

# Intro to Neural Nets

Class Logistics and Introduction

# Session Agenda

## 1. INTRO & LOGISTICS

- Who am I? Why should you care what I have to say?
- Schedule and evaluation criteria.
- Delivery format, etc.

## 2. WHAT IS DEEP LEARNING?

- How did we get here? When are these techniques this useful?
- Examples of good, useful, and problematic applications

## 3. THE BASICS & A FIRST EXAMPLE

- What is a neural network?
- How does it work?
- Training our first neural network.





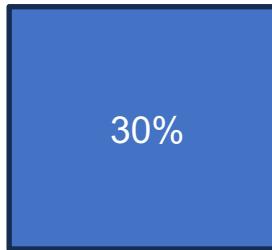
# About Me



**MSI** at the  
**ARF**



# Grading & Evaluation



Premodule quiz



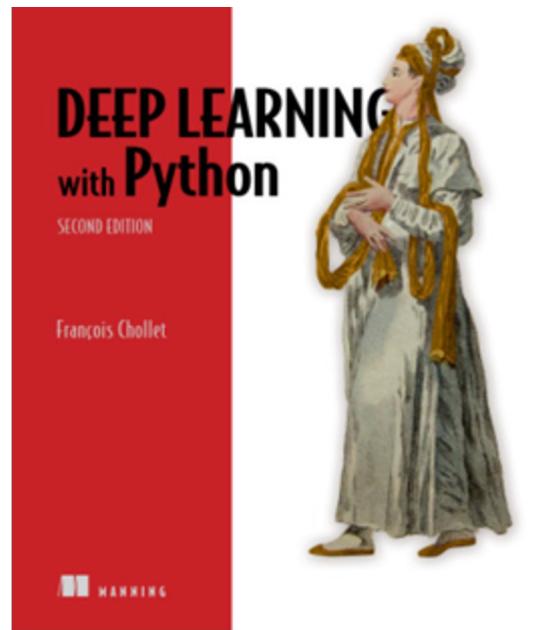
Post module

(40%) INDIVIDUAL PROJECT - Complete the Deep Learning assignment posted on NYU Brightspace . Please feel free to write to me and TF for any questions. The exercise should be done ***independently***. Discussions with other students are allowed, but each student needs to submit his/her individual write-up. Please submit a single Colab/RMD/HTML/PDF file as your homework submission (with your name and ID)

**PLUS**

30% GROUP PROJECT -- Assume you are a technical leader or business manager of a company, Brainstorm the main business of your company and the size of the business. Please describe the details of your virtual company. Describe how AI can potentially help your business. Describe in more detail about an AI strategy that you would like to implement. You should provide the timeline, the team, and the data acquisition strategy, and any other ideas that you would like to share. Based on the information above, please organize them and write an essay on the plan of implementing an AI strategy in your "virtual" or "real" company (12 point font, 1.5 line space, at maximum 2 pages).

# Textbook



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

# We Will Be Using Keras...

## SOFTWARE CONFIGURATION

- You can access Google Colab at <https://colab.research.google.com>. You should have setup a colab account already.



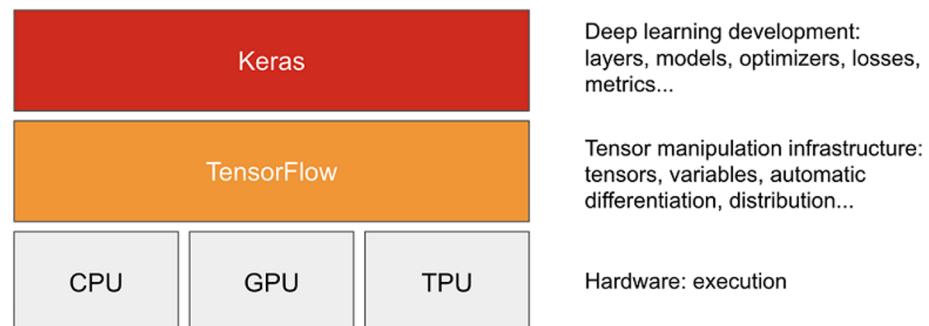
# Keras and Tensorflow

## 1. Tensorflow

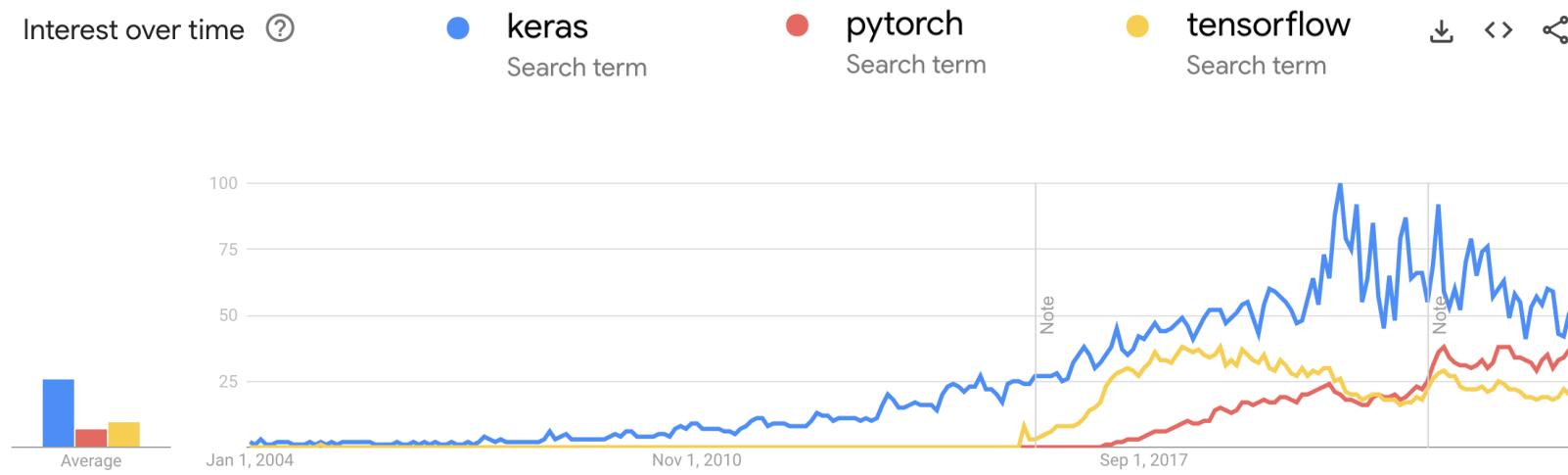
- A Python platform for working with tensors, implementing automatic differentiation, providing access to repositories of (well-known) pre-trained models.

## 2. Keras

- A higher-level API that wraps common usage patterns with Tensorflow functions, pre-defined loss functions, optimization algorithms, etc.
- Keras simplifies data scientists' interaction with Tensorflow.



# Why Keras?



Course



## LECTURE AND DISCUSSION

- For each session block over the next two days, I will begin by presenting concepts, describing logic, discussing implementation considerations.

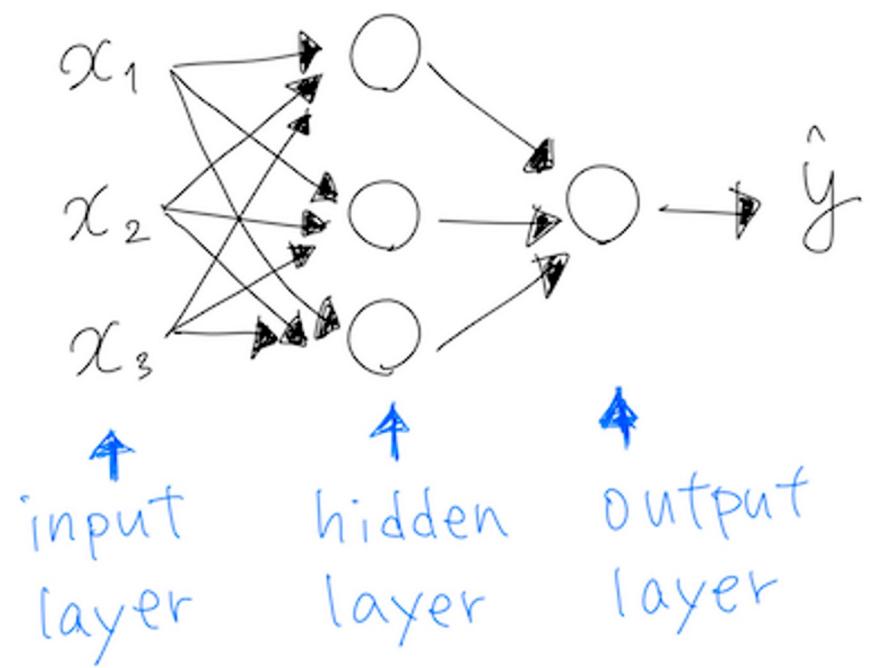
## HANDS-ON EXERCISES

- The latter half of each session will be dedicated to hands on examples in Colab notebooks, where I will walk through the code and explain the logic of what we are doing, allow you to ask questions, etc.

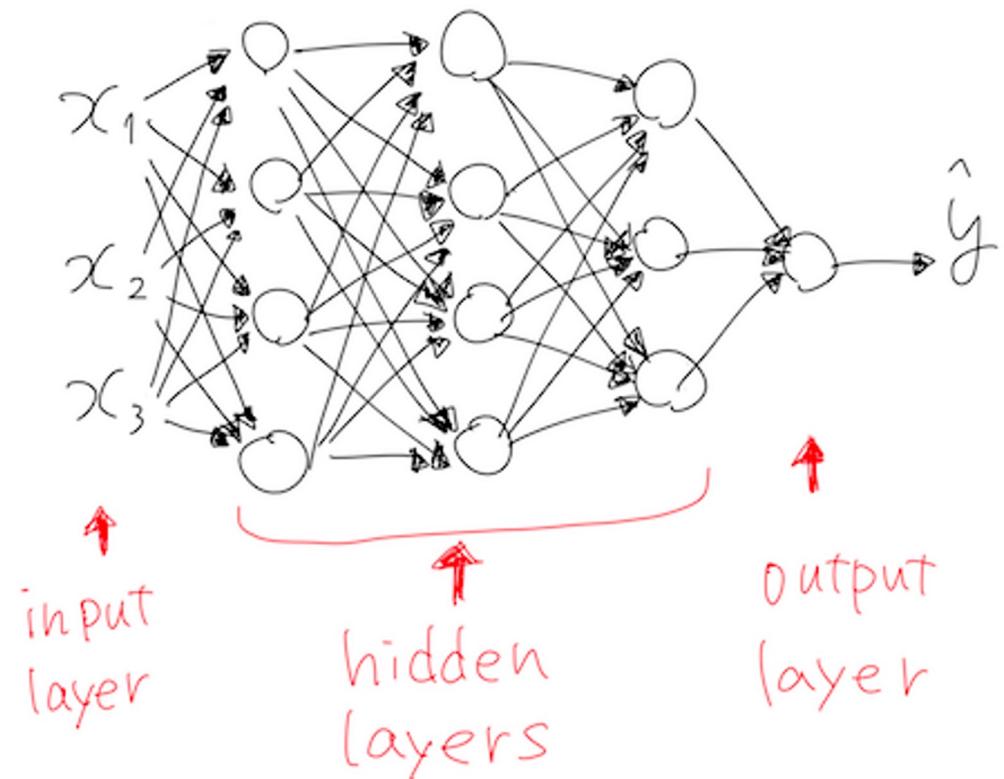
Question: What is ‘Deep’  
(vs. Shallow) Learning?

# What is 'Deep' Learning?

Shallow Neural Network



Deep Neural Network



# A Very Simple Application

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.

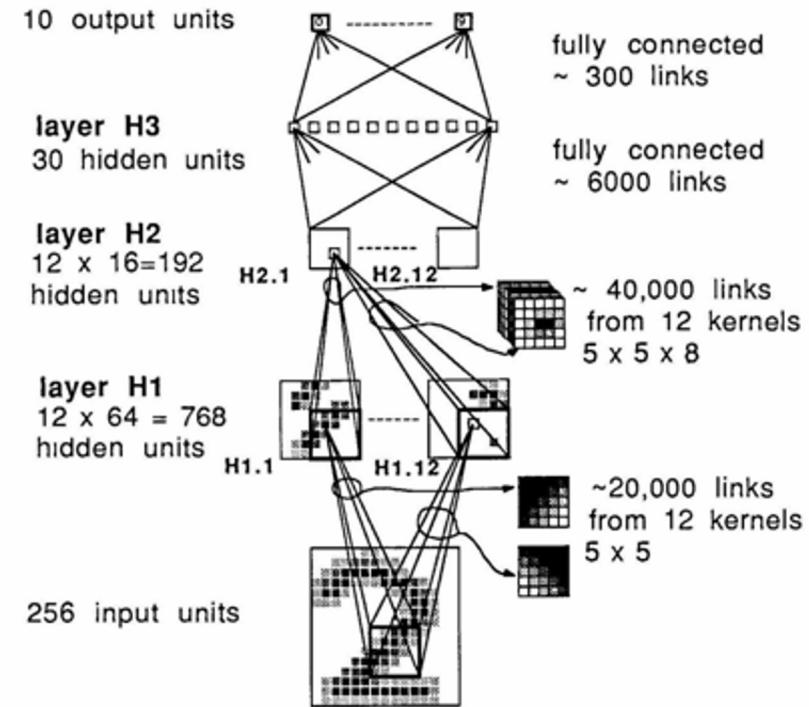


Figure 3 Log mean squared error (MSE) (top) and raw error rate (bottom) versus number of training passes

# Then It Shuffled Along for Decades...

## What was actually wrong with backpropagation in 1986?

- We all drew the wrong conclusions about why it failed.  
The real reasons were:
  1. Our labeled datasets were thousands of times too small.
  2. Our computers were millions of times too slow.
  3. We initialized the weights in a stupid way.
  4. We used the wrong type of non-linearity.

A few years ago, Jeff Dean decided that with enough computation, neural networks might do amazing things.

He built a lot of infrastructure to allow big neural nets to be trained on lots of cores in Google data centers.

THE  
ROYAL  
SOCIETY

A photograph of a man in a suit standing behind a wooden podium, speaking into a microphone. Behind him is a blue banner with the text "Breakthrough science", "Transforming technologies", and "Improving our future". There are two bottles on the podium.

A small video thumbnail showing a group of people seated in an audience, watching a presentation.

Watch more vid  
[royalsociety.org](http://royalsociety.org) 42:50

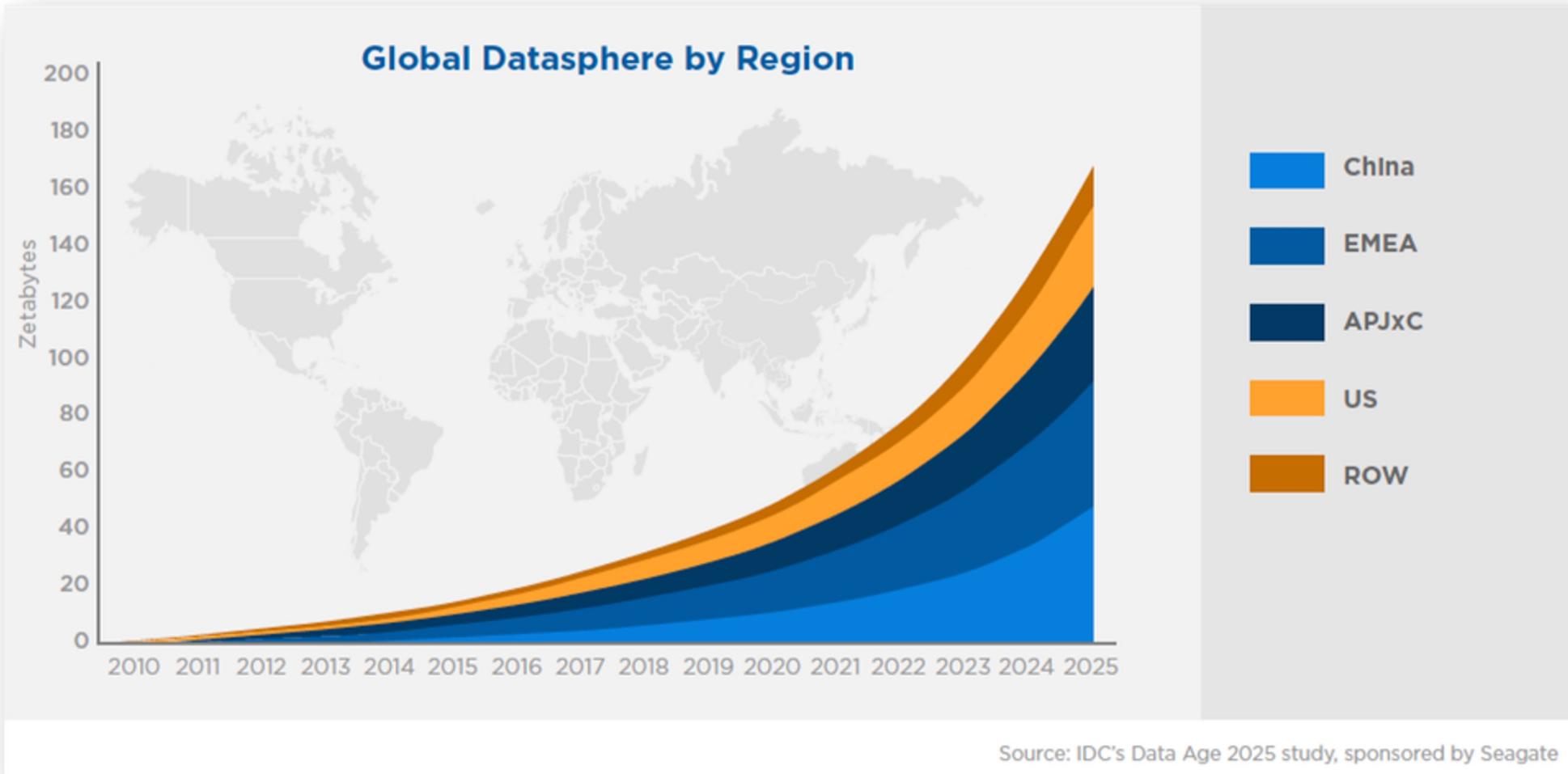
# Now...

# Why Did Deep Learning Take Off?

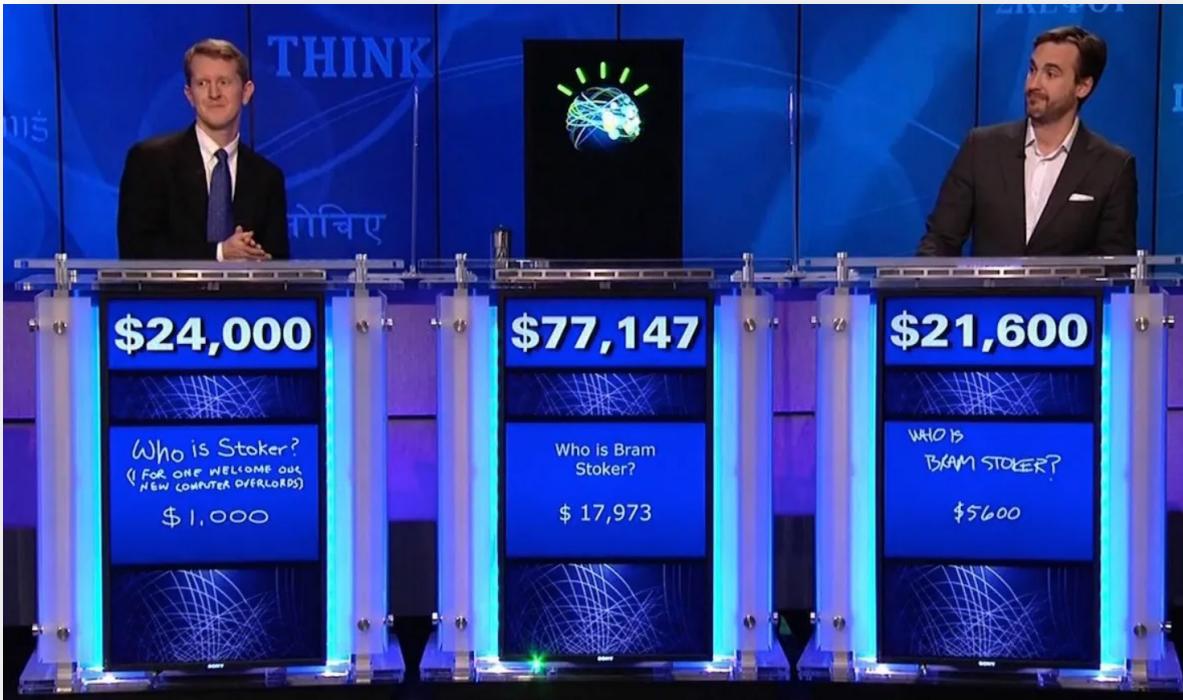


# Video Games

# Data



# Caveat: Lots of Data != Good Data



## M. D. Anderson Breaks With IBM Watson, Raising Questions About Artificial Intelligence in Oncology

By Charlie Schmidt

In 2012, the University of Texas M. D. Anderson Cancer Center in Houston partnered with IBM to develop the artificial intelligence program, called IBM Watson, as a clinical decision tool in oncology. Five years and \$62 million later, M. D. Anderson let its contract with IBM expire before anyone used Watson on actual patients.

Last February, a university audit of the project exposed many procurement problems, cost overruns, and delays. Although the audit took no position on Watson's scientific basis or functional capabilities, it did describe challenges with assimilating Watson into the hospital setting.



Amy Abernethy, M.D.  
Photo credit:  
Flatiron Health

Experts familiar with Watson's applications in oncology describe problems with the system's ability to digest written case reports, doctors' notes, and other text-heavy information generated in

resources. IBM Watson Health, a specialized division commercializing the system in genomics and drug discovery, as well as oncology, has relationships with roughly 50 other institutions.

*"The M. D. Anderson experience is telling us that solving data quality problems in unstructured data is a much bigger challenge for artificial intelligence*

*than was first anticipated."*

Watson's selling point is that it helps doctors stay current with the volume of new findings published every day. IBM has been working on natural language processing capabilities that allow Watson to increasingly understand human speech. Oncologists have been trying to teach the computer system to think like a cancer doctor by training it with real and made-up cases. At M. D. Anderson,

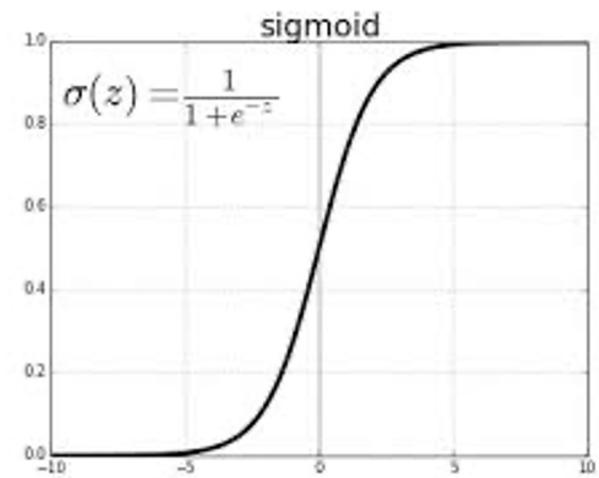
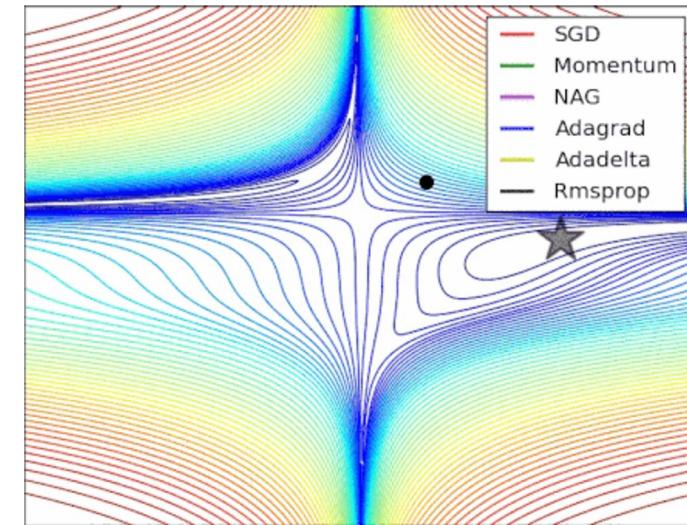
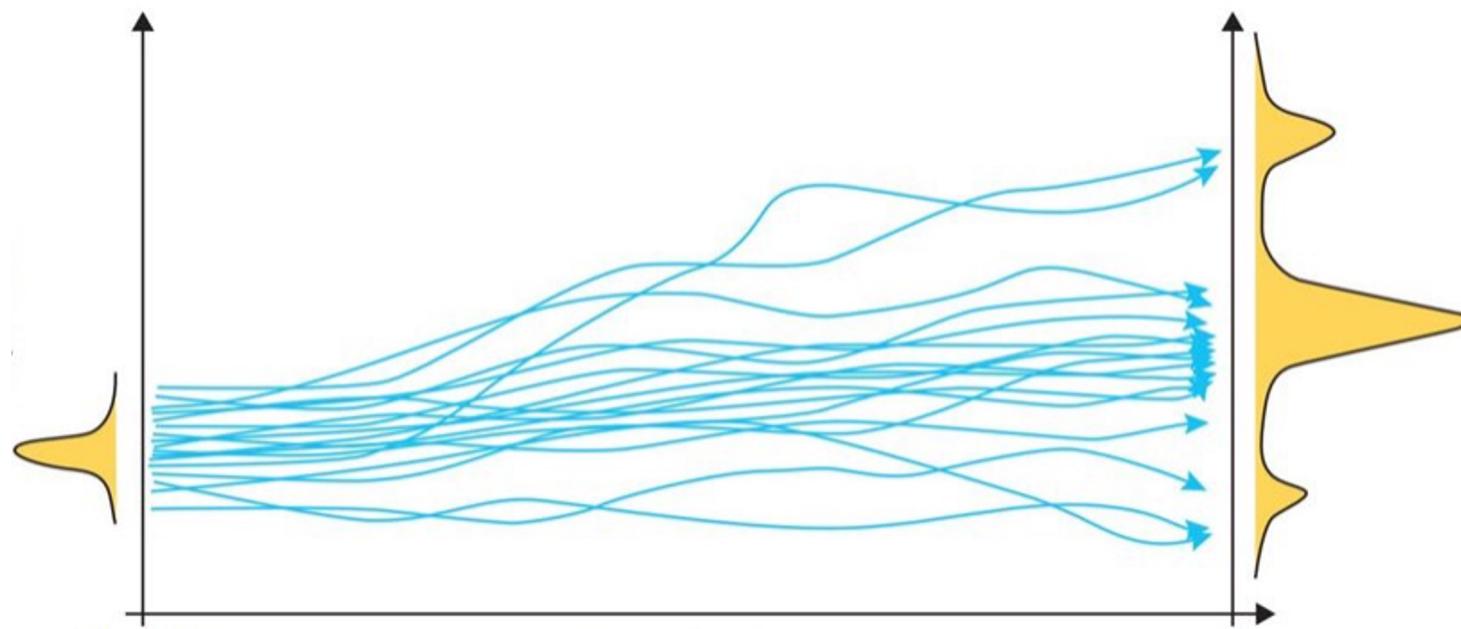
timely completion," the UT audit stated.

Andrew Seidman, M.D., a medical oncologist at the Memorial Sloan Kettering Cancer Center in New York, has been trying to teach Watson how to treat breast cancer. He said that to bring its accumulated knowledge to bear on treating patients at the hospital, Watson needs to be better integrated with electronic medical records.

"The more efficient way to do that would be for it to extract medical attributes automatically," Seidman said. "But many of the attributes we want Watson to digest aren't found in categorical structured data—they're buried in narrative form in doctors' consultation notes." And though natural language processing is central to Watson's functionality, Seidman said that "it remains a work in progress."

Andrew Norden, M.D., deputy chief health officer at IBM Watson Health, acknowledged the need to clarify abbreviations in unstructured medical notes and "other things Watson hasn't seen before." He added, "But this isn't rocket science—the trick is getting the right digital content to Watson so that Watson can read it." In M. D. Anderson's case, a third party—the London-based consult-

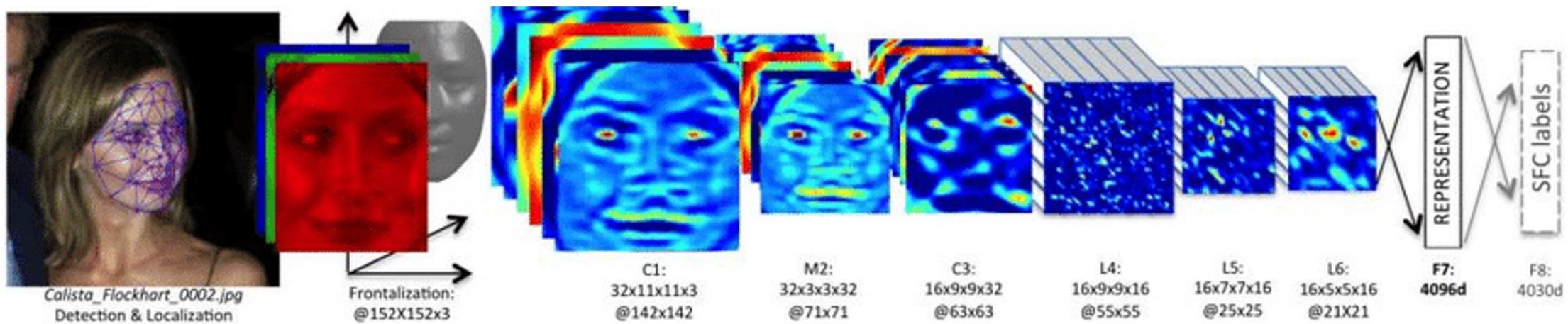
# Algorithmic Improvements



# When to Learn Deeply (vs. Not)

## COMPLEX RELATIONSHIPS

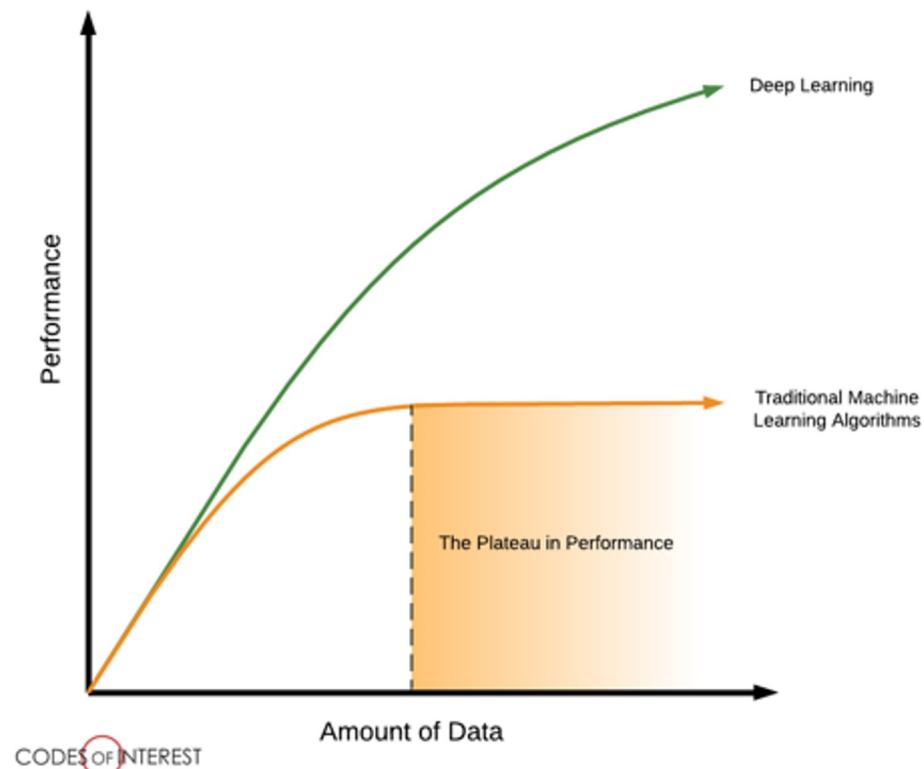
- Complex, non-linear, interactive relationships and mappings; common use cases involve unstructured (high dimensional) data. Deep learning techniques remove the need for feature engineering, a daunting task.



# When to Learn Deeply (vs. Not)

## LOTS OF DATA ON HAND

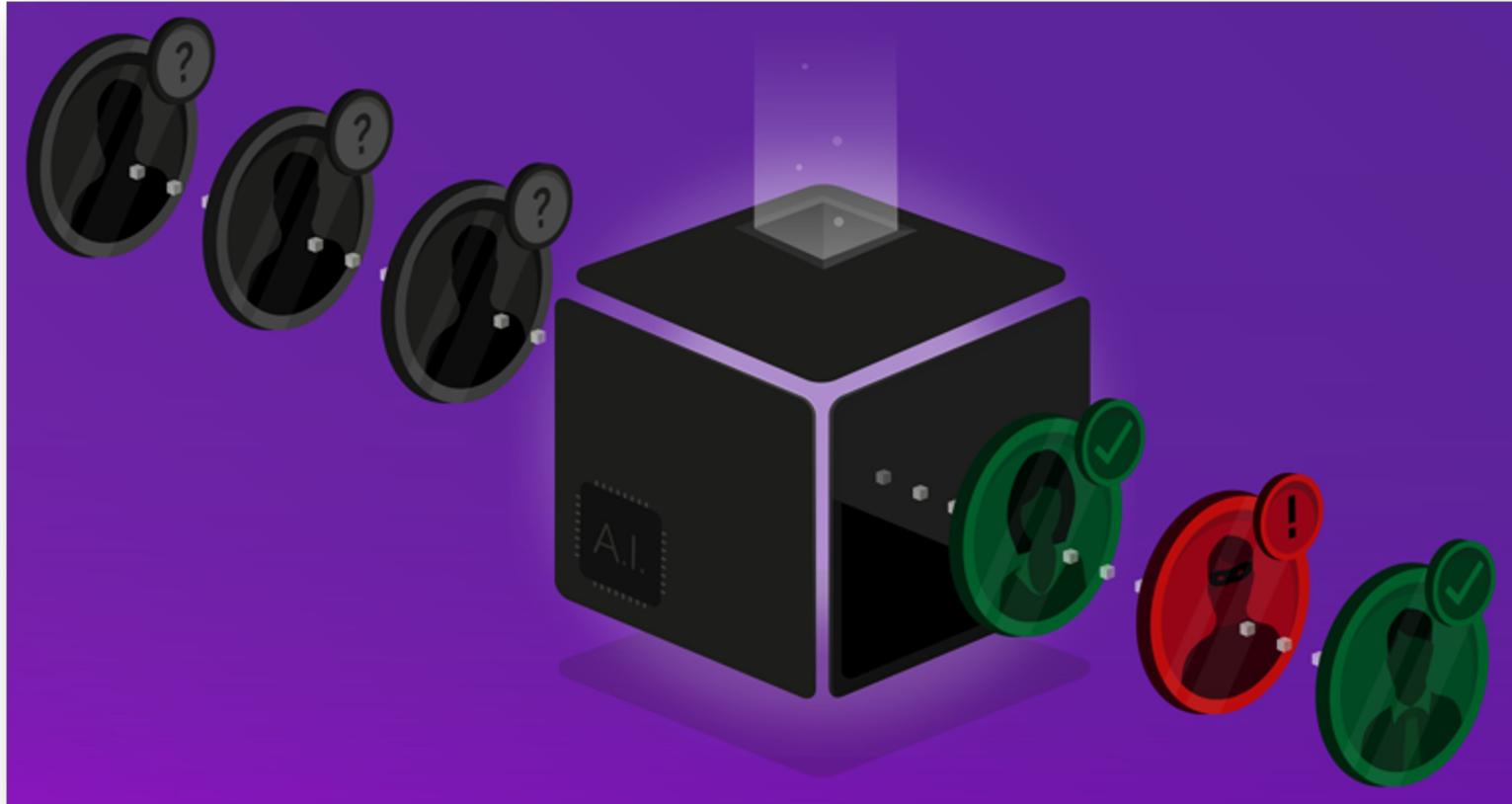
- To be able to learn those complex mappings, typically requires many, many, many training examples.



# 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.



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.

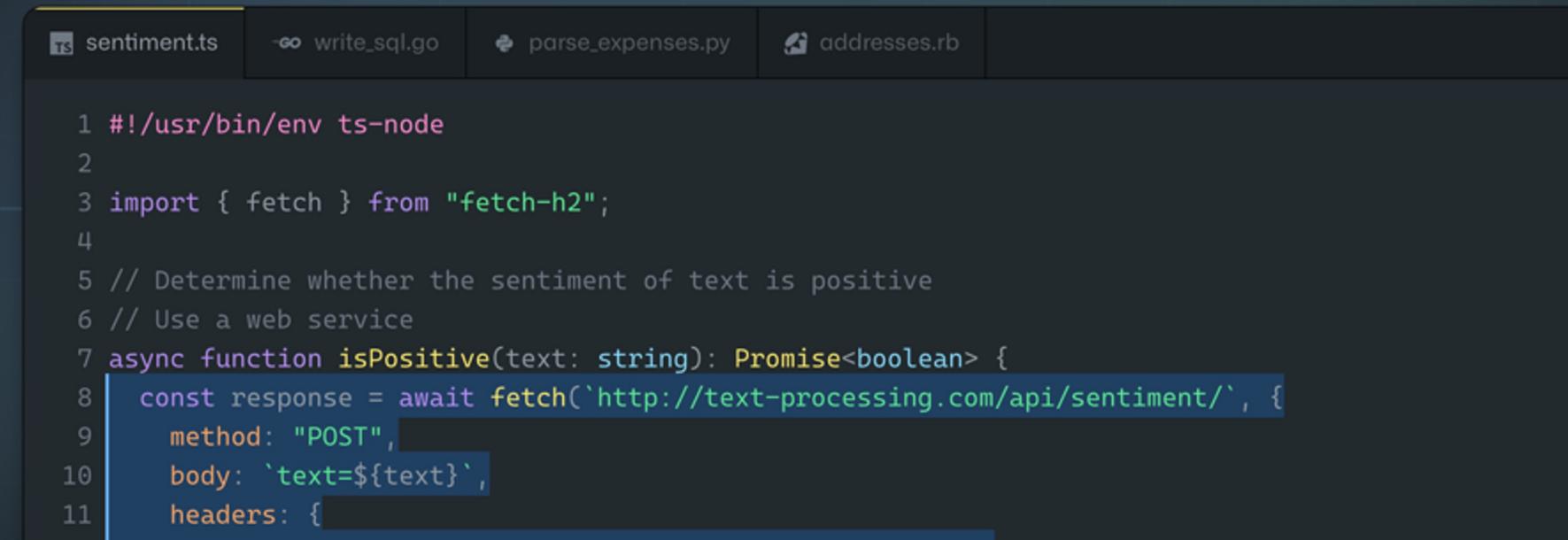
By Will Douglas Heaven

October 8, 2020

Technical Preview

# Your AI pair programmer

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

[Sign up >](#)

A screenshot of a code editor interface. At the top, there are tabs for 'sentiment.ts', 'write\_sql.go', 'parse\_expenses.py', and 'addresses.rb'. The 'sentiment.ts' tab is active. Below the tabs, a snippet of TypeScript code is shown:

```
1 #!/usr/bin/env ts-node
2
3 import { fetch } from "fetch-h2";
4
5 // Determine whether the sentiment of text is positive
6 // Use a web service
7 async function isPositive(text: string): Promise<boolean> {
8   const response = await fetch(`http://text-processing.com/api/sentiment/`, {
9     method: "POST",
10    body: `text=${text}`,
11    headers: {
```

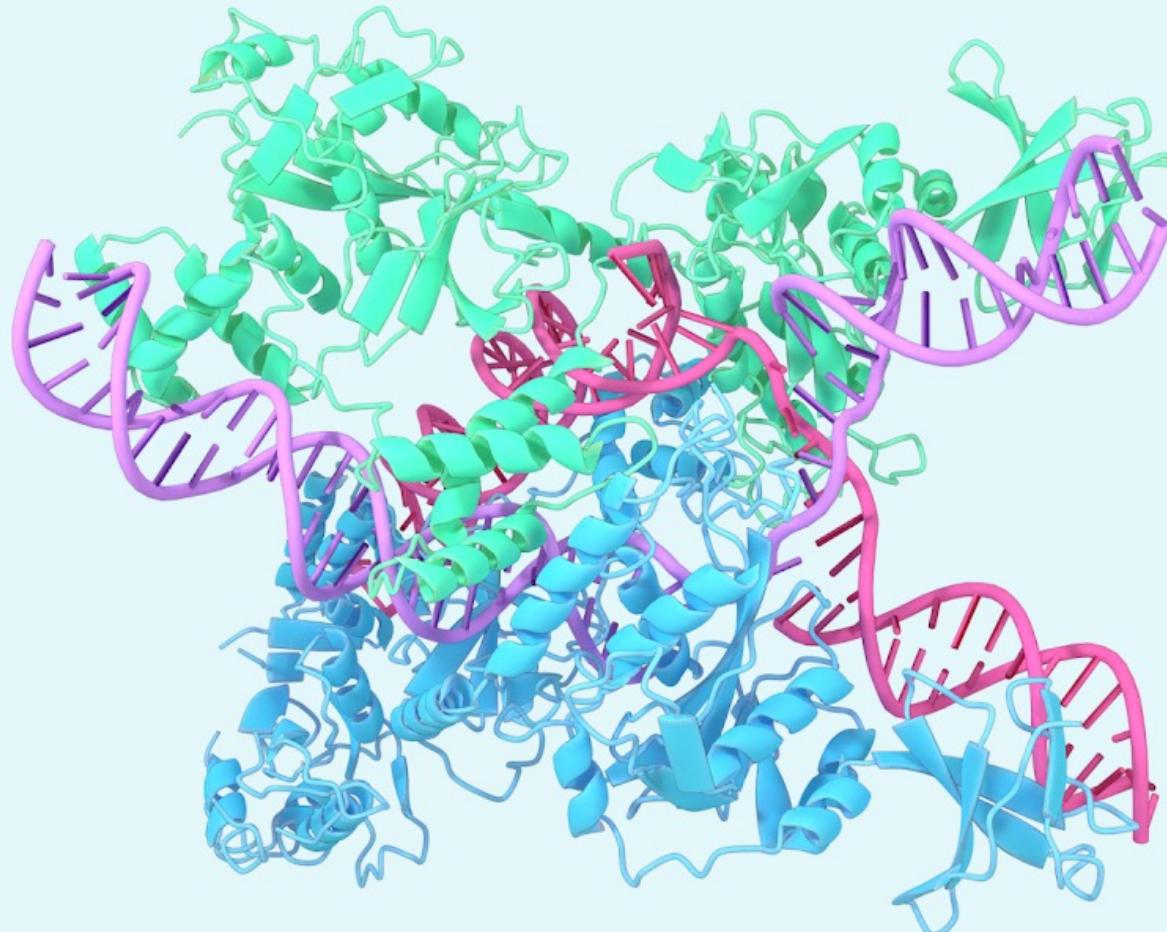
The code editor has a light blue selection bar highlighting the line 'const response = await fetch(`http://text-processing.com/api/sentiment/`, {'. This indicates that GitHub Copilot is currently suggesting or completing this line of code.

TECHNOLOGY

# AlphaFold

AlphaFold is accelerating research in nearly every field of biology.

Google DeepMind



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



ZEGNA

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.*

f t in





Gord Burtch  
@gburtch

...

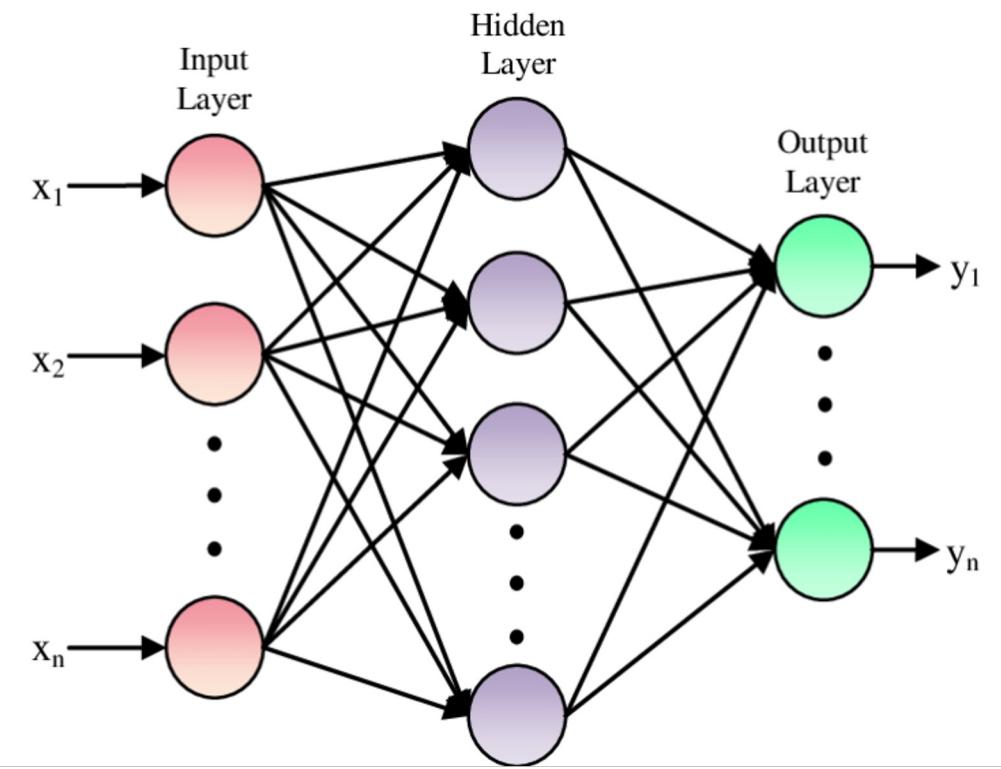
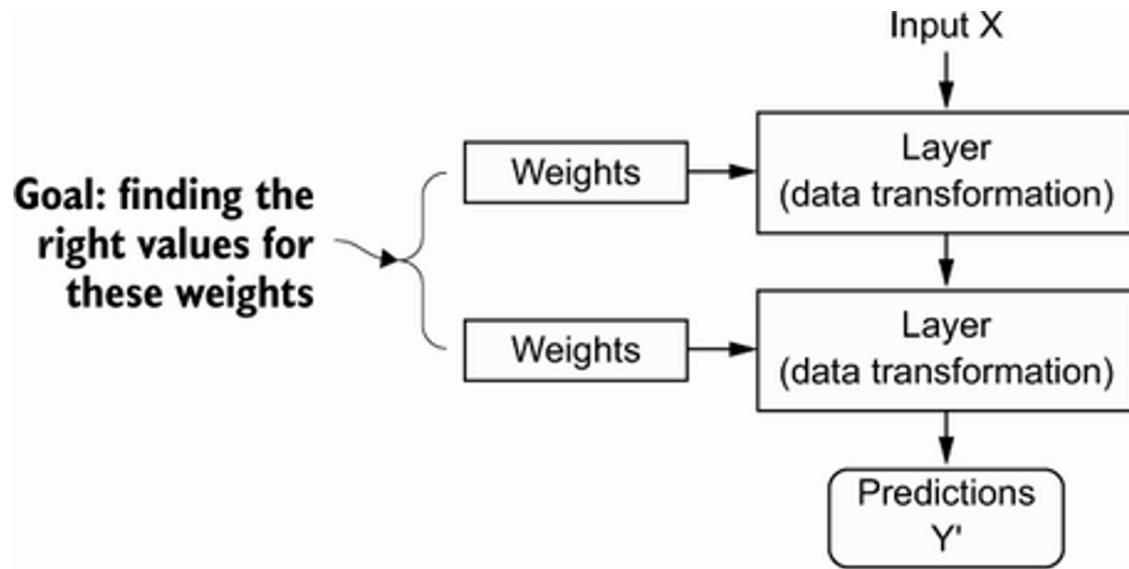
I made this video for free in 5-10 minutes, using my phone, a couple git repositories, ffmpeg and Google colab. The image is based on one picture of trump, and 5 seconds of audio from a speech. Just imagine what you someone can do with time, effort and skill...



<https://twitter.com/gburtch/status/1366806289192517634>

# How Neural Nets Work, Conceptually

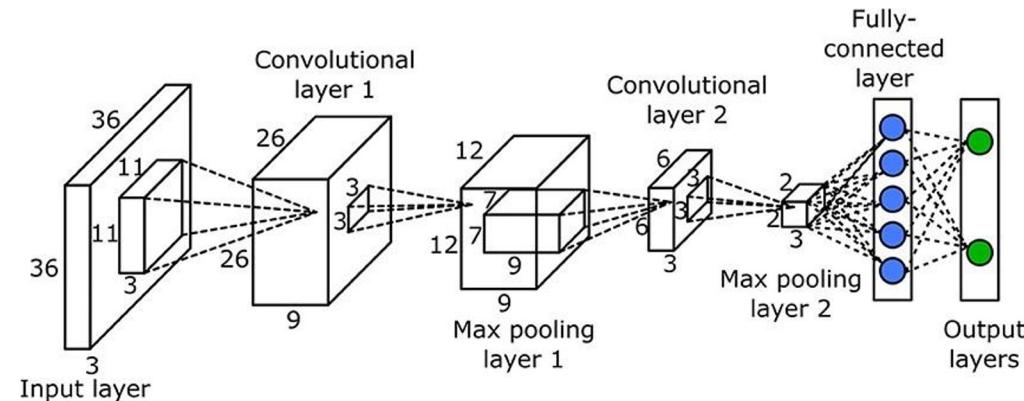
# Model Parameters



# Neural Network Layers

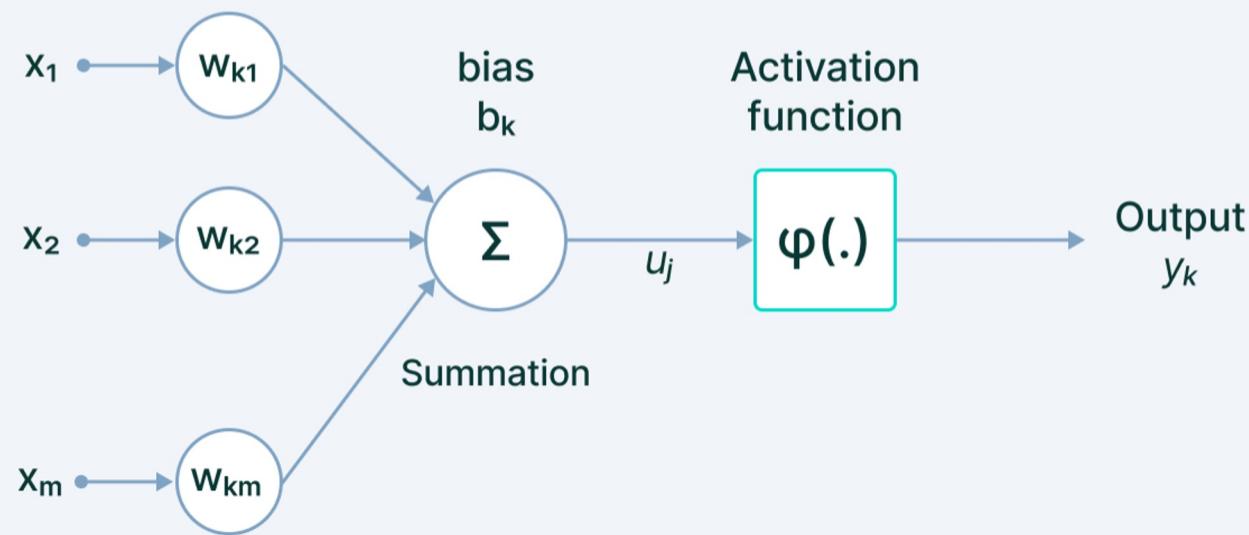
## Layers are the Key Building Block of NNs

- Dense layers are the simplest form, `layers.Dense()`, but there are others, e.g., convolutional layers, max-pooling layers, recurrent layers, and so on:  
<https://keras.io/api/layers/>.
- These are different architectural components that can be mixed and matched in different ways to create different network topologies. It is also possible to construct custom layers.

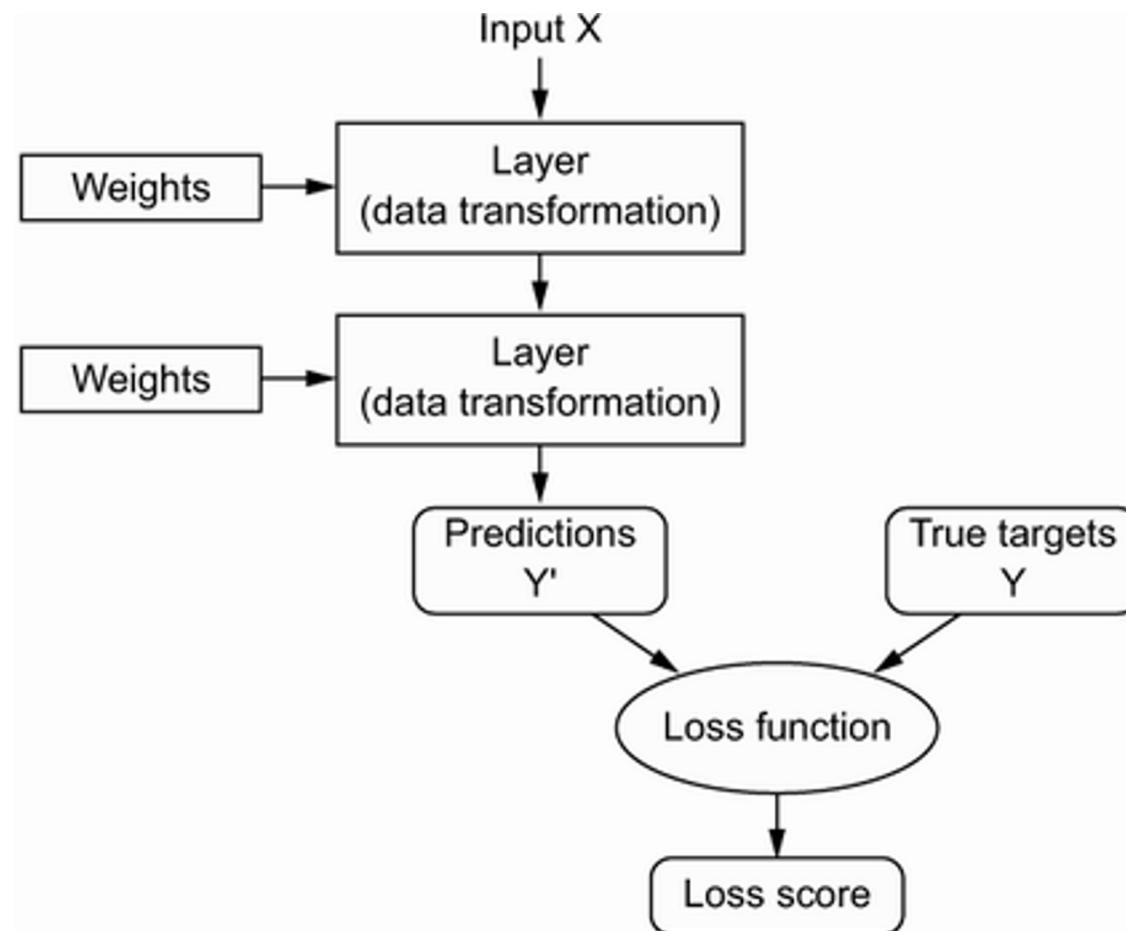


# Model Parameters

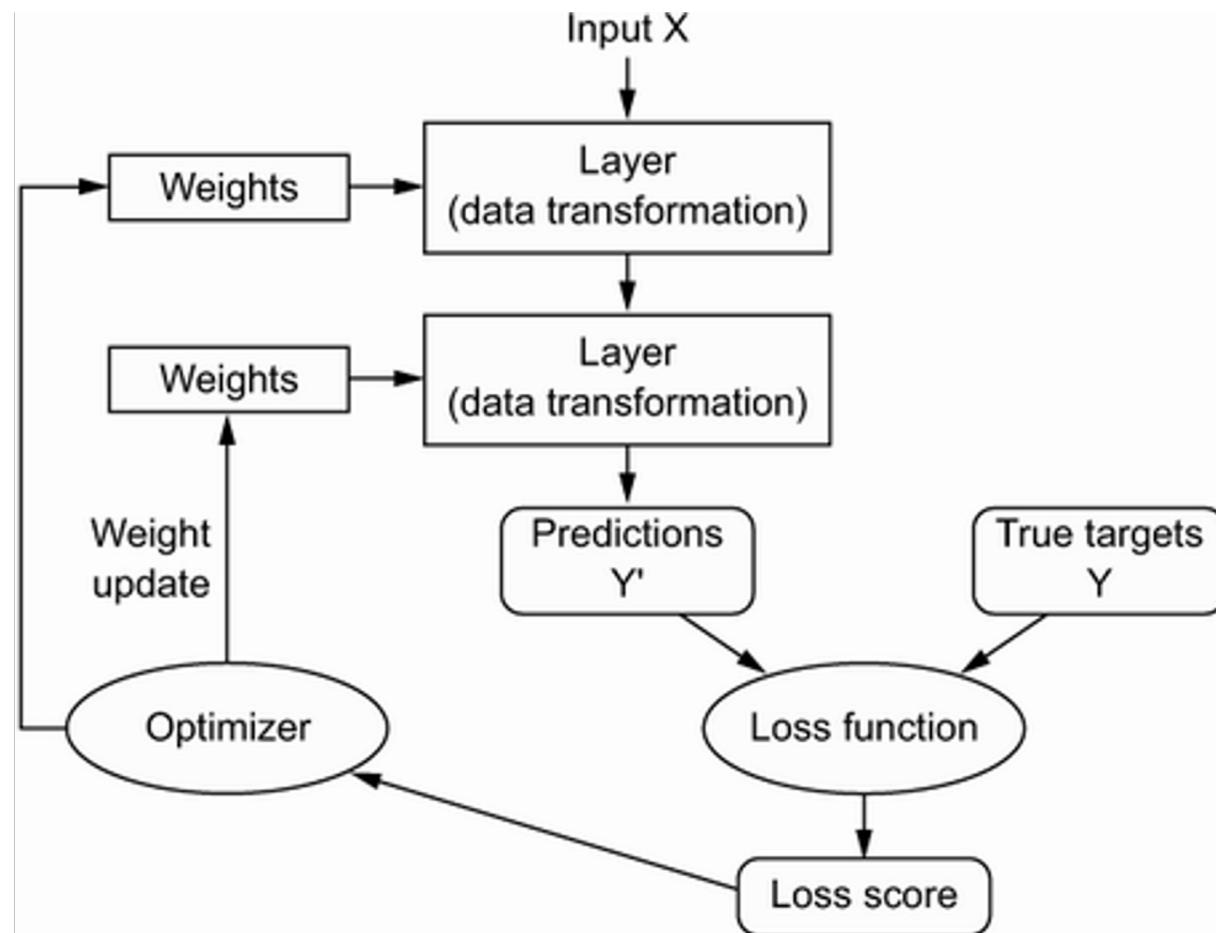
## Neuron



# Loss Function (Error)



# Optimization



# Questions?

# First Example

