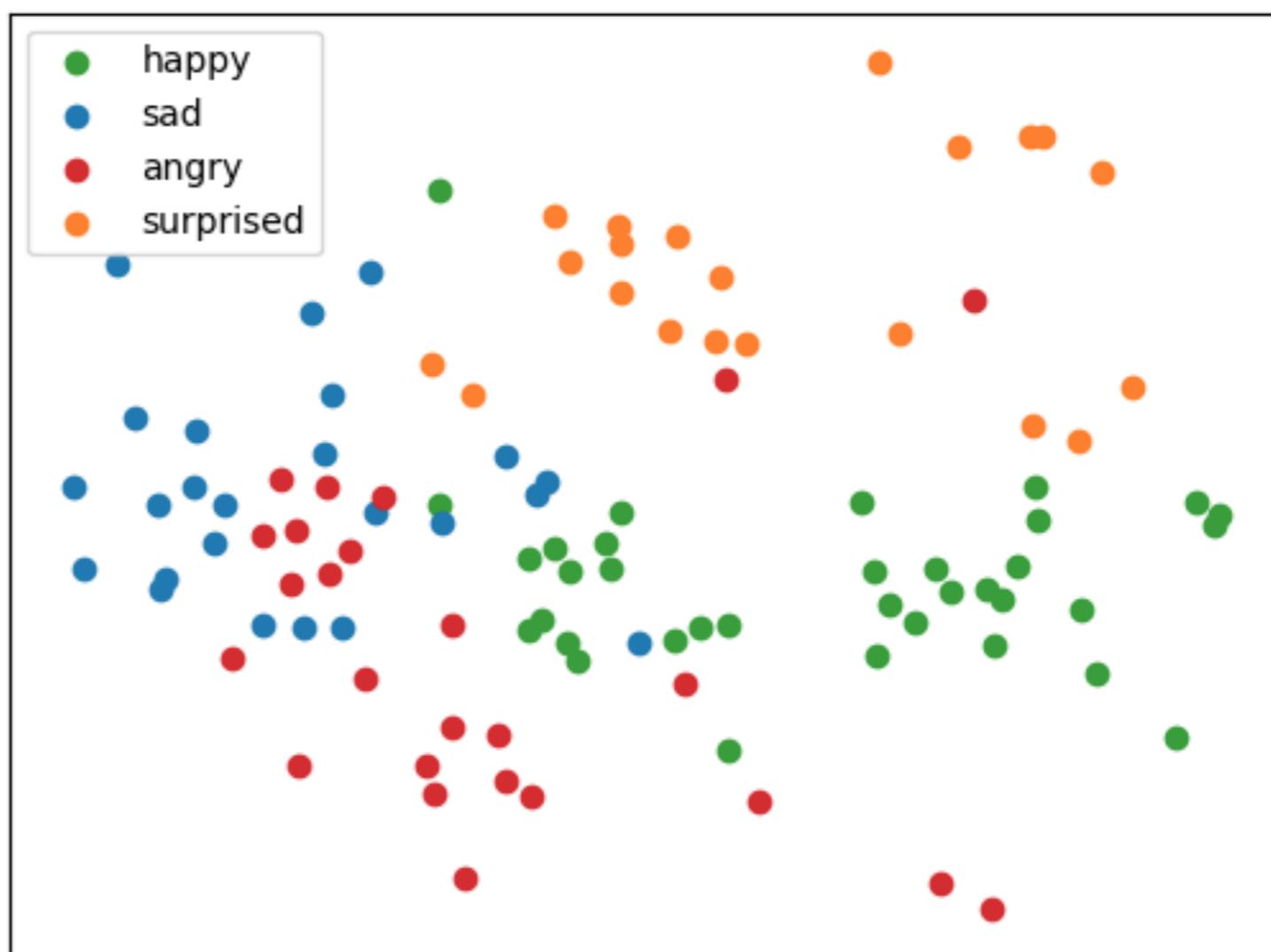
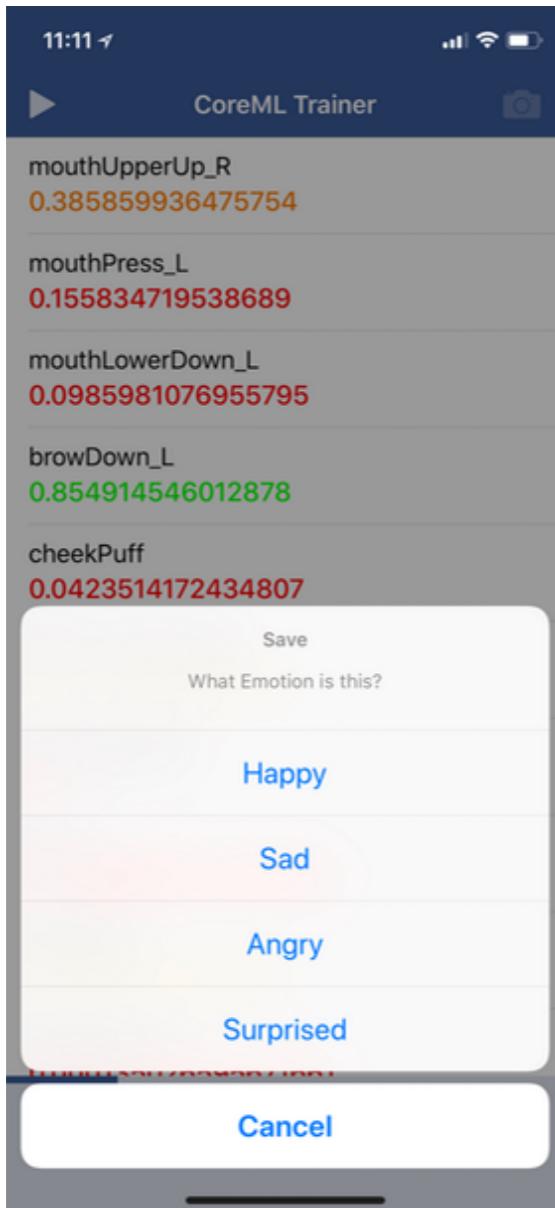
Regression



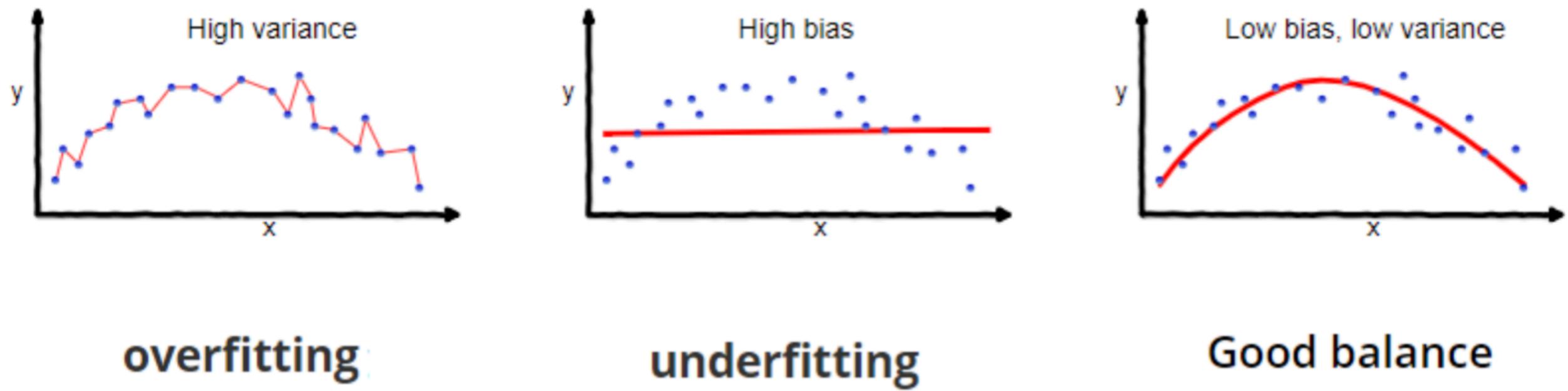
Classification



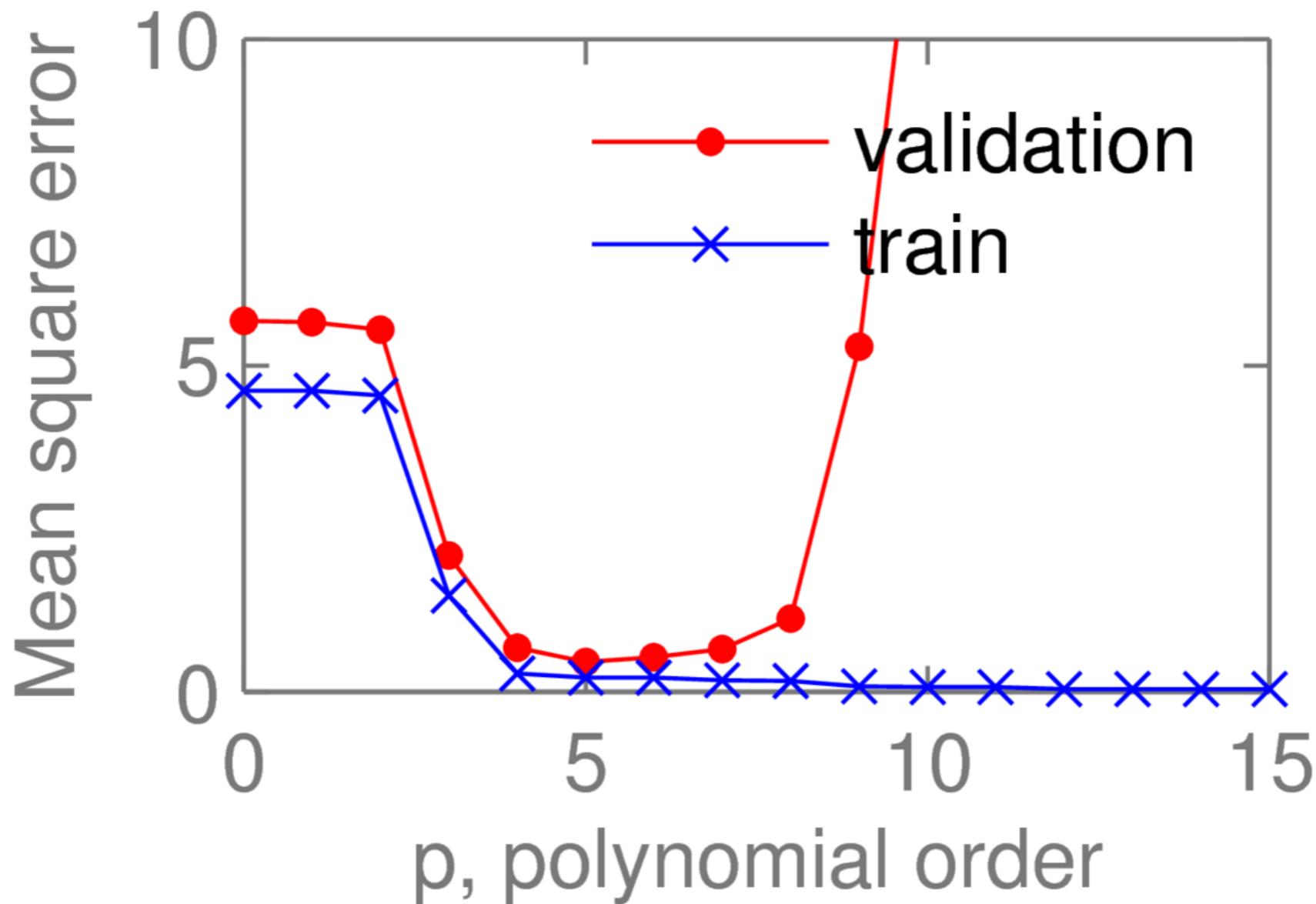
Classification



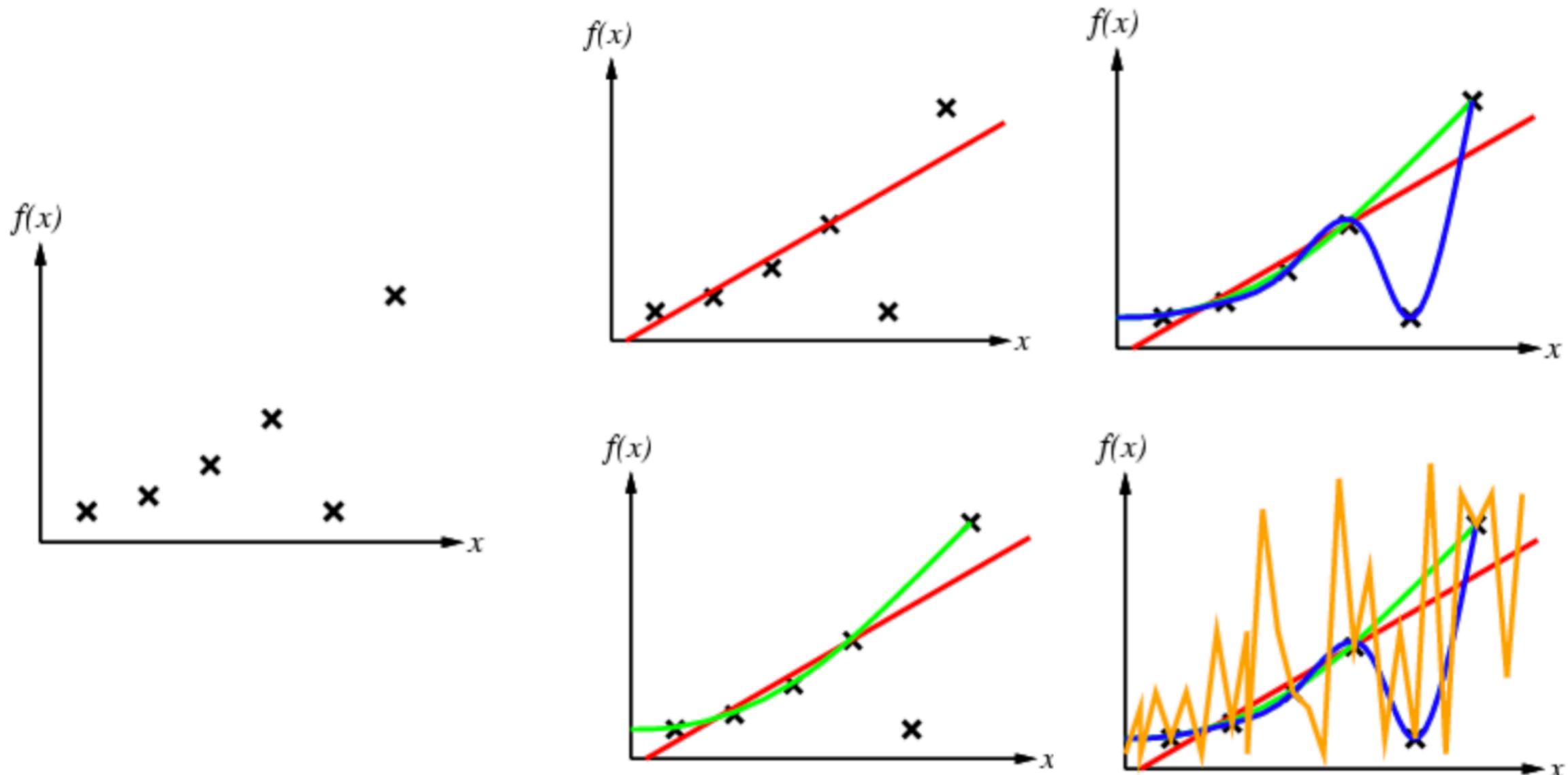
Dangers of Overfitting



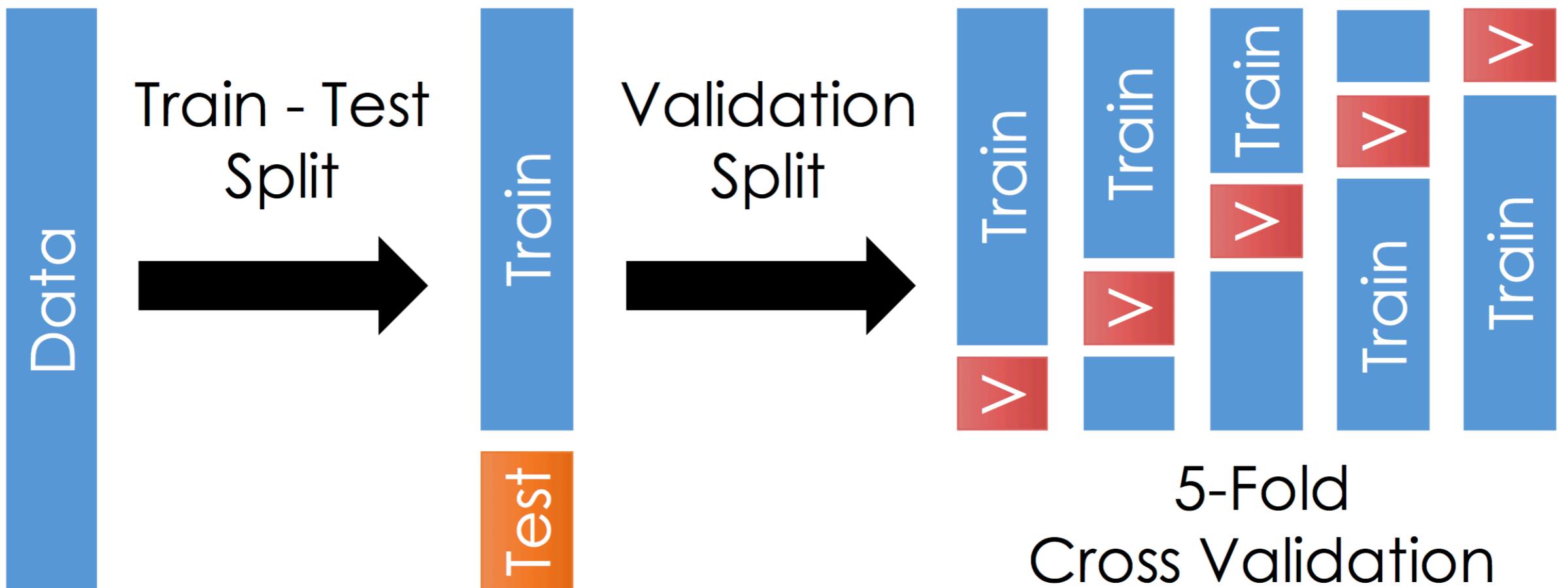
Dangers of Overfitting



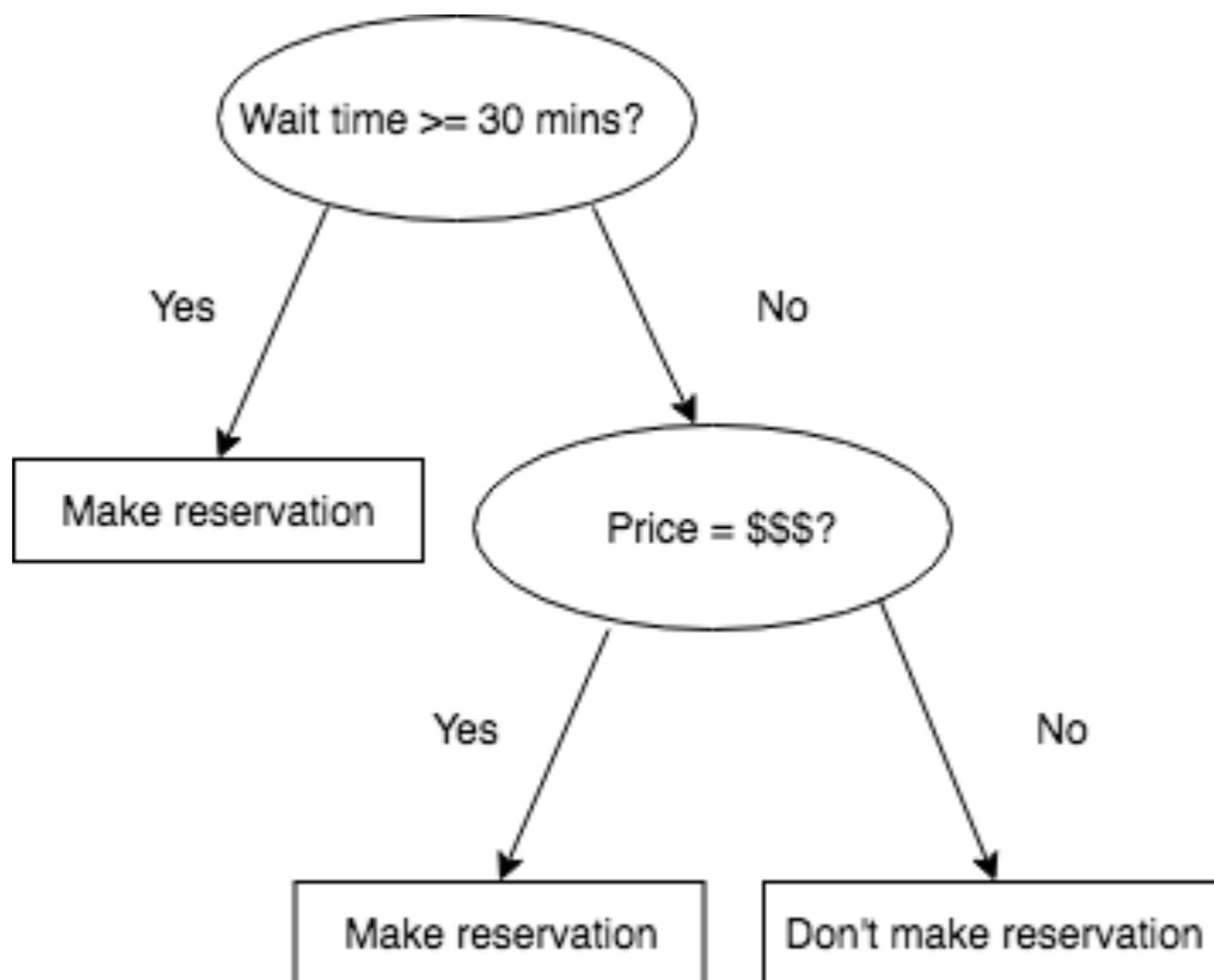
Dangers of Overfitting



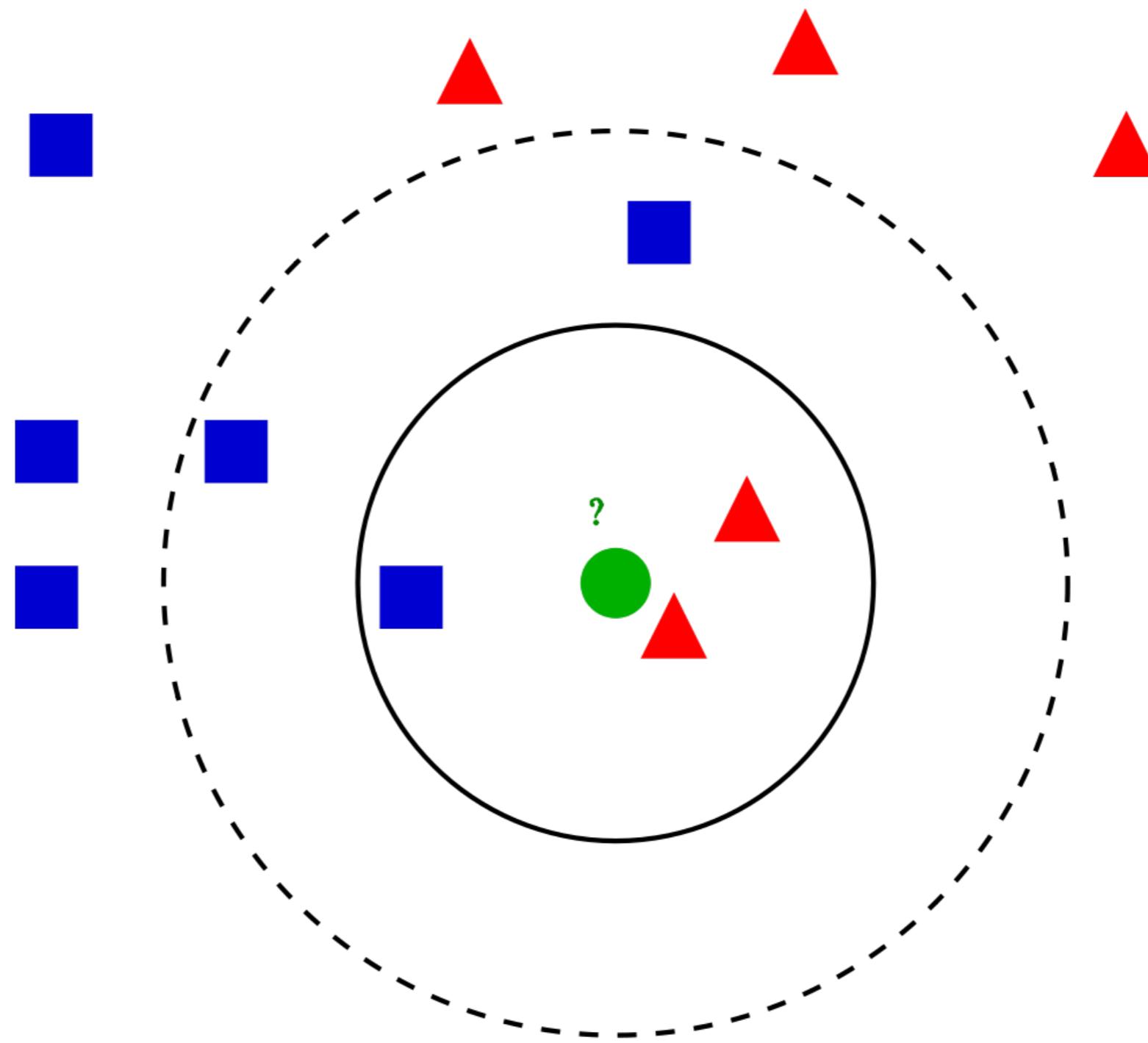
Training / Test / Validation



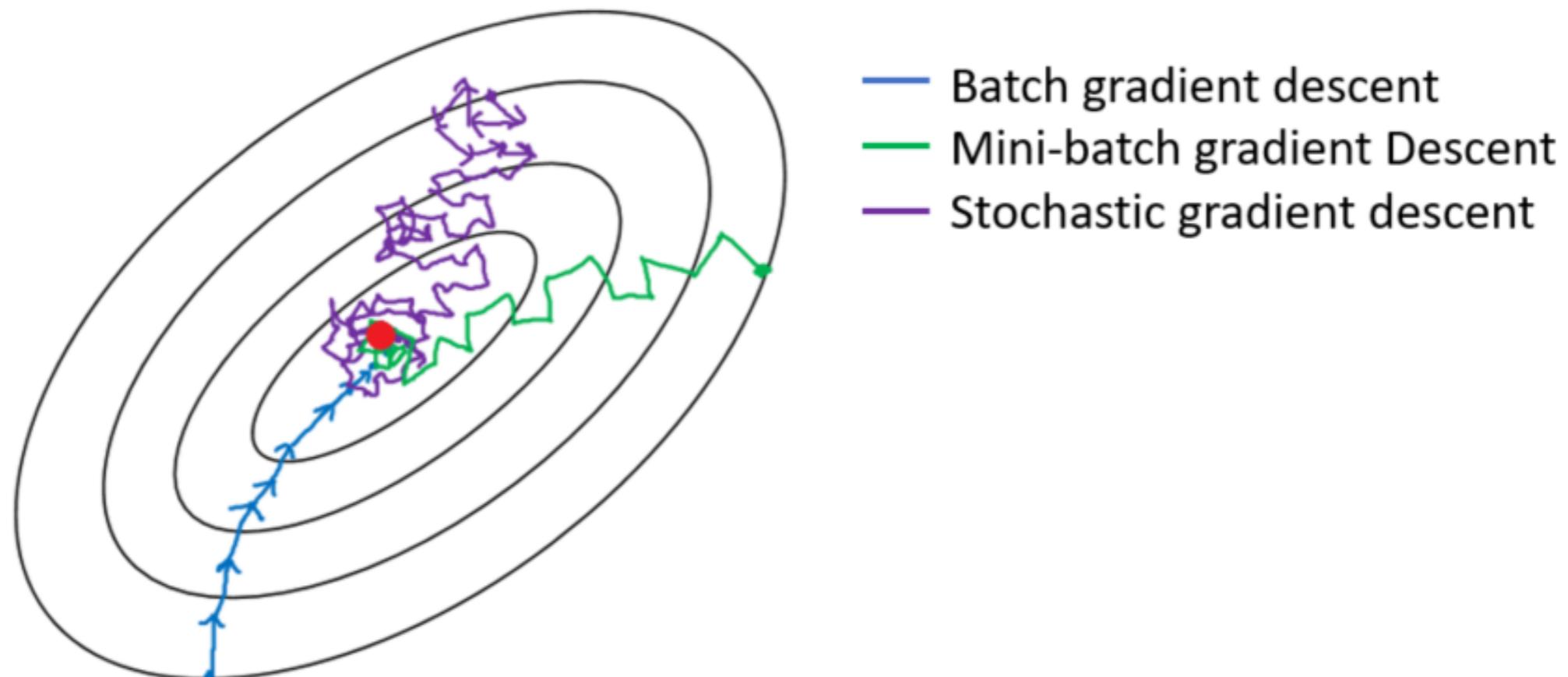
Decision Trees (Boolean Logic)



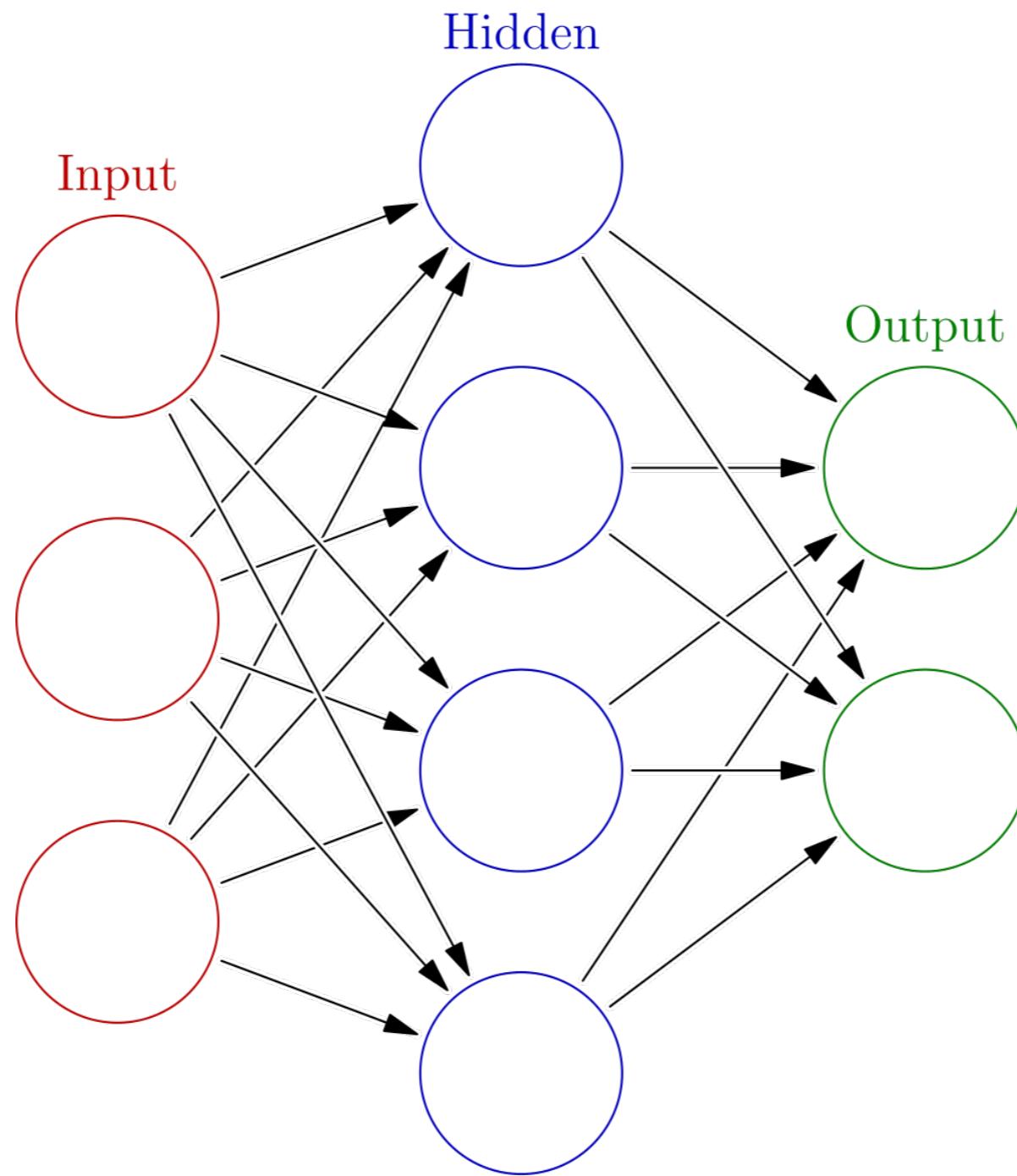
k-Nearest Neighbours



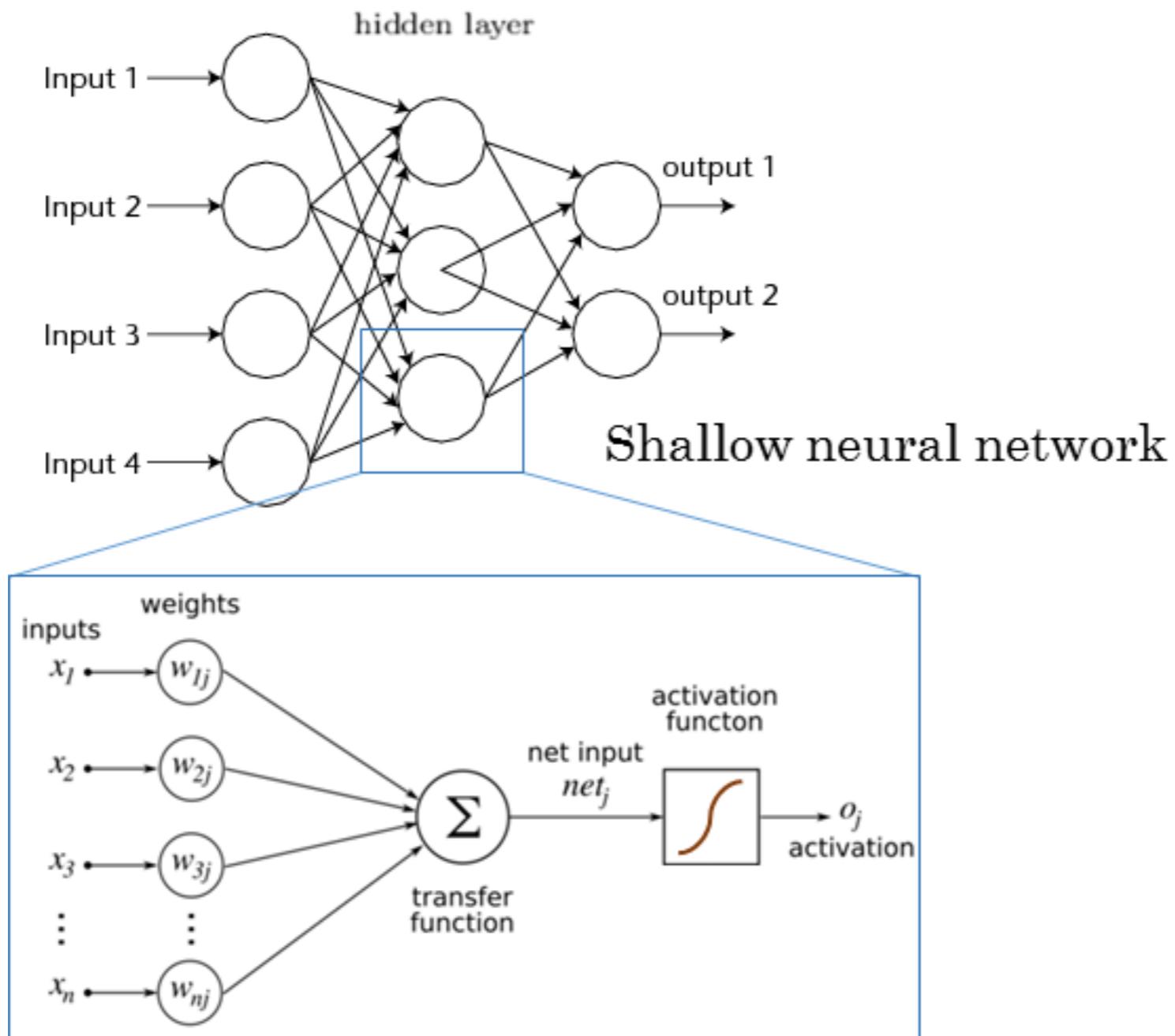
Stochastic Gradient Descent



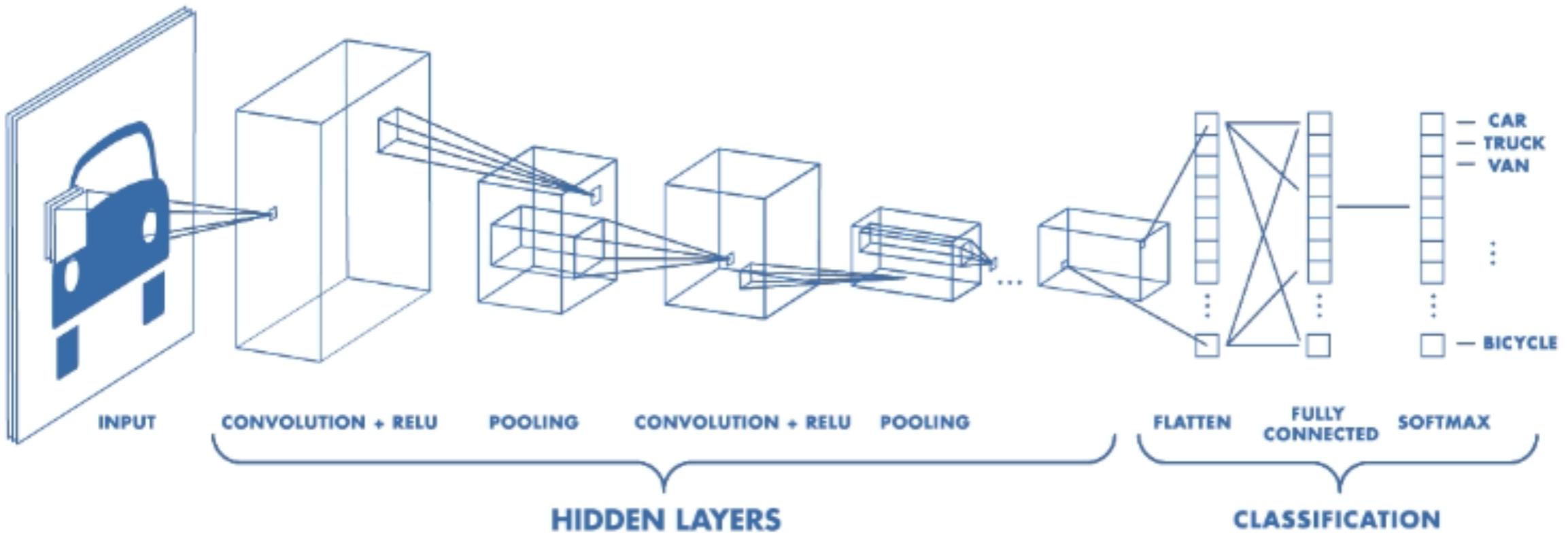
Neural Networks



Neural Networks



Convolutional Neural Network



Resources

- 3Blue1Brown on neural networks (<http://3b1b.co/neural-networks>)
- Welch Labs on computer vision + ML (<https://youtu.be/i8D90DkCLhI>)
- Google's crash course (<https://developers.google.com/machine-learning/crash-course/ml-intro>)
- CPSC 340 (<https://ubc-cs.github.io/cpsc340/>)

Questions?

Workshop Time!

- Colaboratory (Jupyter Notebooks + Google Docs)
- https://colab.research.google.com/drive/1Rk9pF_ZjhVHLGDCJZAN47xayqatkrsKW