Intro to Neural Nets

Course Logistics and Introduction

Today’s Agenda

1. COURSE LOGISTICS

- Website, schedule, grading and evaluation criteria.
- Course textbook, lecture format, etc.

2. INTERESTING USE CASES

- Frivolous, academic, and practically useful.
- A recent failure, and societal concerns.

3. QUICK INTRODUCTION

- What is a neural network?
- How does it work?

About Me

accenture

MSI at the ARF

NSF

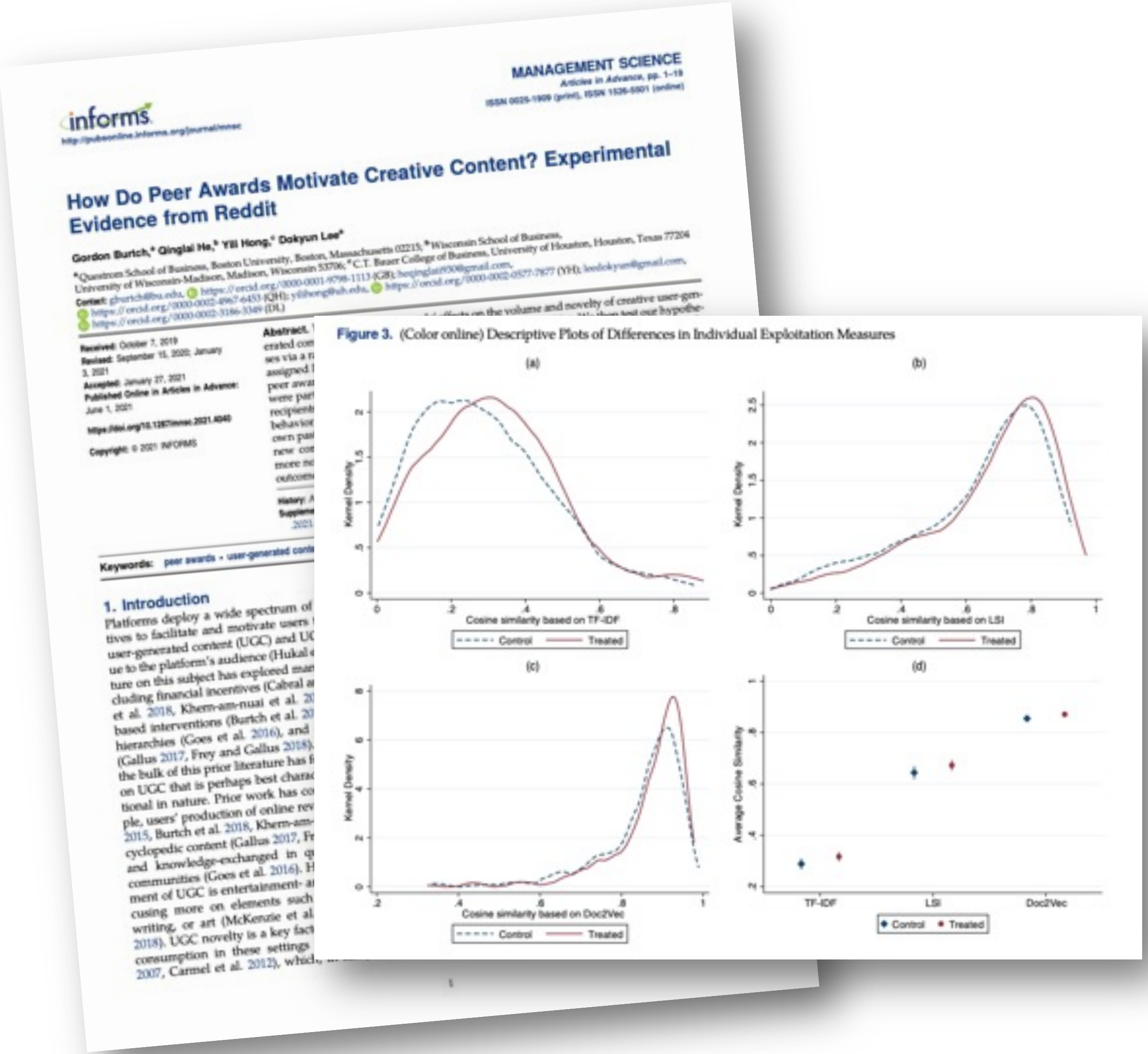
FFG Forschung wirkt.

Adobe

3M

Microsoft

My Research



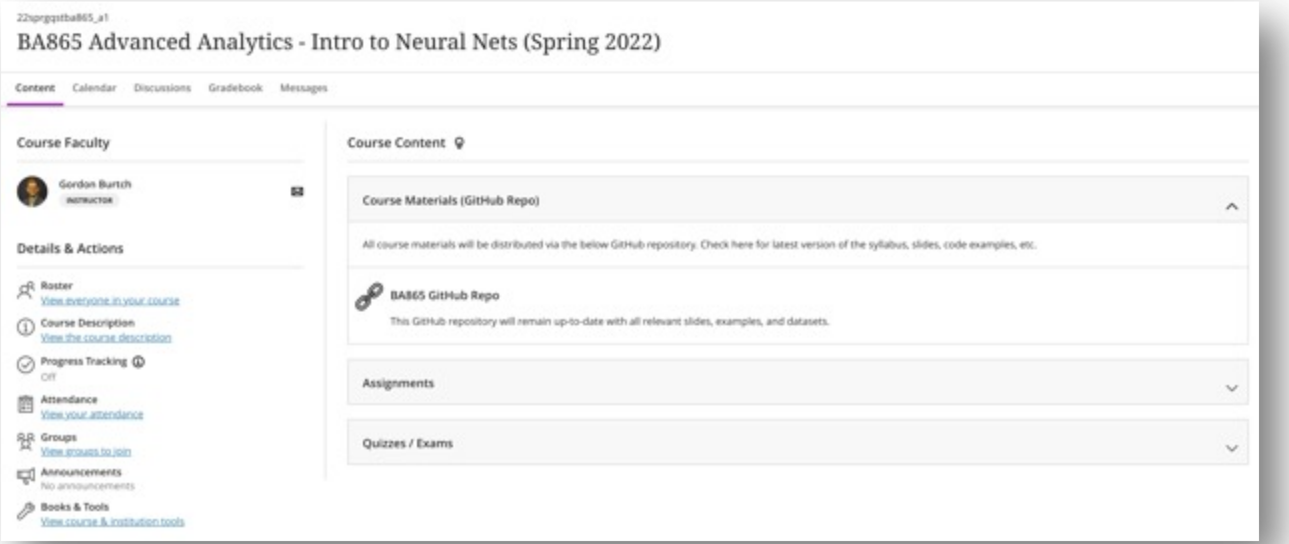
MANAGEMENT SCIENCE
Articles in Advance, pp. 1–19
informs
How Do Peer Awards Motivate Creative Content? Experimental Evidence from Reddit
Gordon Burtch, Qinglai He, Yili Hong, Dokyun Lee
Keywords: peer awards • user-generated content
1. Introduction
Figure 3. (Color online) Descriptive Plots of Differences in Individual Exploitation Measures
(a)
(b)
(c)
(d)
Kernel Density
Cosine similarity based on TF-IDF
Cosine similarity based on LSI
Cosine similarity based on Doc2Vec
Average Cosine Similarity
TF-IDF
LSI
Doc2Vec
Control
Treated

Achieving Reliable Causal Inference with Data-Mined Variables: A Random Forest Approach to the Measurement Error Problem
Mochen Yang, Edward McFowland III, Gordon Burtch, Gediminas Adomavicius
University of Minnesota, Carlson School of Management
March 13, 2020
Abstract
1 Introduction

Course Materials

COURSE WEBSITE

- The course website is on Blackboard – I assume you have looked at it by this point (if you aren't enrolled in the site, let me know)
- You will submit all assignments and receive relevant course announcements via this site.
- I will post lecture materials and in-class exercises / examples in the GitHub Repository that is linked on Blackboard.



GOOGLE COLAB

- All homework and exercises in this course will be implemented in Python. You are welcome to employ a local instance of Python, but for simplicity's sake I would recommend you work in Google Colab, because we (the TA and I) will not be able to provide technical support. We will rely heavily on Jupyter Notebooks.

Grading and Evaluation

25%

25%

HOMEWORK

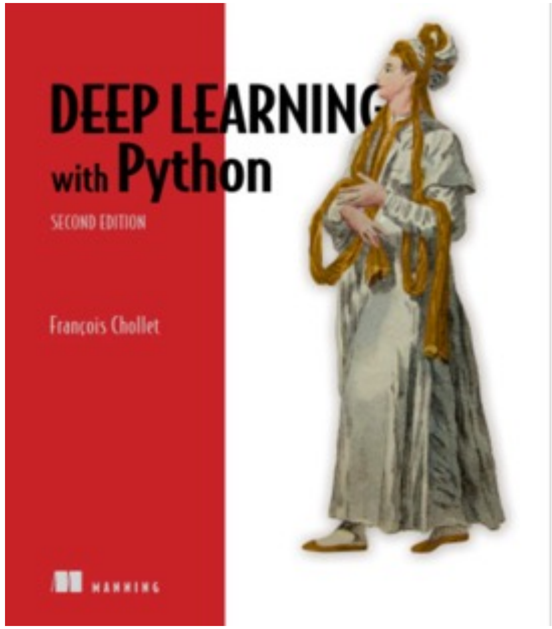
- Two (2) assignments worth 25% of your final grade, each, to be completed in groups of no more than 2 students (indicate group members as part of your submission); you can also complete these assignments by yourself, if you prefer.
- Each homework problem set is based on preceding lab / course material. Due before the start of class on the indicated submission date – **submit via Blackboard only!**
- Per the syllabus, late submissions will not be accepted.

50%

FINAL PROJECT

- The final project will be presented during the last regularly scheduled classroom session.
- You will work in groups of 2-3 (your choice) to implement a neural network-based predictive model that addresses a practical problem of interest to you! You will source your own dataset, motivate the prediction problem (explain why it's practically relevant), and then implement the predictive model. Your goal should be to make this project useful for you!
- The deliverables will include a project proposal due in Week 2 (5%), a mid-point check-in meeting with me to ensure you are on track (5%)

Course Textbook



Chollet, François. (2021). *Deep Learning with Python (2nd Edition)*. Manning Publications Co. **ISBN-13: 978-1617296864.**
https://www.manning.com/books/deep-learning-with-python-second-edition

Required Software

SOFTWARE CONFIGURATION

- You can access Google Colab at https://colab.research.google.com. You probably want to use your BU Google account credentials.
- If you wish, you may run Python / Jupyter notebooks locally, installing the relevant packages, but please note that we (the TA and I) cannot provide any technical support.

Why Keras?

Figure 1.12 Machine learning tools used by top teams on Kaggle

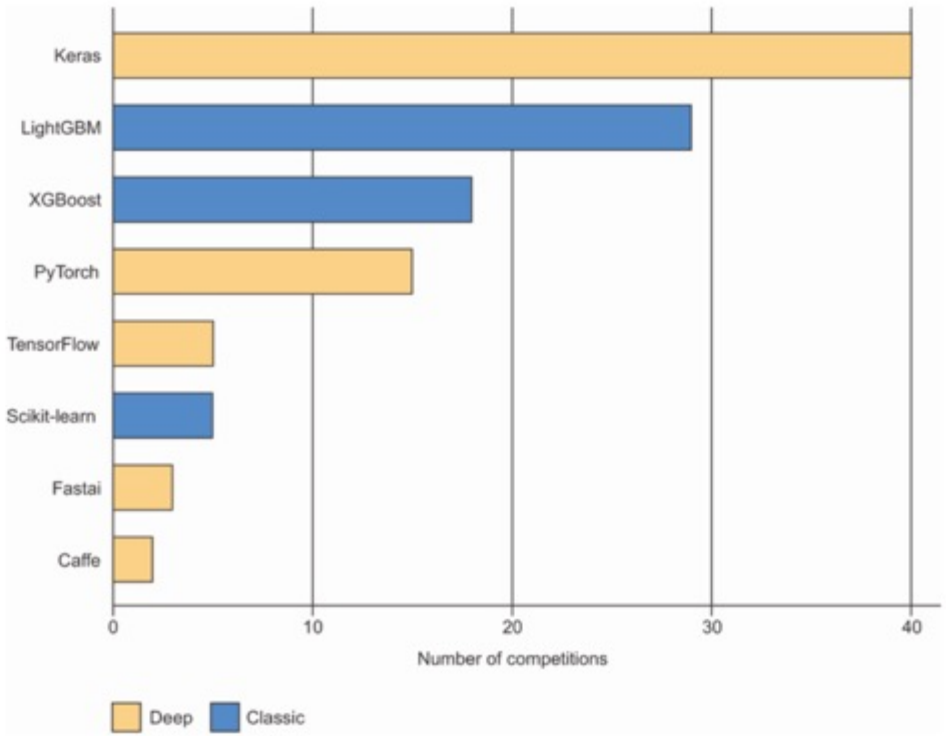
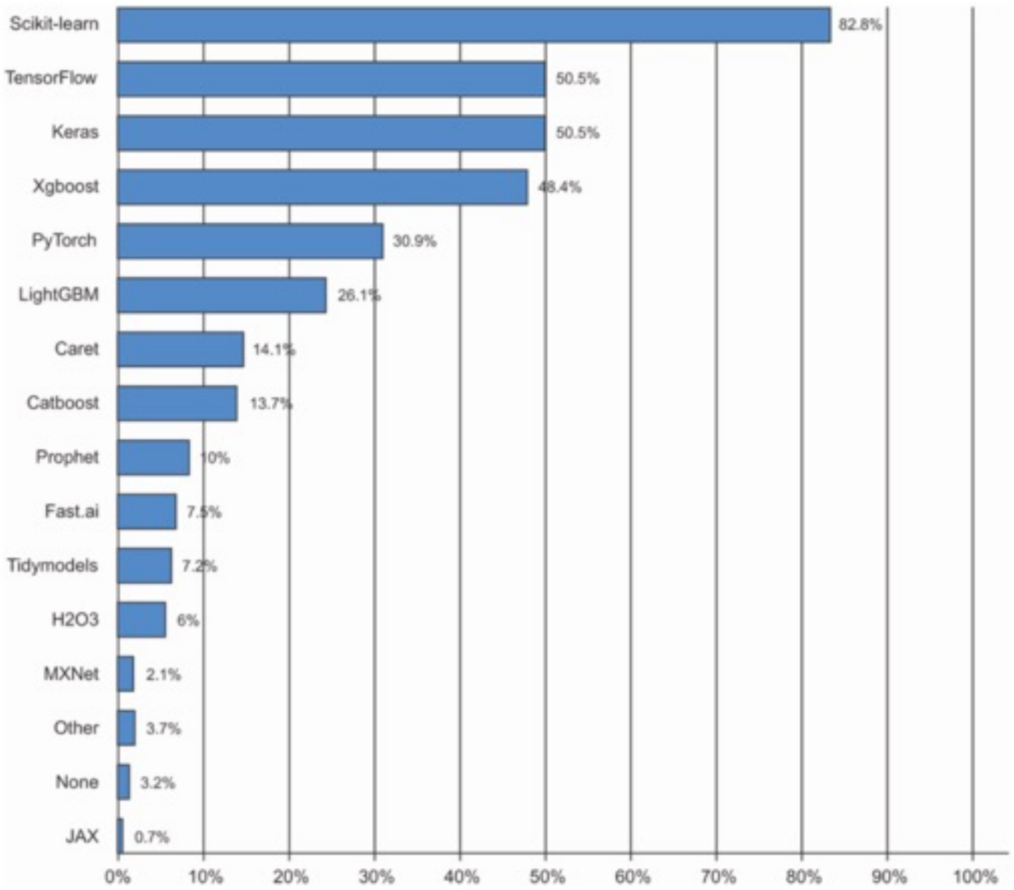
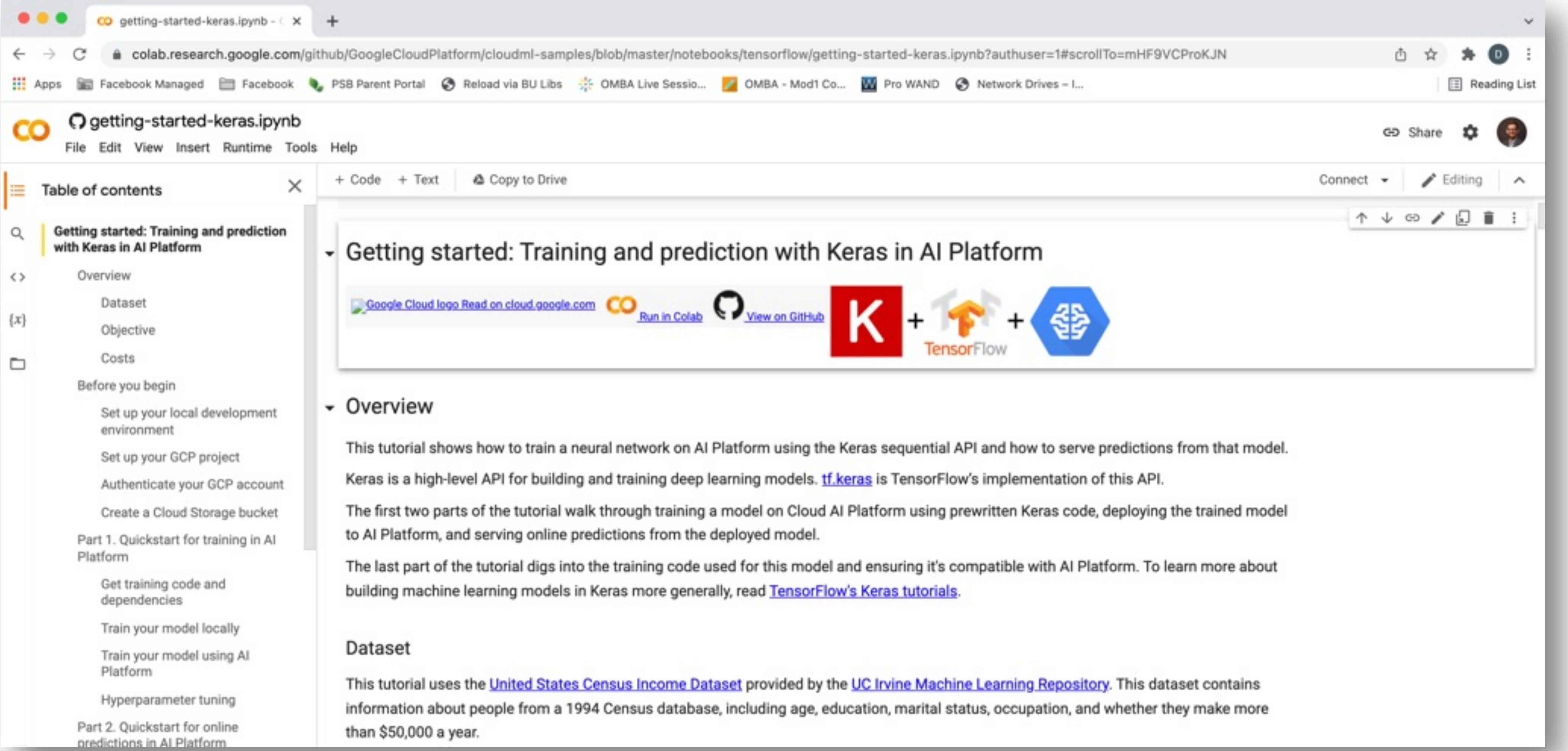


Figure 1.13 Tool usage across the machine learning and data science industry (Source: www.kaggle.com/kaggle-survey-2020)



Google Colab



Course Timeline

AGENDA

- We will start with the concepts.
- We will then get into basic neural networks for simple prediction problems.
- Finally, we will get into more complex use cases, namely computer vision, NLP, and time series.

NOTE TIMING OF DELIVERABLES

- First homework on basic prediction.
- Second homework on image-based prediction.
- There will be sign-ups for the week of 2/8 – 2/10, for your mid-point project check-in meeting.

ONE FLEX WEEK

- This course is significantly revamped, so I reserve the right to revise our schedule depending on how we progress. Week 7 is tentatively going to cover autoencoders and perhaps another topic, but we may be playing catch-up at that point!

COURSE SCHEDULE
(*Subject to Revision Depending on Progress*)

Week	Date	Topic	Assignments	Readings
1	1/20 (Th)	**Course Logistics / Introduction**	--	Chapter 1
2	1/25 (Tu) 1/27 (Th)	**The Math Behind NNs** **Introduction to Keras and TensorFlow**	-- --	Chapter 2 Chapter 3
3	2/1 (Tu) 2/3 (Th)	**Neural Networks for Classification & Regression** **Model Tuning + Deep Dive on Keras**	HW1 Posted Project Proposal Due --	Chapter 4 Chapters 5 & 7
4	2/8 (Tu) 2/10 (Th)	**Mid-Point Project Check-in Meetings (via Zoom – Sign-up)**	HW1 Due --	Chapter 6
5	2/15 (Tu) 2/17 (Th)	**CNNs (Computer Vision)**	HW2 Posted --	Chapter 8 Chapter 9
6	2/22 (Tu) 2/24 (Th)	**RNNs (Time Series)** **RNNs (Text)**	HW2 Due --	Chapter 10 Chapter 11
7	3/1 (Tu) 3/3 (Th)	**Other Topics (e.g., Autoencoders, GANs)**	-- In-Class Exam	
8	*Spring Break*			
9	3/15 (Tu) 3/17 (Th)	**Final Project Presentations**	-- Final Report Due	

Course Format

LECTURES

- We will meet twice weekly (beginning from Week 2), for ~2.5 hours each session. I will incorporate a 20-minute break in the middle of each session.
- The first half of each session will focus on lecture / concepts / explanation.

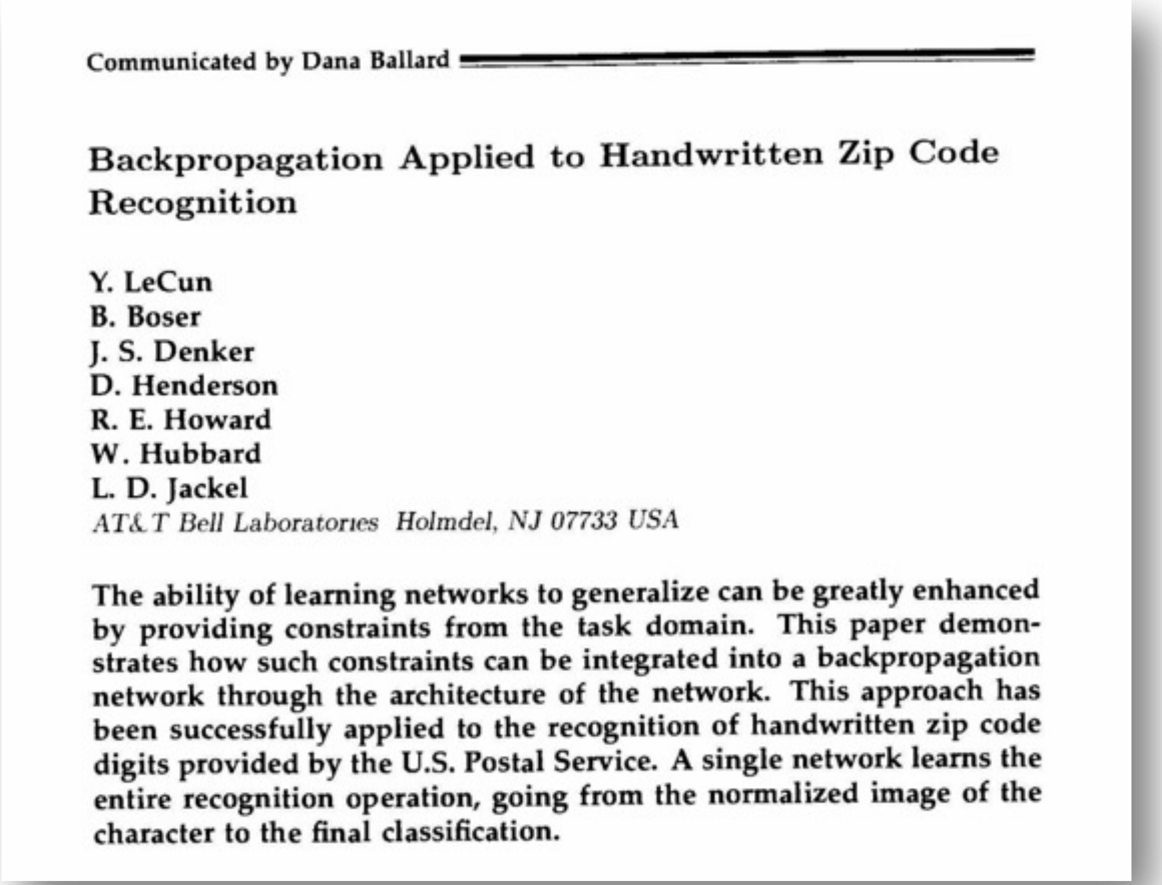
HANDS-ON EXERCISES

- The second half of each session will be focused on implementing hands-on examples / demonstration in Python.
- I will provide you with Jupyter Notebooks and data-sets via GitHub, which we will walk through together.
- You are encouraged to ask questions as we progress.
- Note that the two homework assignments will be based on the in-class lab exercises / questions.

Any Questions?

What is ‘Deep’ (vs. Shallow) Learning?

Where Deep Learning Started


Communicated by Dana Ballard

Backpropagation Applied to Handwritten Zip Code Recognition

Y. LeCun
B. Boser
J. S. Denker
D. Henderson
R. E. Howard
W. Hubbard
L. D. Jackel
AT&T Bell Laboratories Holmdel, NJ 07733 USA

The ability of learning networks to generalize can be greatly enhanced by providing constraints from the task domain. This paper demonstrates how such constraints can be integrated into a backpropagation network through the architecture of the network. This approach has been successfully applied to the recognition of handwritten zip code digits provided by the U.S. Postal Service. A single network learns the entire recognition operation, going from the normalized image of the character to the final classification.

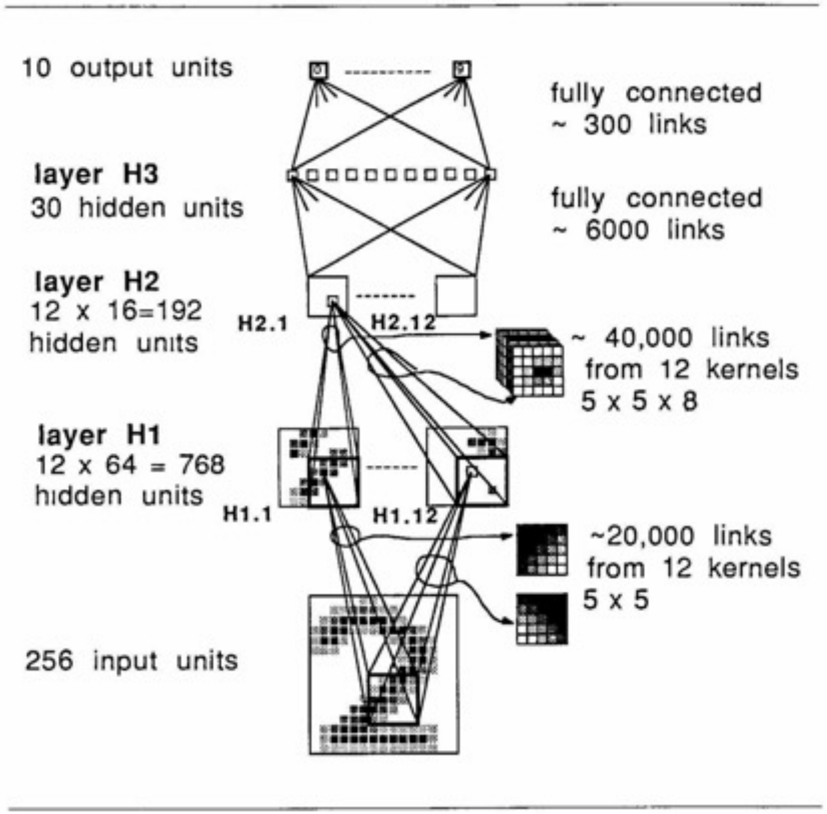
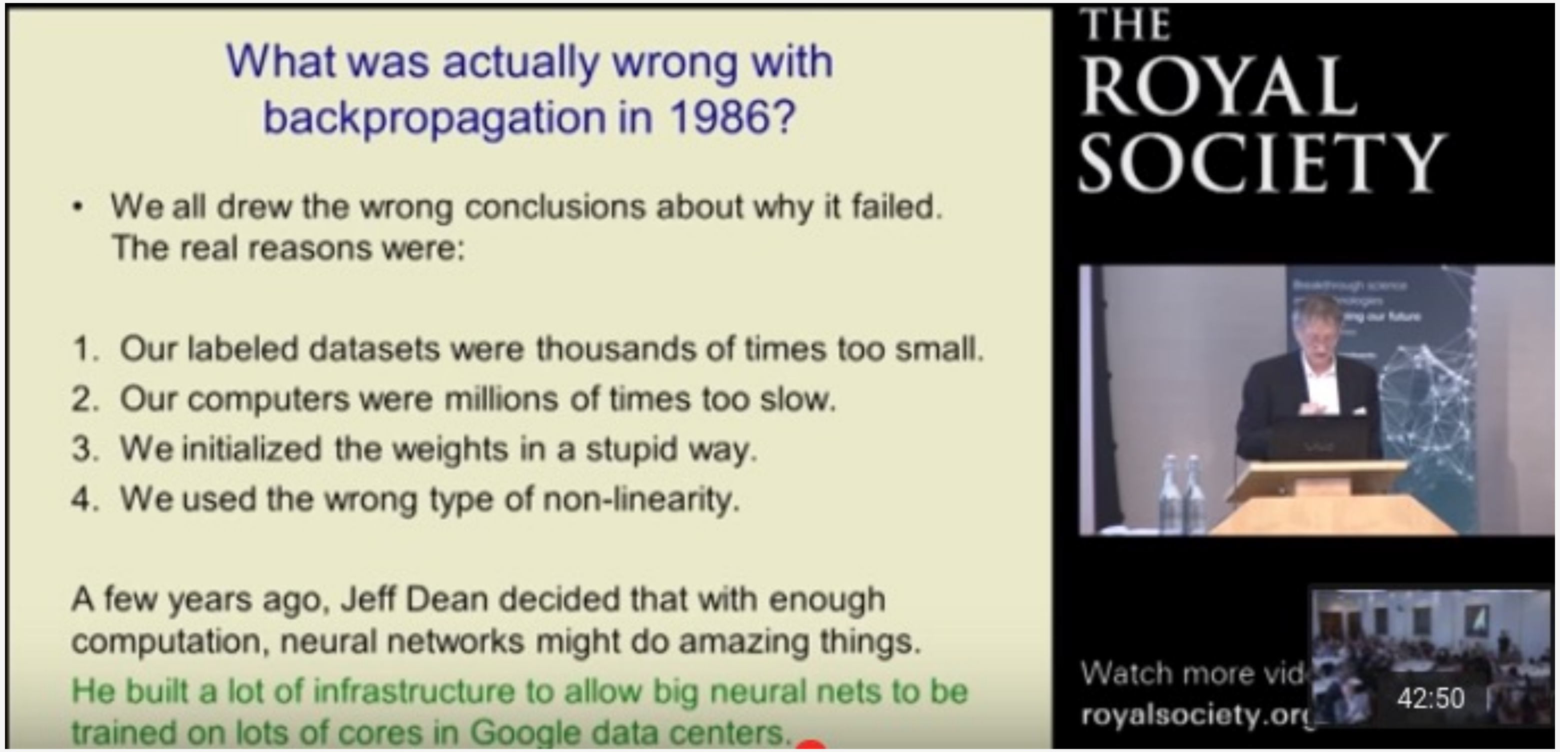


Figure 3 Log mean squared error (MSE) (top) and raw error rate (bottom) versus number of training passes

Then It Shuffled Along for Decades…



Now…

ARTIFICIAL INTELLIGENCE

A GPT-3 bot posted comments on Reddit for a week and no one noticed

Under the username /u/thegentlemetre, the bot was interacting with people on /r/AskReddit, a popular forum for general chat with 30 million users.

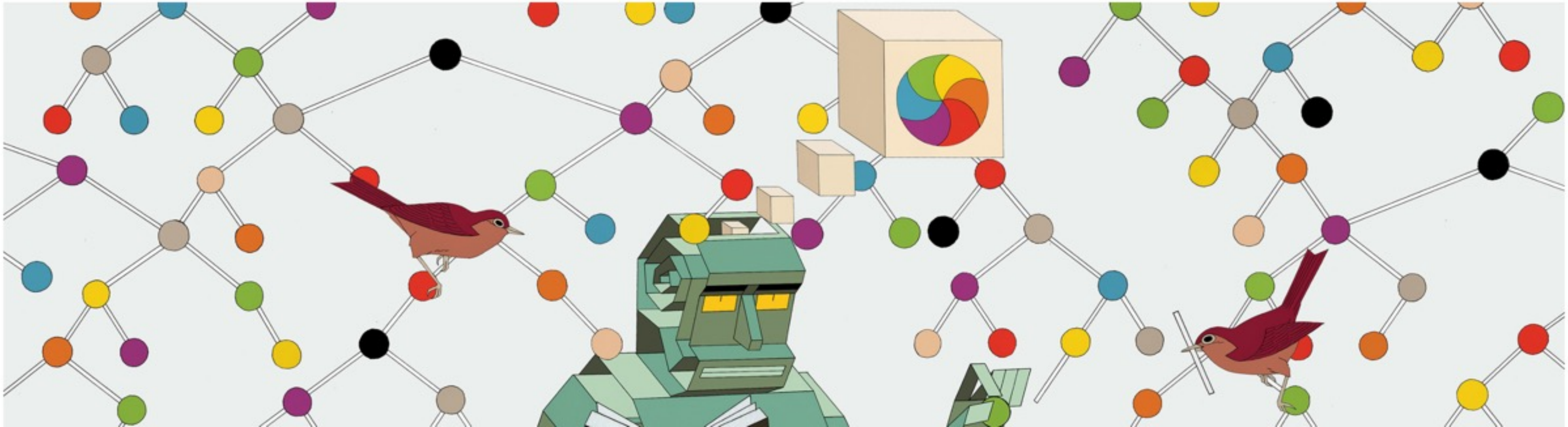
By Will Douglas Heaven

October 8, 2020

ARTIFICIAL INTELLIGENCE

Symbolic Mathematics Finally Yields to Neural Networks

After translating some of math's complicated equations, researchers have created an AI system that they hope will answer even bigger questions.



Predicting Parameters in Deep Learning

Misha Denil[1] **Babak Shakibi**[2] **Laurent Dinh**[3]
Marc'Aurelio Ranzato[4] **Nando de Freitas**[1,2]
[1]University of Oxford, United Kingdom
[2]University of British Columbia, Canada
[3]Université de Montréal, Canada
[4]Facebook Inc., USA
{misha.denil,nando.de.freitas}@cs.ox.ac.uk
laurent.dinh@umontreal.ca
ranzato@fb.com

Abstract

We demonstrate that there is significant redundancy in the parameterization of several deep learning models. Given only a few weight values for each feature it is possible to accurately predict the remaining values. Moreover, we show that not only can the parameter values be predicted, but many of them need not be learned at all. We train several different architectures by learning only a small number of weights and predicting the rest. In the best case we are able to predict more than 95% of the weights of a network without any drop in accuracy.

GitHub Copilot

Learn more >

Technical Preview

Your AI pair programmer

With GitHub Copilot, get suggestions for whole lines or entire functions right inside your editor.

Sign up >

sentiment.ts | write_sql.go | parse_expenses.py | addresses.rb

```
#!/usr/bin/env ts-node

import { fetch } from "fetch-h2";

// Determine whether the sentiment of text is positive
// Use a web service
async function isPositive(text: string): Promise<boolean> {
  const response = await fetch(`http://text-processing.com/api/sentiment/`, {
    method: "POST",
    body: `text=${text}`,
    headers: {
```

Bloomberg Opinion

Money Stuff

Sorry, Zillow's Computer Can't Buy Your House Right Now

Also CEO pay, the Boredom Markets Hypothesis and Big Short guys being big short.

By Matt Levine +Sign Up

October 18, 2021, 1:18 PM EDT

Matt Levine is a Bloomberg Opinion columnist covering finance. He was an editor of Dealbreaker, an investment banker at Goldman Sachs, a mergers and acquisitions lawyer at Wachtell, Lipton, Rosen & Katz, and a clerk for the U.S. Court of Appeals for the 3rd Circuit.

Read more opinion

Zillow

Deciding how much you should pay for a share of large-cap publicly traded stock is not an *entirely* solved problem, but it's pretty close. If someone comes to you and says "hey I have 100 shares of Microsoft Corp. stock for sale, how much will you pay me for it," a pretty decent answer would be to look at the last price at which Microsoft traded – like a millisecond ago – and subtract, you know, one cent from that price. That will get you a price that is likely to be competitive (the seller might actually sell to you), likely to be profitable (you might be able to sell it for more than you paid), and

Forbes

Sub

EDITORS' PICK | Oct 14, 2021, 07:01am EDT | 79,274 views

Fraudsters Cloned Company Director's Voice In $35 Million Bank Heist, Police Find

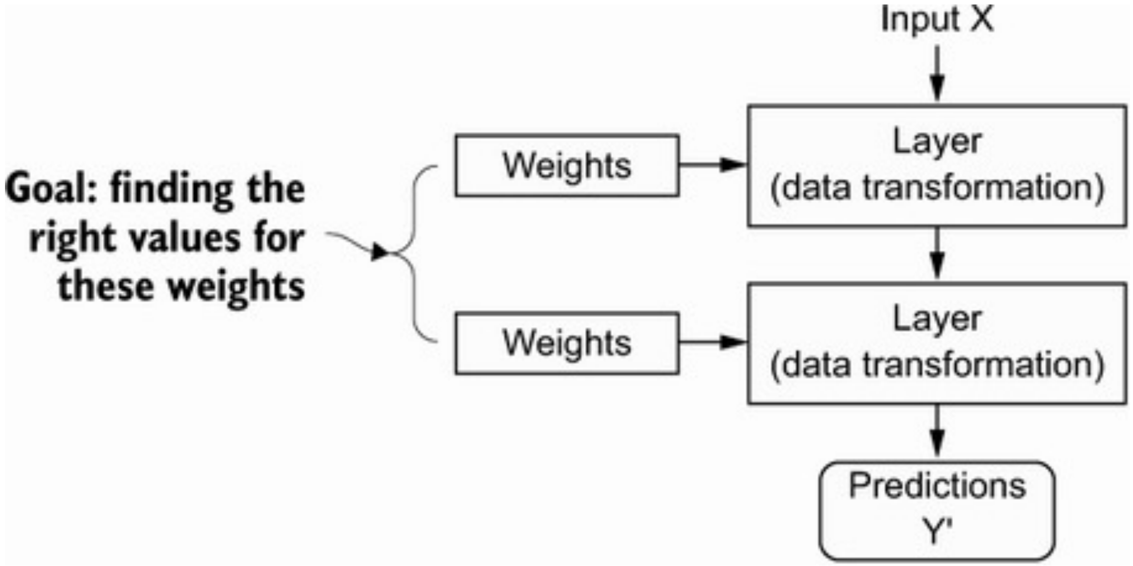
Thomas Brewster Forbes Staff
Cybersecurity
Associate editor at Forbes, covering cybercrime, privacy, security and surveillance.

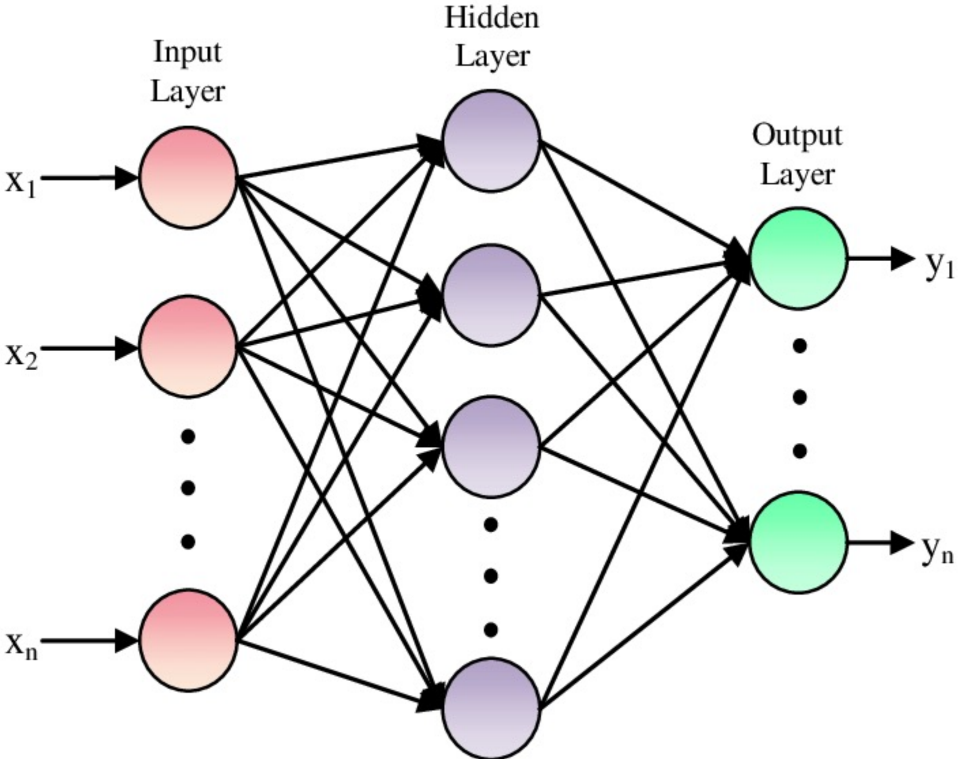
Gord Burtch
@gburtch

I made this video for free in 5-10 minutes, using my phone, a couple git repositories, ffmpeg and Google colab. The image is based on one picture of trump, and 5 seconds of audio from a speech. Just imagine what you someone can do with time, effort and skill…

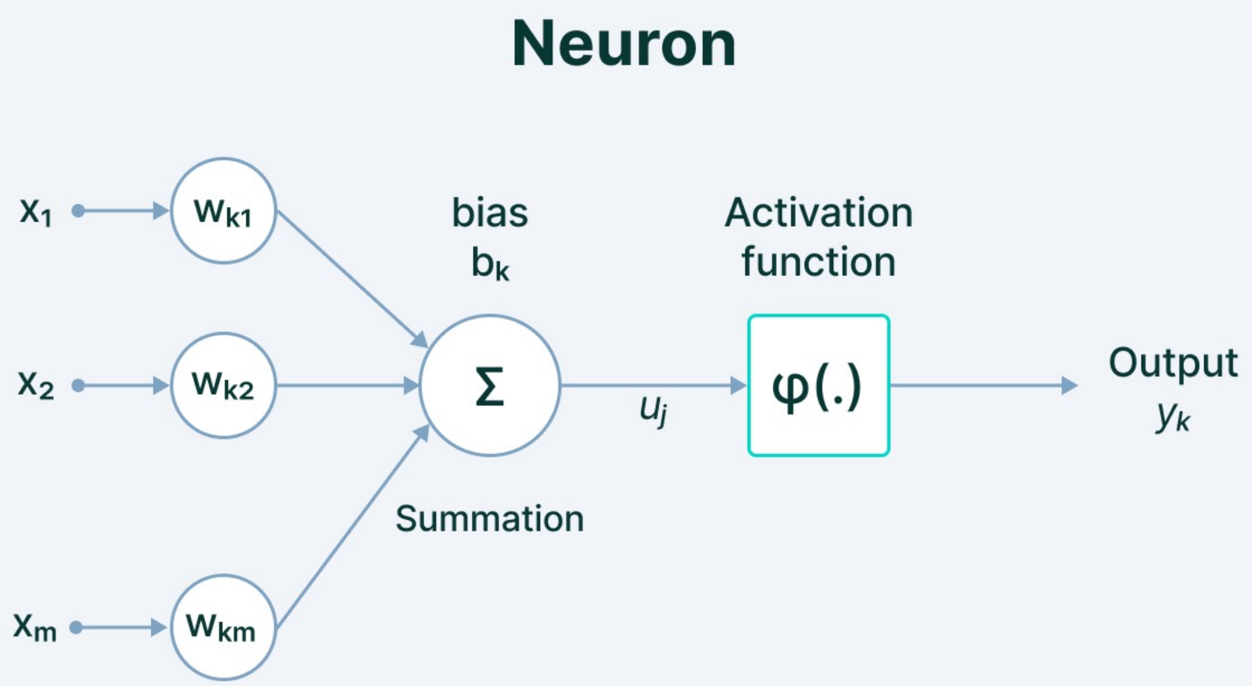
How It Works, Conceptually

Model Parameters

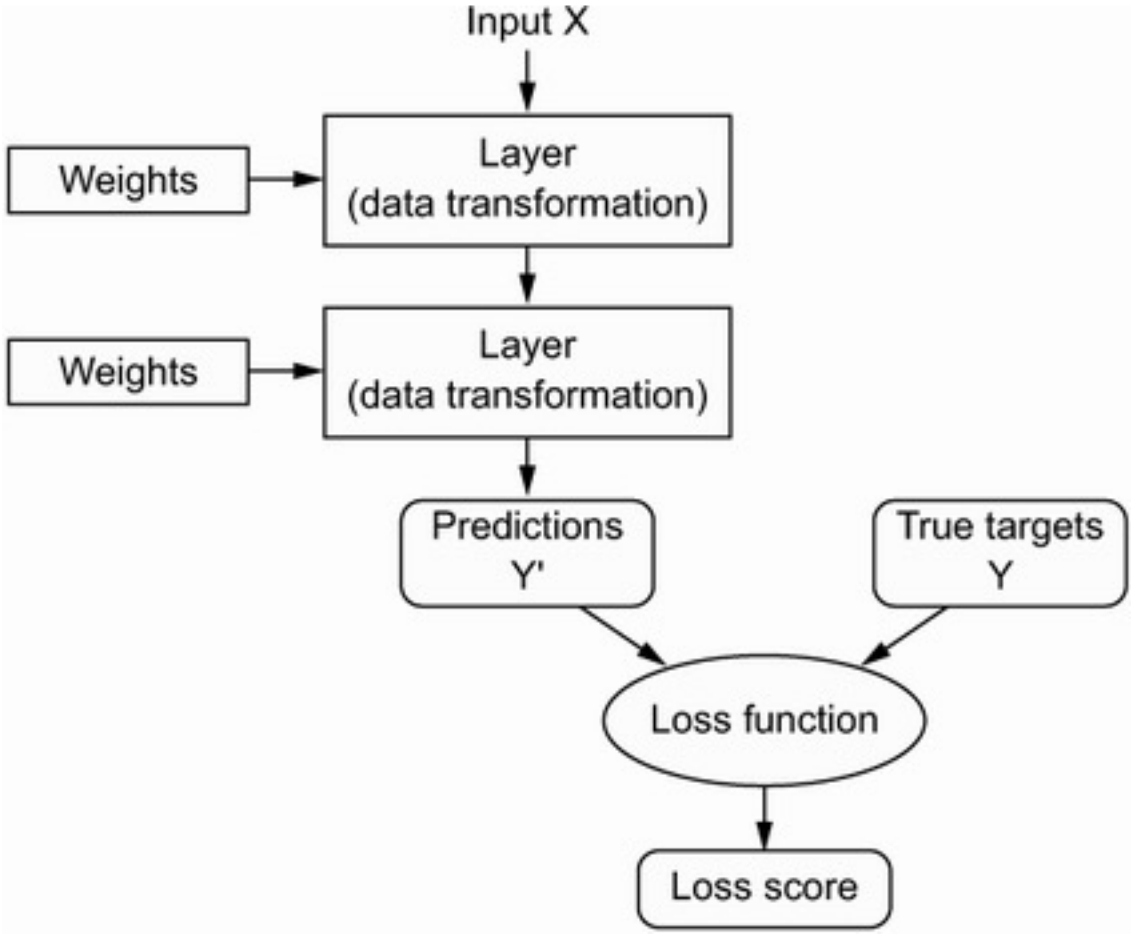




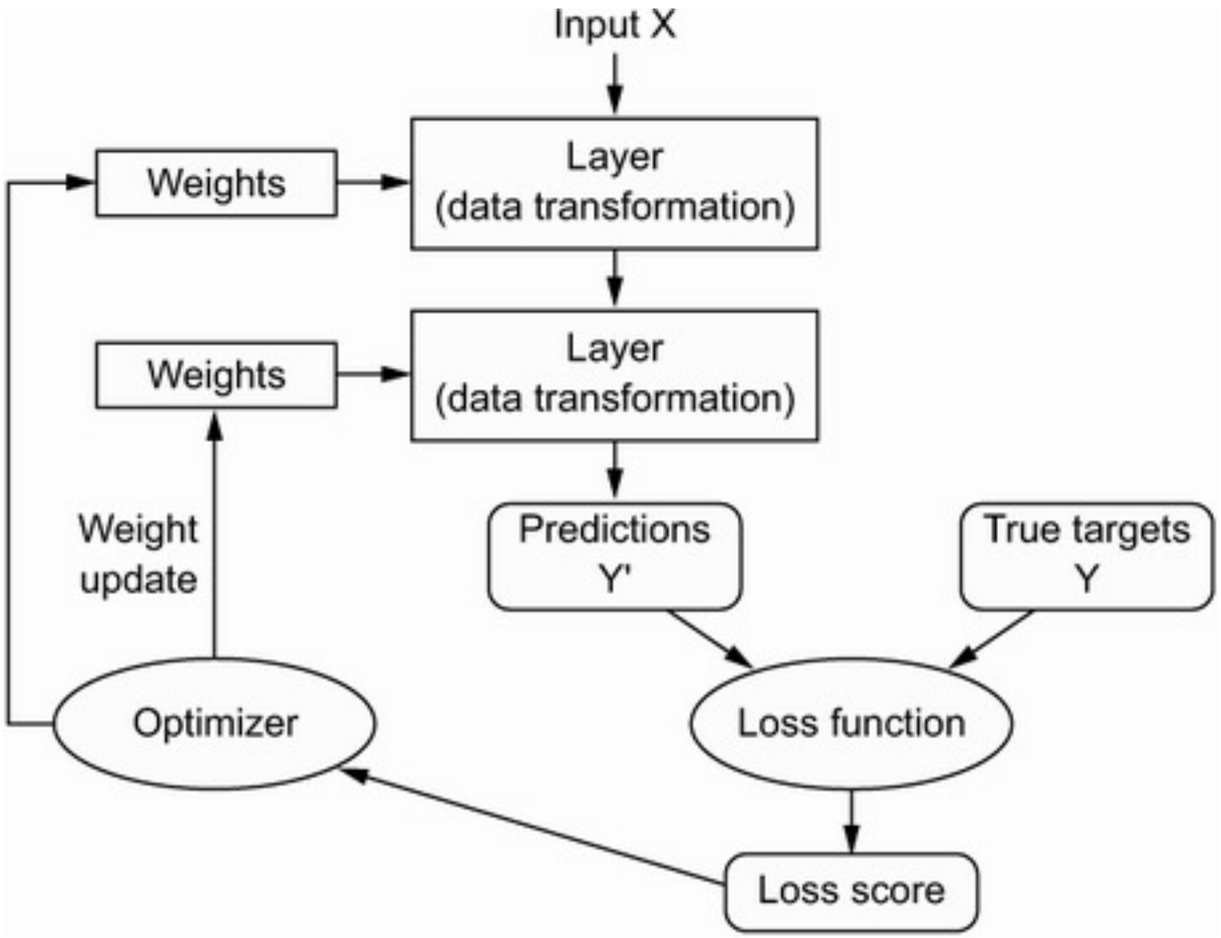
Model Parameters



Loss Function (Error)



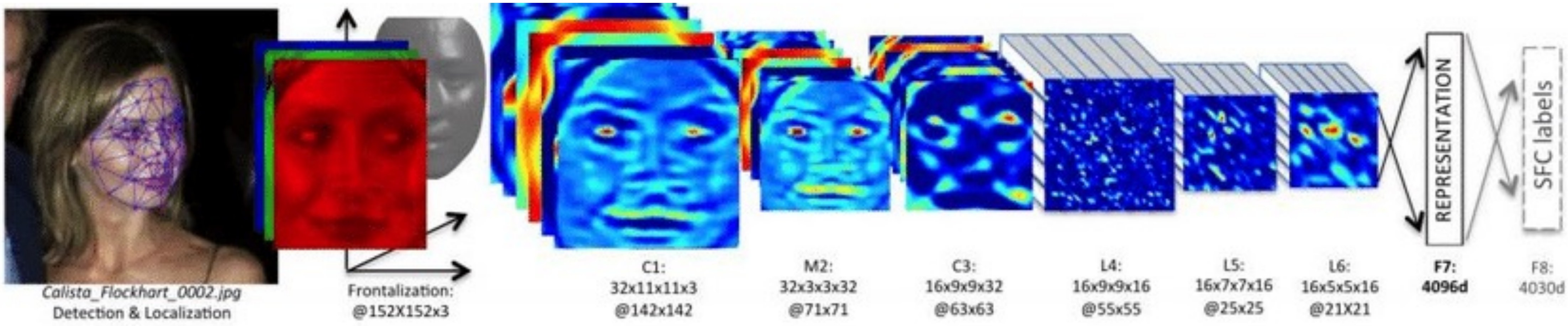
Optimization



When to Learn Deeply (vs. Not)

COMPLEX RELATIONSHIPS

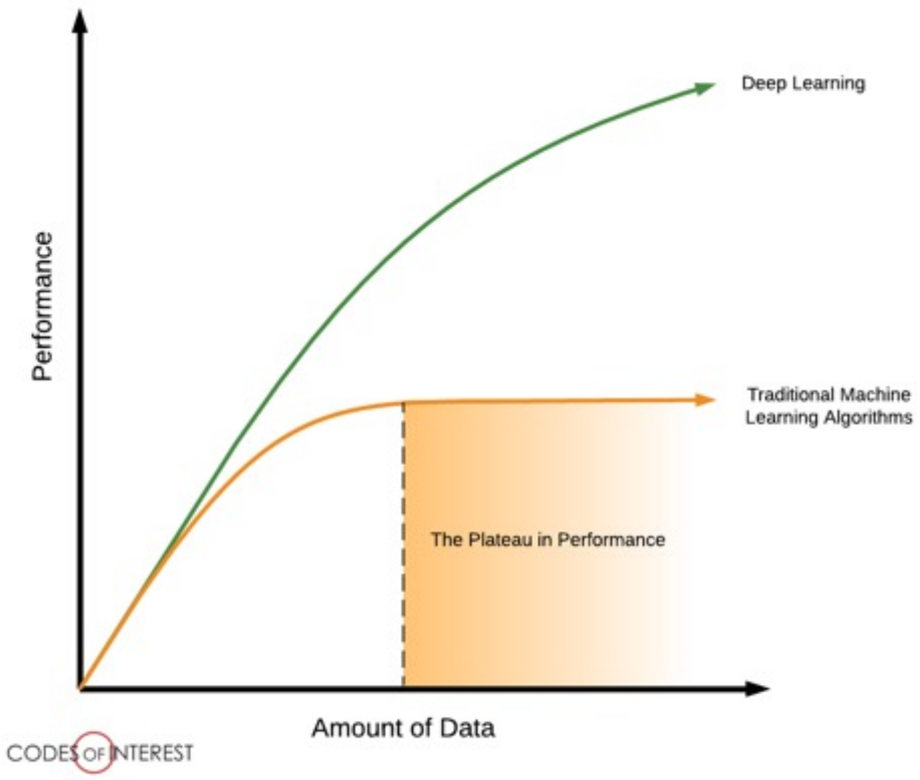
- Complex, non-linear, interactive relationships and mappings; common use cases involve unstructured (high dimensional) data. Deep learning techniques remove the need for feature engineering, a daunting task.



When to Learn Deeply (vs. Nσt)

LOTS OF DATA ON HAND

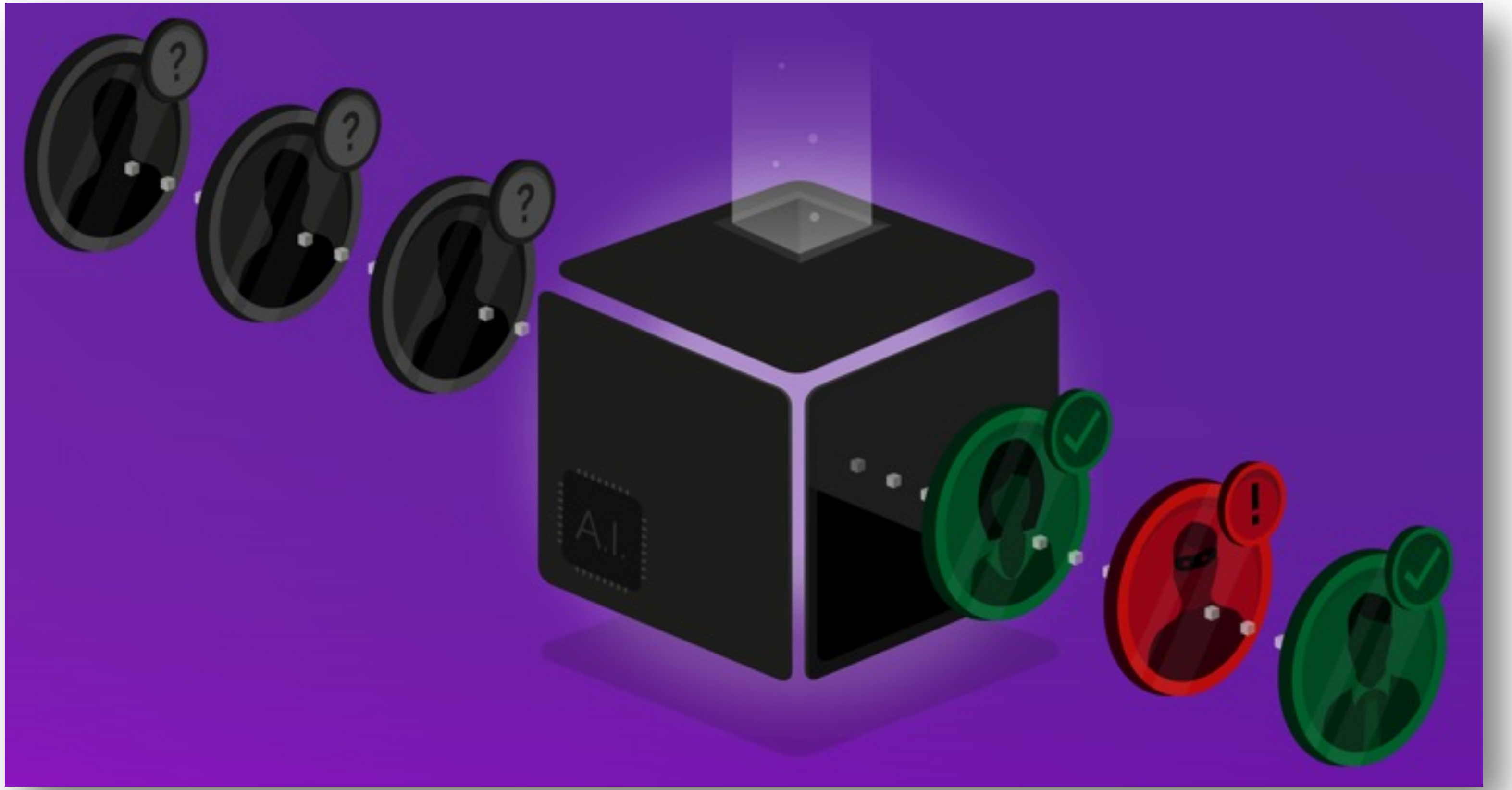
- To be able to learn those complex mappings, typically requires many, many, many training examples.



When to Learn Deeply (vs. Not)

LITTLE NEED FOR UNDERSTANDING

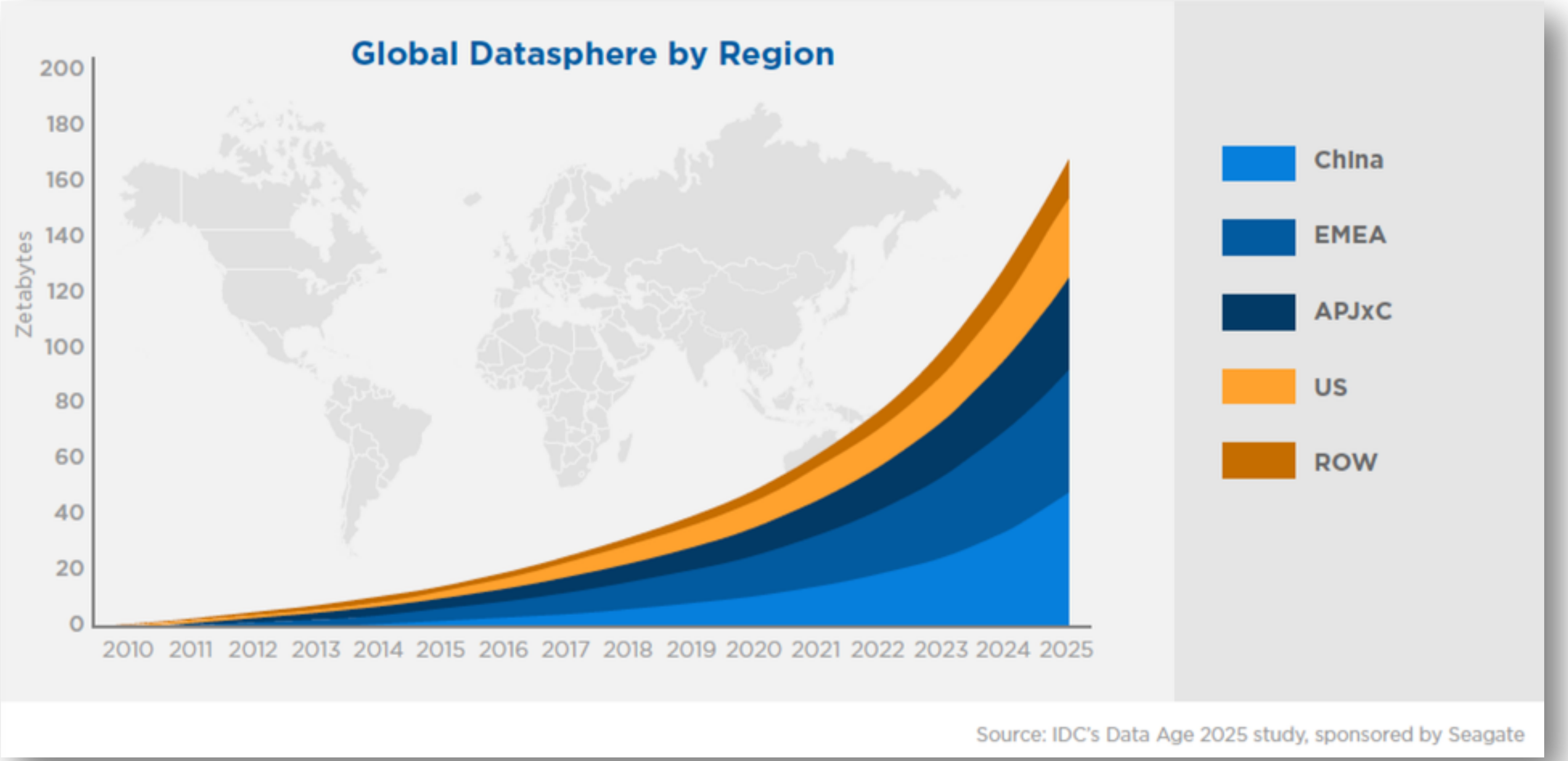
- Although there have been advancements in explainable and interpretable AI, deep nets are notoriously “black box” algorithms.



Why Did Deep Learning Take Off?

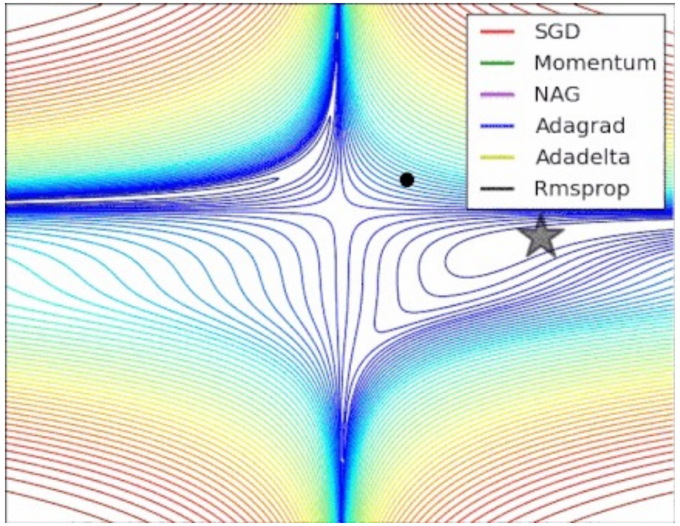
Video Games

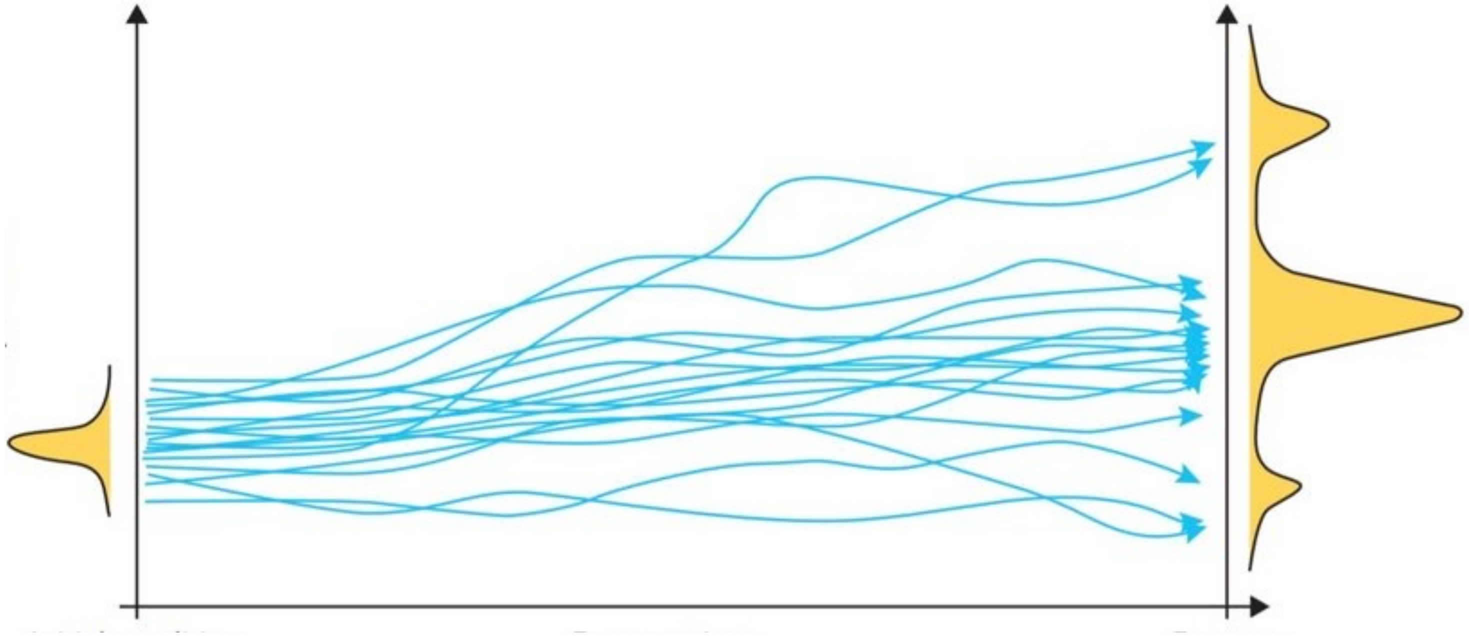
Data

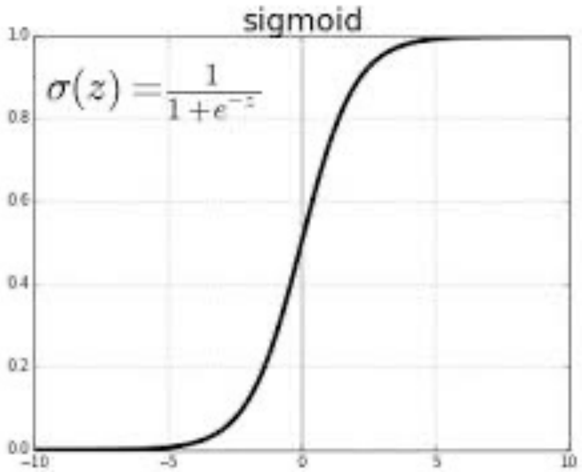


Global Datasphere by Region
Zetabytes
200
180
160
140
120
100
80
60
40
20
0
2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025
China
EMEA
APJxC
US
ROW
Source: IDC's Data Age 2025 study, sponsored by Seagate

Algorithmic Improvements







Questions?

Walk-Through of Google Colab