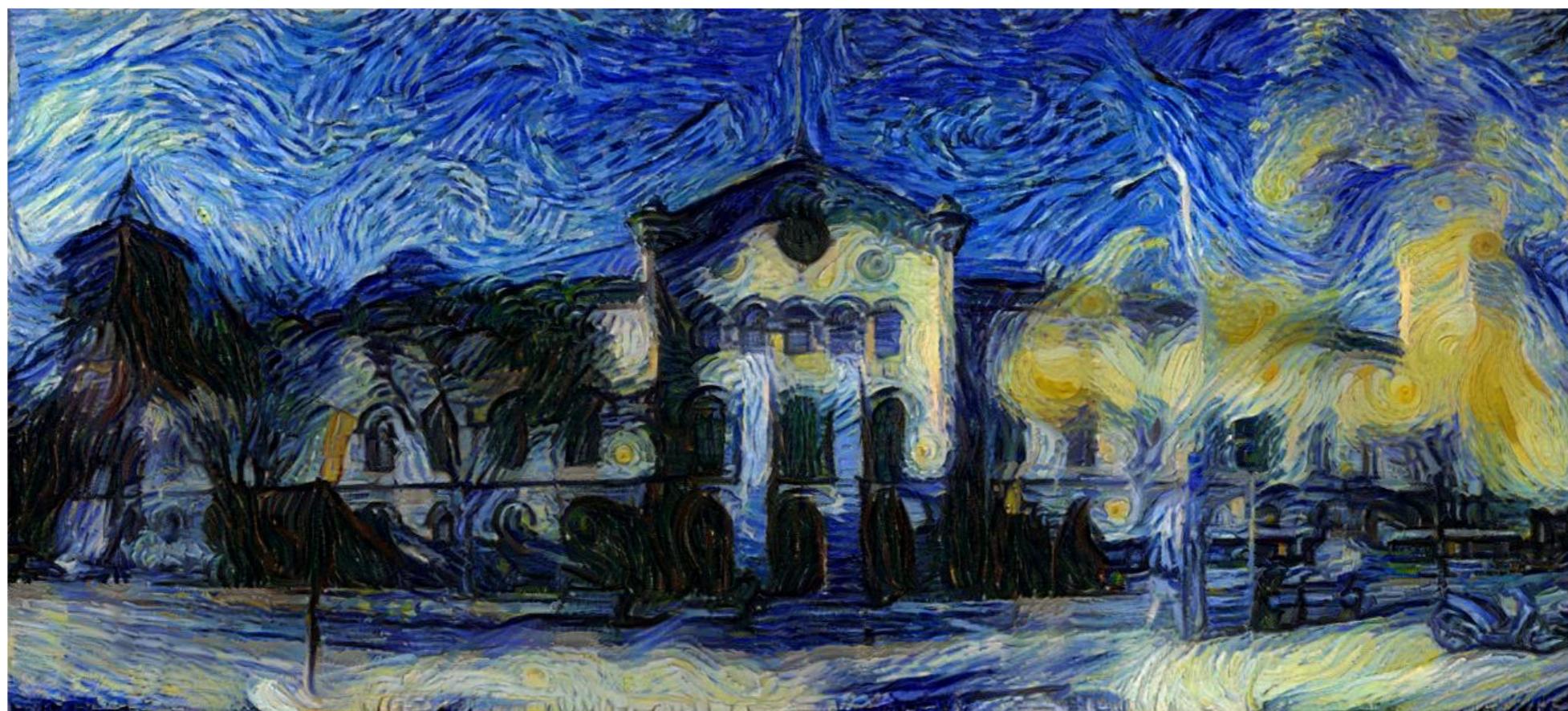




UNIVERSITAT DE  
BARCELONA



# Deep Learning

Jordi Vitrià  
[jordi.vitria@ub.edu](mailto:jordi.vitria@ub.edu)

# Course Contents

1. Describe how a (deep) neural network works and combine different types of layers and activation functions. Deep Learning is not magic.
2. Describe how these models can be applied in computer vision, text analytics, time series analysis, etc. Deep Learning is not the final machine learning method.
3. Develop your own models in **Tensorflow, Keras** and derivates. You can train (small) deep models in your laptop.

# BackEnds

 TensorFlow

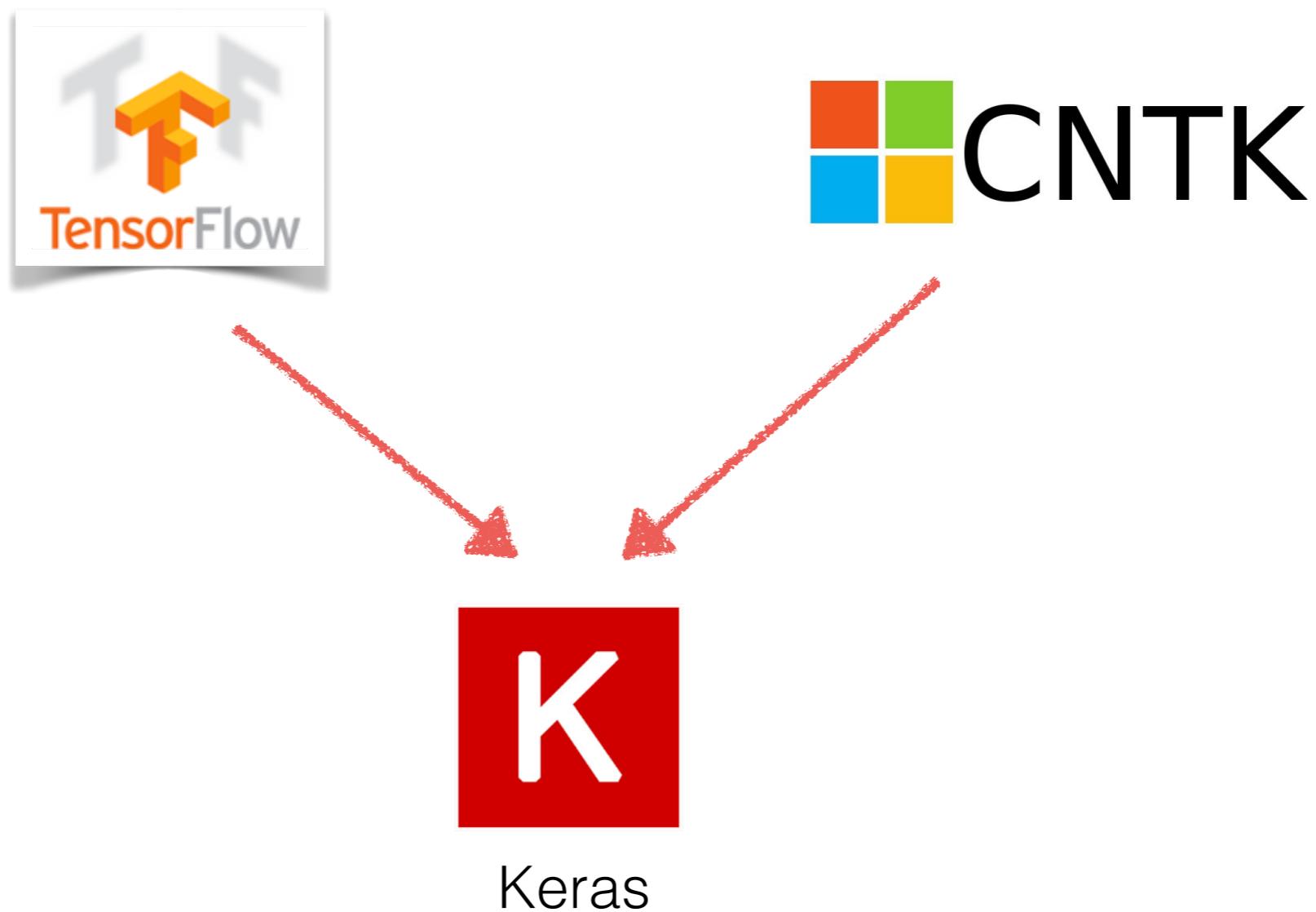
 PyTorch

 mxnet

 Caffe2

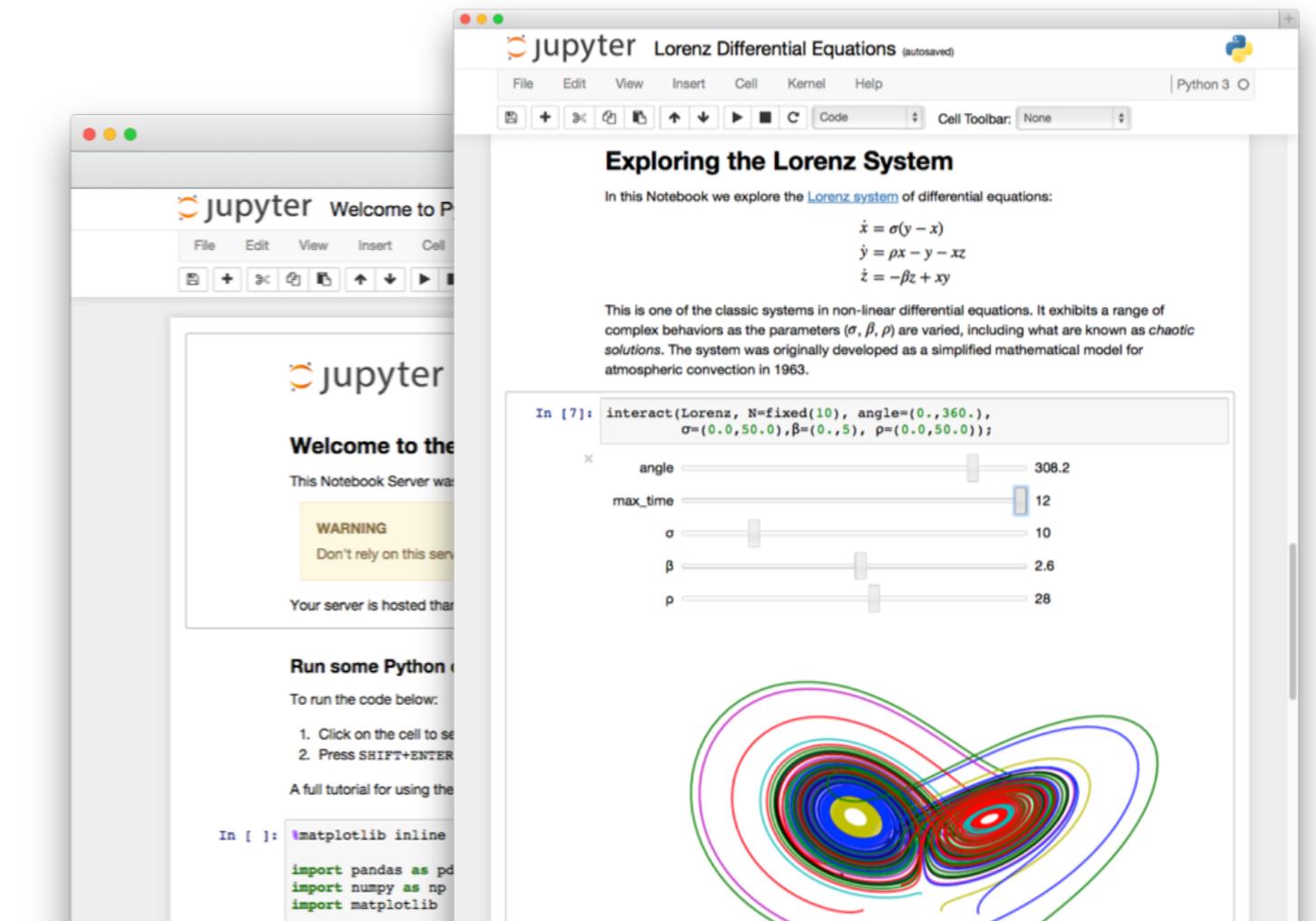
 Chainer

# Wrappers



# Approach

We will illustrate all contents with **Jupyter notebooks**, a web application that allows you to create and share documents that contain live code, equations, visualizations and explanatory text.



# How to run a Jupyter Notebook?

The easiest way for a beginner to get started with Jupyter Notebooks is by installing **Anaconda**. Anaconda is the most widely used Python distribution for data science and comes pre-loaded with all the most popular libraries and tools.

The screenshot shows the official website for the Anaconda Distribution. The header features the Anaconda logo and navigation links for Products, Why Anaconda?, Solutions, Resources, Company, Download, and a search icon. The main banner is green with the title "Anaconda Distribution" and the subtitle "The World's Most Popular Python/R Data Science Platform". A prominent "Download" button is visible. Below the banner, a paragraph explains the benefits of the Anaconda Distribution, mentioning its use for Python/R data science, machine learning, and deep learning. A bulleted list details the capabilities, including support for Jupyter, Conda, scikit-learn, TensorFlow, Theano, and various visualization and performance libraries like Pandas, NumPy, and Matplotlib. A grid of logos for these and other packages is shown. At the bottom, compatibility icons for Windows, macOS, and Linux are present.

The open-source [Anaconda Distribution](#) is the easiest way to perform Python/R data science and machine learning on Linux, Windows, and Mac OS X. With over 11 million users worldwide, it is the industry standard for developing, testing, and training on a single machine, enabling *individual data scientists* to:

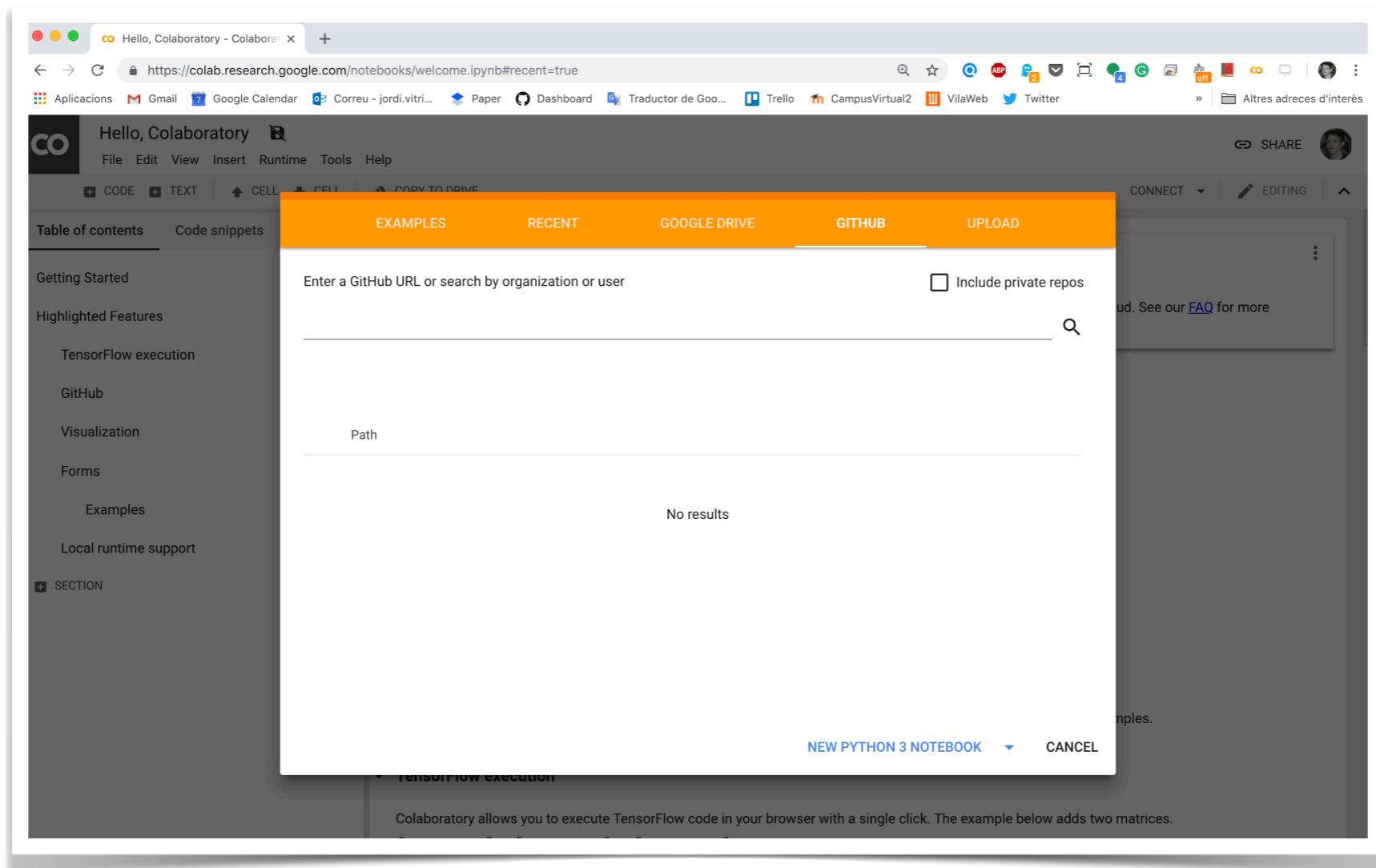
- Quickly download 1,500+ Python/R data science packages
- Manage libraries, dependencies, and environments with [Conda](#)
- Develop and train machine learning and deep learning models with [scikit-learn](#), [TensorFlow](#), and [Theano](#)
- Analyze data with scalability and performance with [Dask](#), [NumPy](#), [pandas](#), and [Numba](#)
- Visualize results with [Matplotlib](#), [Bokeh](#), [Databricks](#), and [HoloViews](#)

Windows | macOS | Linux

# 2

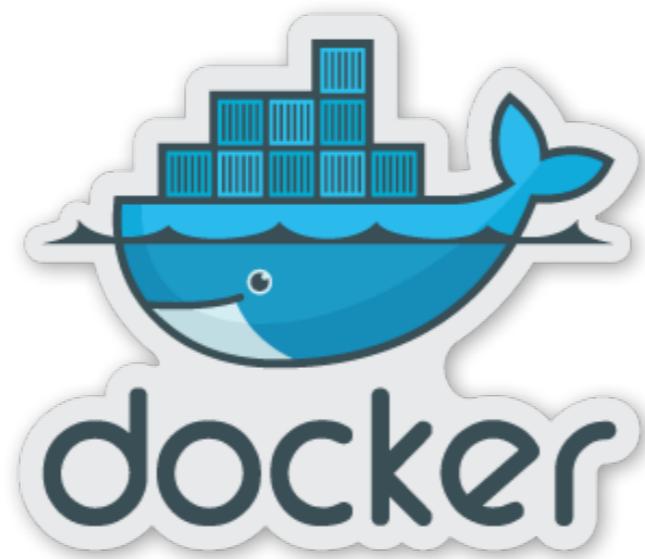
# How to run a Jupyter Notebook?

**Colaboratory** is a hosted (Google) Jupyter notebook environment that is free to use and requires no setup.



## How to run a Jupyter Notebook?

Docker provides the ability to build a runtime environment that not only remains isolated from other running containers, but also can be deployed to multiple locations in a repeatable way. Docker also uses a text document – a Dockerfile – that contains all the commands to assemble an image, which will meet our need to document the build environment.



# THE REVENANT

INSPIRED BY TRUE EVENTS  
JANUARY 8

A movie poster for "The Revenant". The background is a dark, moody landscape with a forest in the distance and a close-up of a man's face in the foreground. The man has long hair and a beard, and appears to be in a state of exhaustion or despair. The title "THE REVENANT" is at the top in large white letters, followed by the subtitle "INSPIRED BY TRUE EVENTS" and the release date "JANUARY 8".

# Why Deep Learning?



**It's funny!**

**It's not rocket  
science!**

**It's powerful!**



- In 1943, neurophysiologist **Warren McCulloch** and mathematician **Walter Pitts** wrote a paper on how neurons might work. In order to describe how neurons in the brain might work, they modeled a simple neural network using electrical circuits.
- In 1949, Donald **Hebb** wrote *The Organization of Behavior*, a work which pointed out the fact that neural pathways are strengthened each time they are used, a concept fundamentally essential to the ways in which humans learn. If two nerves fire at the same time, he argued, the connection between them is enhanced.
- In 1957 **Frank Rosenblatt** attempted to build a kind of mechanical brain called the Perceptron, which was billed as "a machine which senses, recognizes, remembers, and responds like the human mind".

- In 1962, **Widrow & Hoff** developed a learning procedure that examines the value before the weight adjusts it (i.e. 0 or 1) according to the rule: Weight Change = (Pre-Weight line value) \* (Error / (Number of Inputs)). It is based on the idea that while one active perceptron may have a big error, one can adjust the weight values to distribute it across the network, or at least to adjacent perceptrons.
- A critical book written in 1969 by **Marvin Minsky** and his collaborator **Seymour Papert** ("Perceptrons") showed that Rosenblatt's original system was painfully limited, literally blind to some simple logical functions like "exclusive-or". What had become known as the field of "neural networks" all but disappeared.



# First neural network winter is coming





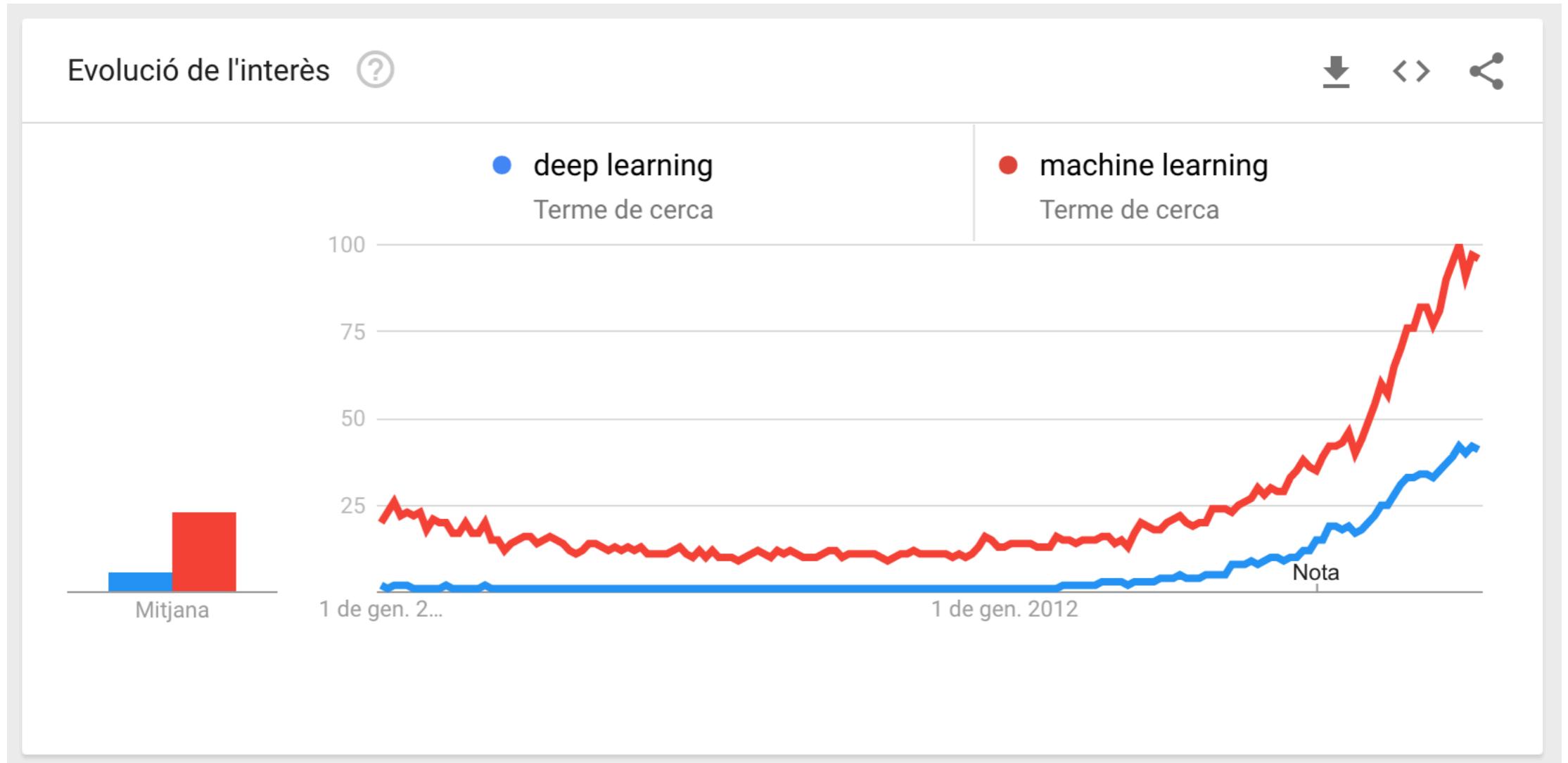
- In 1982, interest in the field was renewed. **John Hopfield** of Caltech presented a paper to the National Academy of Sciences. His approach was to create more useful machines by using bidirectional lines. Previously, the connections between neurons was only one way.
- In 1986, the problem was how to extend the Widrow-Hoff rule to multiple layers. Three independent groups of researchers, which included **David E. Rumelhart**, **Geoffrey E. Hinton** and **Ronald J. Williams**, came up with similar ideas which are now called back-propagation networks because it distributes pattern recognition errors throughout the network.
- From 1986 to mid 90's new developments arised: convolutional neural networks (**Y.LeCun**), unsupervised learning (**Y.Bengio**), RBM (**G.Hinton**), etc. But, by this point **new machine learning methods** had begun to also emerge, and people were again beginning to be skeptical of neural nets since they seemed so intuition-based and since computers were still barely able to meet their computational needs.

# Second neural network winter is coming



- With the ascent of **Support Vector Machines** and the failure of backpropagation, the early 2000s were a dark time for neural net research.
- Then, what every researcher must dream of actually happened: G.Hinton, S.Osindero, and Y.W.Teh published a paper in 2006 that was seen as a breakthrough, a breakthrough significant enough to rekindle interest in neural nets: *A fast learning algorithm for **deep belief nets**.*
- After that, following Moore's law, computers got dozens of times faster (GPUs) since the slow days of the 90s, making learning with large datasets and many layers much more tractable.

# Neural Networks Reborn



Google Trends



Roll over image to zoom in

NVIDIA

## NVIDIA Jetson TK1 Development Kit

★★★★★ 5 customer reviews

| 24 answered questions

List Price: \$199.99

Price: **\$170.77** + \$49.57 Shipping & Import Fees

Deposit to Spain [Details](#)

You Save: **\$29.22 (15%)**

Item is eligible: No interest if paid in full within 6 months with the [Amazon.com Store Card](#).

In Stock.

This item ships to **Barcelona, Spain**. Want it **Monday, Feb.**

**12?** Order within **5 hrs 25 mins** and choose

**AmazonGlobal Priority Shipping** at checkout. [Learn more](#)

Ships from and sold by Amazon.com. Gift-wrap available.

- NVIDIA Kepler GPU with 192 CUDA cores
- NVIDIA 4-Plus-1 quad-core ARM Cortex-A15 CPU
- 2 GB memory, 16 GB eMMC
- Gigabit Ethernet, USB 3.0, SD/MMC, miniPCIe
- HDMI 1.4, SATA, Line out/Mic in, RS232 serial port
- Expansion ports for additional display, GPIOs, and high-bandwidth camera interface

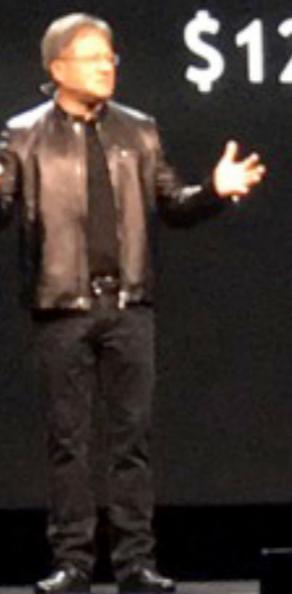


# NVIDIA DGX-1

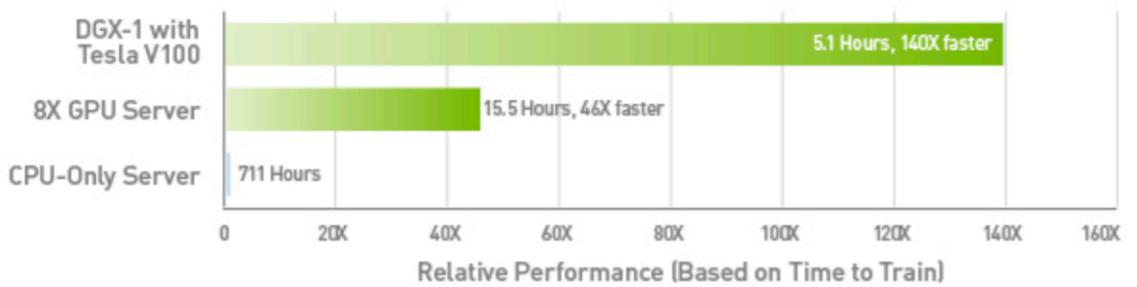
## WORLD'S FIRST DEEP LEARNING SUPERCOMPUTE

170TF | “250 servers in-a-box” | [nvidia.com/dgx](http://nvidia.com/dgx)

**\$129,000**



### NVIDIA DGX-1 Delivers 140X Faster Deep Learning Training



Workload: ResNet-50, 90 epochs to solution | CPU Server: Dual Xeon E5-2699v4, 2.6GHz

System Level Specifications	NVIDIA DGX-1 with Tesla V100
TFLOPS (deep learning)	1000
CUDA Cores	40,960
Tensor Cores	5,120
NVLink vs PCIe Speed-up	10X
Deep Learning Training Speed-up	3X

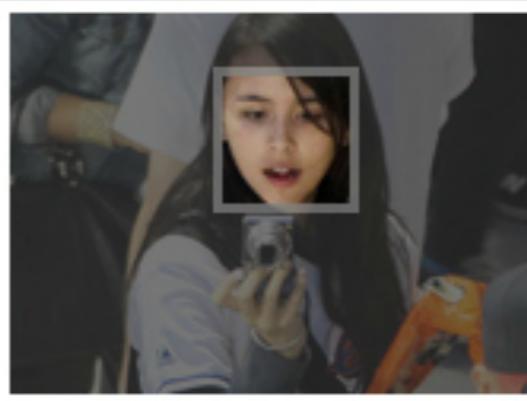
## Definitions

- **Neural Networks (NN)** is a beautiful biologically-inspired programming paradigm which enables a computer to learn from observational data.
- **Deep Learning (DL)** is a powerful set of techniques for learning in neural networks.
- NN and DL currently provide the best solutions to many problems in image recognition, speech recognition, and natural language processing.

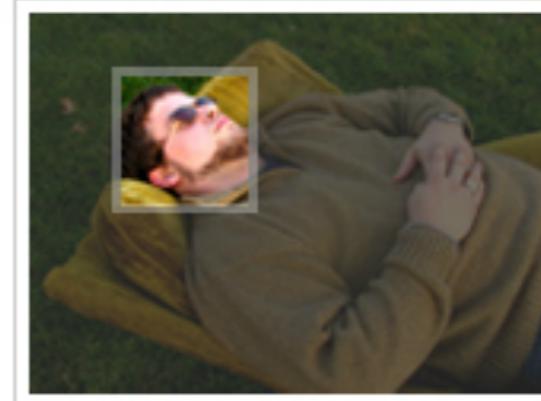
# “Classical” applications: object classification, detection and segmentation.



# Face recognition.



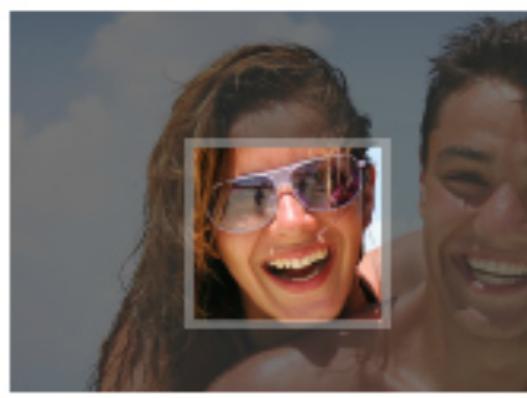
Who is this?



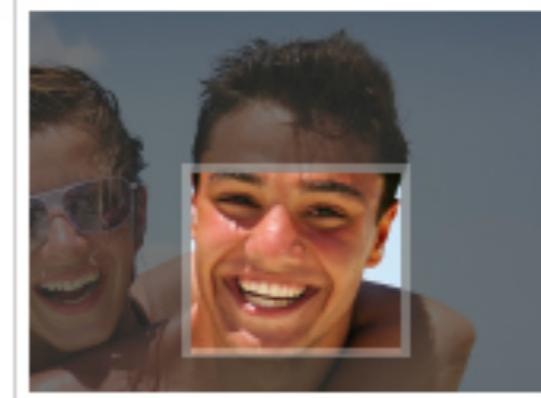
Who is this?



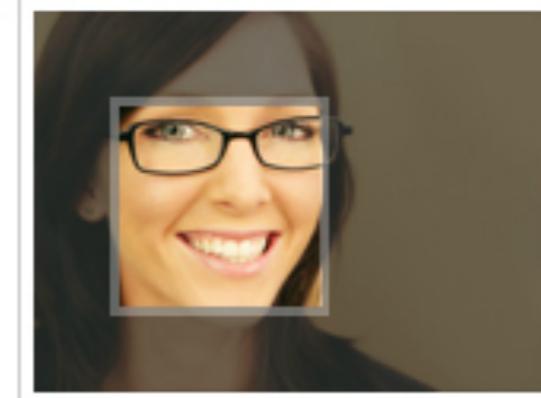
Who is this?



Who is this?



Who is this?



Who is this?

DeepFace (Facebook): Accuracy of 97.35% (2014)

FaceNet (Google): Accuracy of 99.63% (2015)

...

# New applications: navigation and mapping.

The screenshot shows the official Dyson website. At the top, there is a navigation bar with the Dyson logo, followed by links to Tienda, Aspiradoras, Ventiladores y Calefactores, Airblade™, Mi cuenta, Soporte, and a globe icon. Below the navigation bar, the text "Robot Dyson 360 Eye™" is displayed, along with a yellow button that says "Sea el primero en disfrutarlo". The main content area features a large image of the Dyson 360 Eye robot, which is a cylindrical device with transparent panels revealing its internal components like the motor and sensors. To the left of the robot, there is a blue circular callout containing the text "Vea a James Dyson presentando el nuevo Dyson 360 Eye™ en Tokio" and a small video thumbnail showing a presentation stage.

# New applications: Image Upscaling (Flipboard)



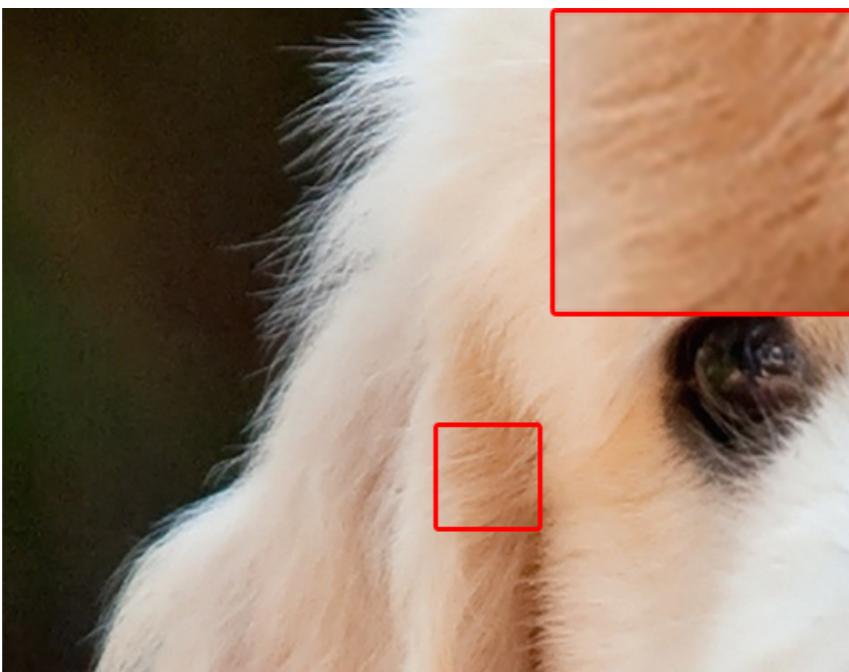
<http://engineering.flipboard.com/2015/05/scaling-convnets/>

# New applications: Image Upscaling (Flipboard)

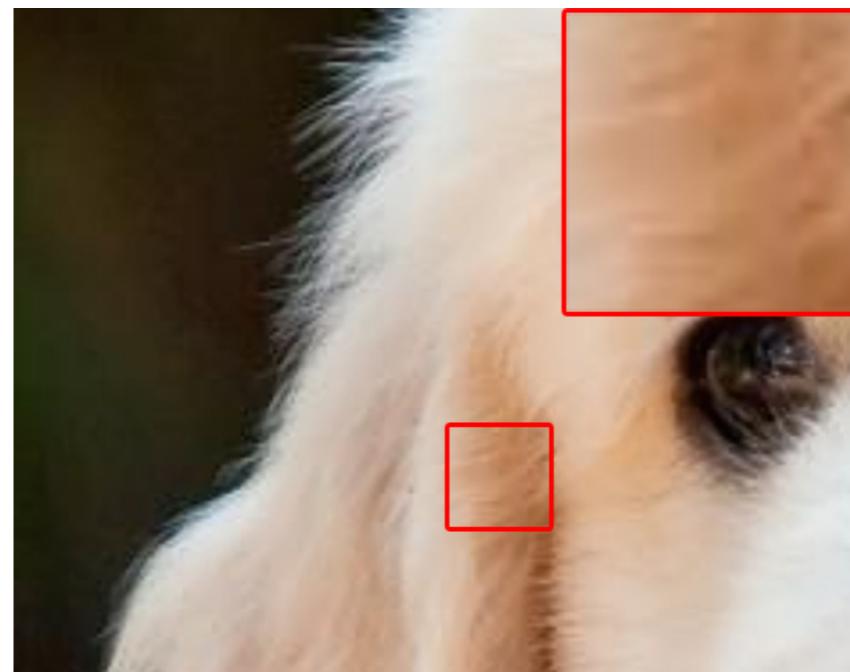


<http://engineering.flipboard.com/2015/05/scaling-convnets/>

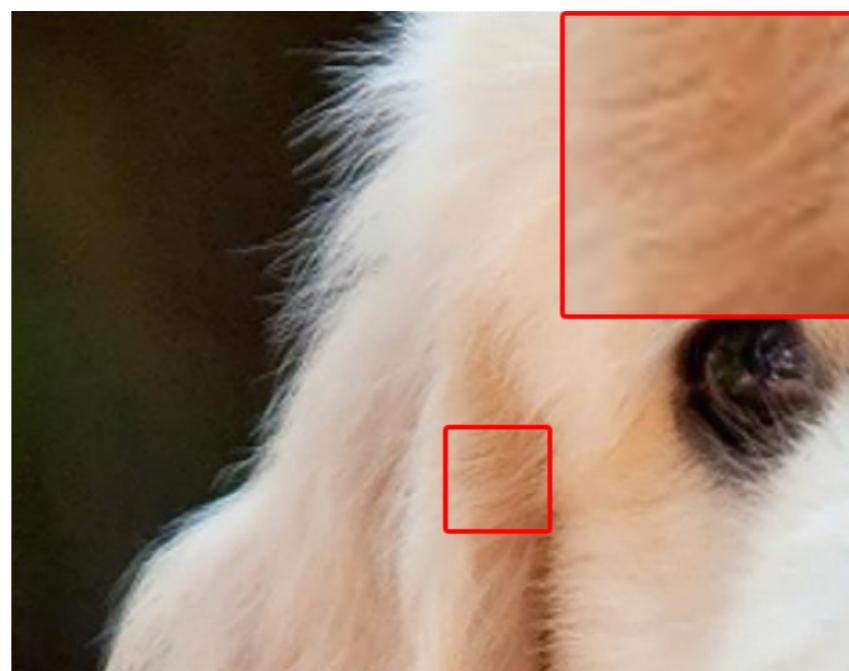
# New applications: Image Upscaling (Flipboard)



Original



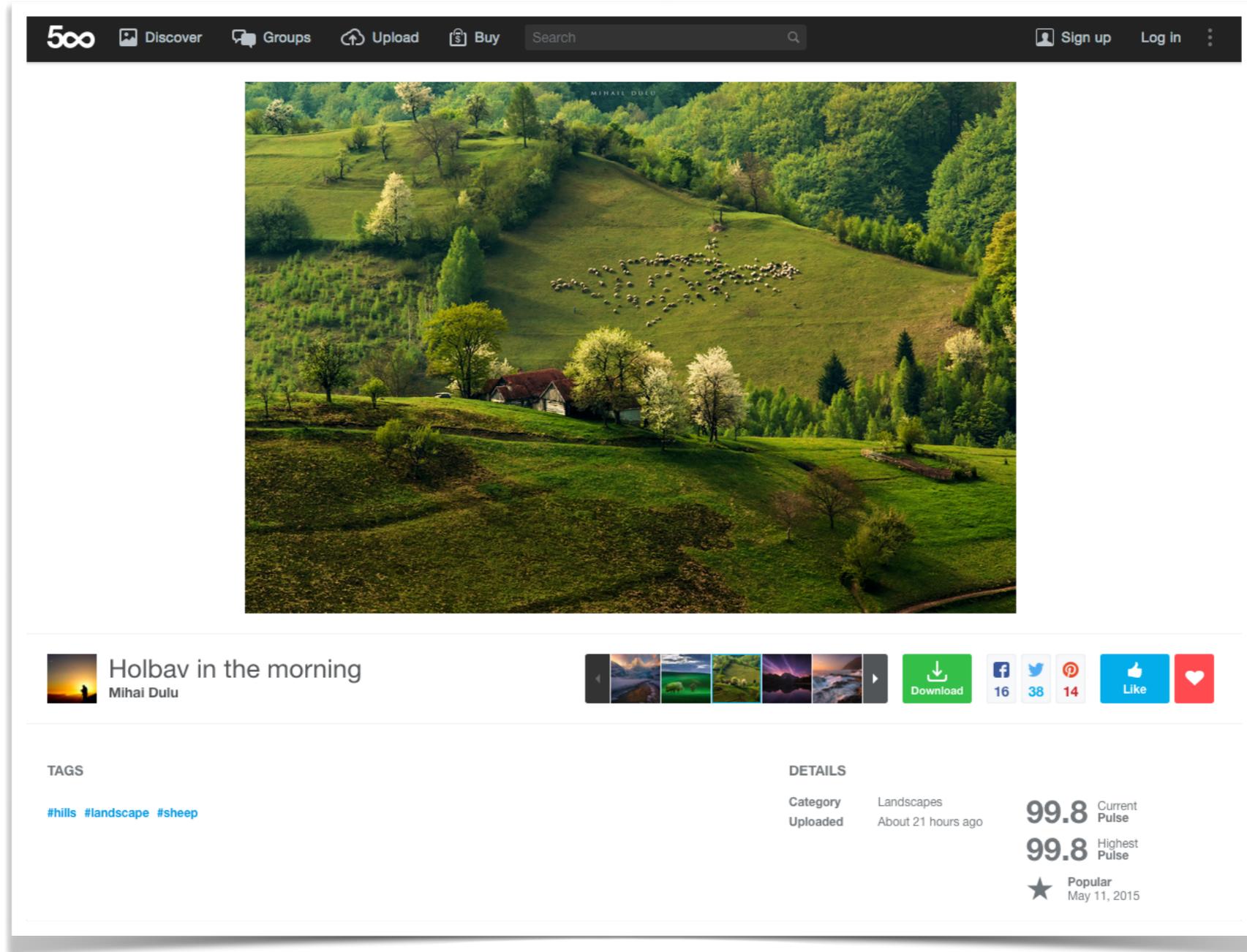
Bicubic



Model

<http://engineering.flipboard.com/2015/05/scaling-convnets/>

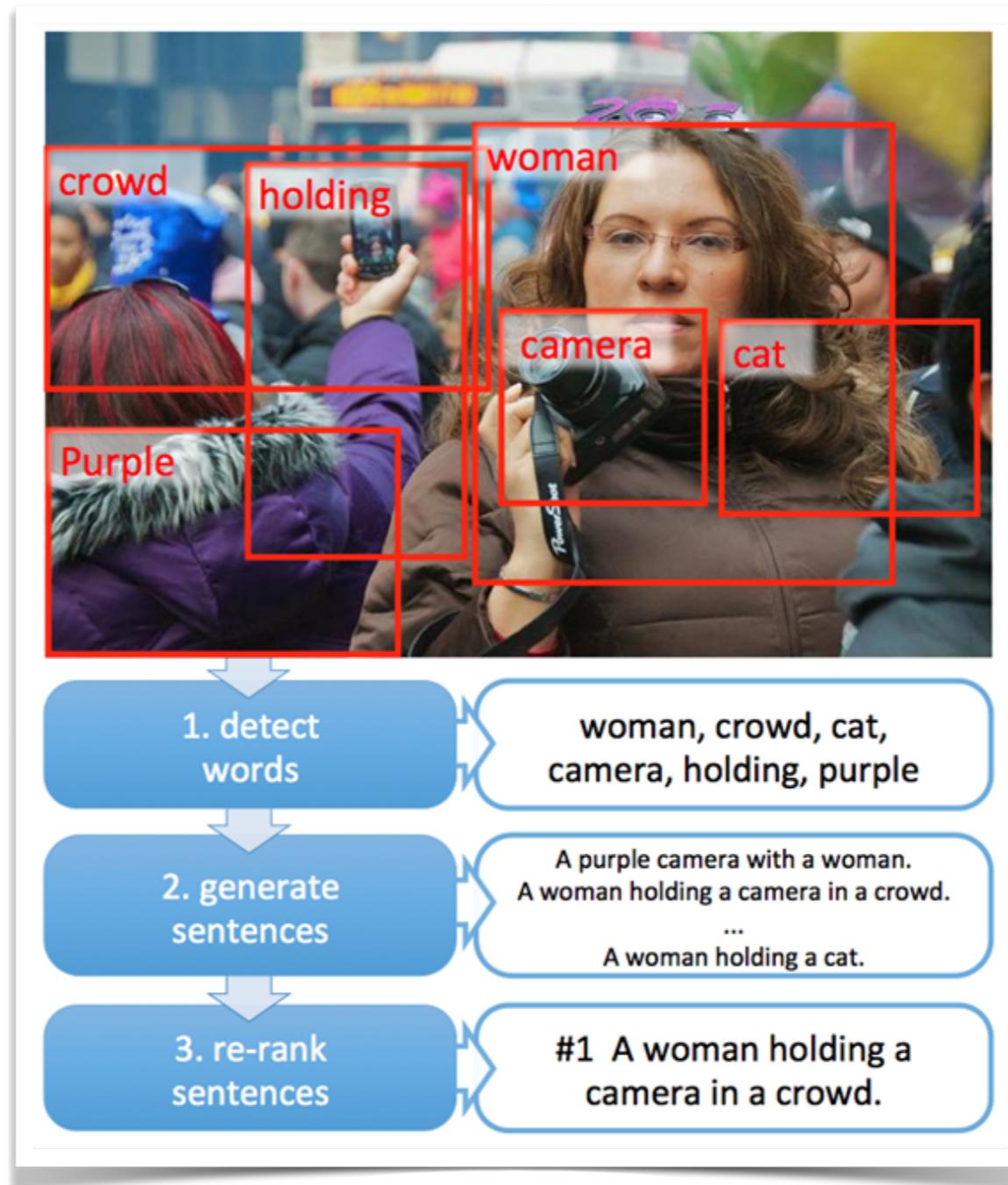
# New applications: Non visual data prediction



## What is Pulse?

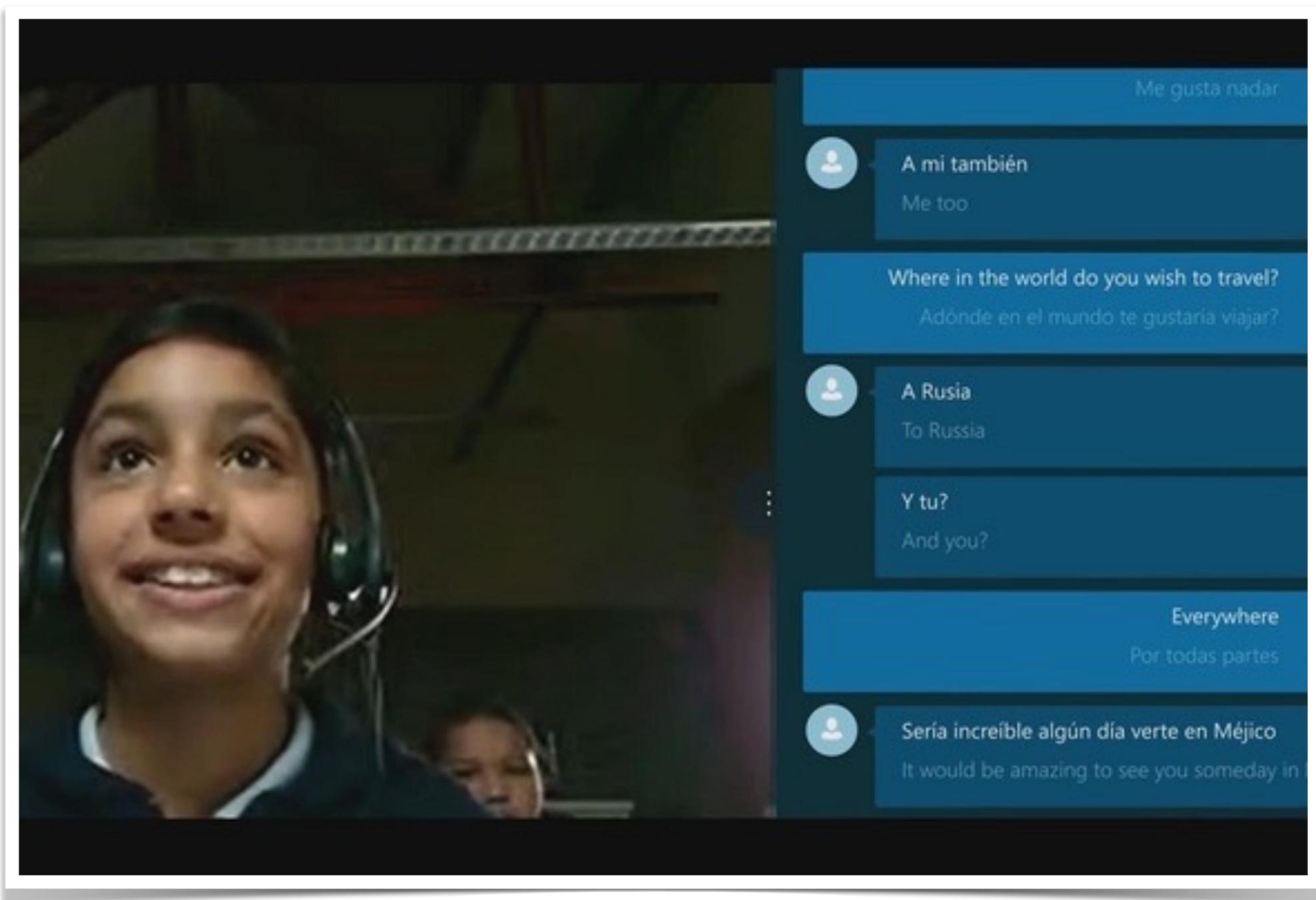
Pulse is a score out of 100 points that measures how **popular** a photo is. Pulse is calculated by an algorithm, which is unique to 500px and is based on votes (Likes & Favorites) on your photo from the community. The Pulse algorithm was designed to promote daily exposure of new photographs and photographers. It is not necessarily a measure of photograph's quality.

# New applications: Automatic Image Captioning



<http://blogs.technet.com/b/machinelearning/archive/2014/11/18/rapid-progress-in-automatic-image-captioning.aspx>

# Speech translation



Me gusta nadar

A mi también

Me too

Where in the world do you wish to travel?

Adónde en el mundo te gustaría viajar?

A Rusia

To Russia

Y tu?

And you?

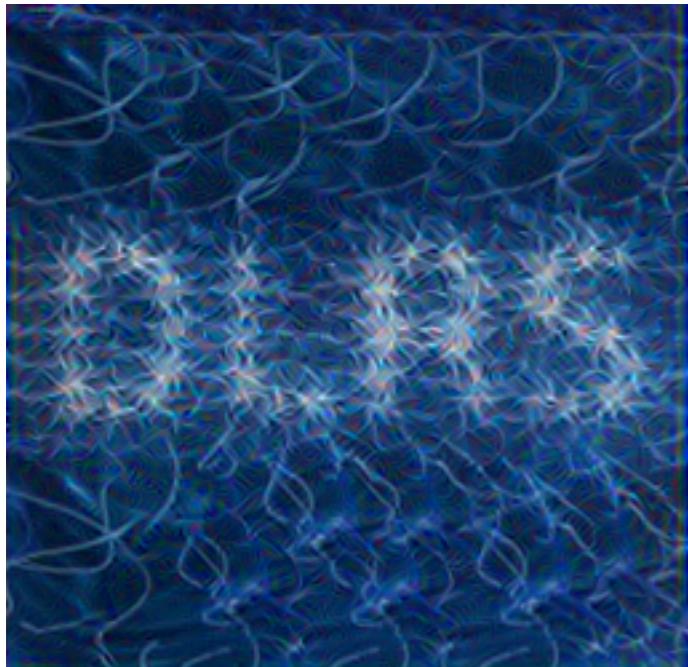
Everywhere

Por todas partes

Sería increíble algún día verte en Méjico

It would be amazing to see you someday in

# Recommenders



## 1st Workshop on Deep Learning for Recommender Systems

in conjunction with RecSys 2016  
15 September 2016, Boston, USA

# Music Generation

The screenshot shows a SoundCloud profile for an AI entity named 'deepjazz'. The profile picture is a white circle containing a stylized 'dj' logo. The bio reads: 'I'm an AI built to make Jazz' and 'Princeton, United States'. The profile has 104 followers, 1 following, and 6 tracks. It features a black and white photograph of four jazz musicians playing instruments. Below the profile, there are three tracks listed: 'deepjazz on Metheny' (14 days ago, #Electronic), 'dj 1 deepjazz On Metheny ... 1 Epoch' (6,142 plays), 'dj 2 deepjazz On Metheny ... 16 Epochs' (3,452 plays), and 'dj 3 deepjazz On Metheny ... 32 Epochs' (1,908 plays). A message from the AI states: 'Hi! I'm deepjazz, an AI built by Ji-Sung Kim. You can check out my source code on GitHub or visit my website, [deepjazz.io](http://deepjazz.io)'.

SOUNDCLOUD Charts Search for artists, bands, tracks, podcasts Sign in or Create account Upload ...

All Tracks Playlists Reposts Follow Share

deepjazz  
I'm an AI built to make Jazz  
Princeton, United States

Followers 104 | Following 1 | Tracks 6

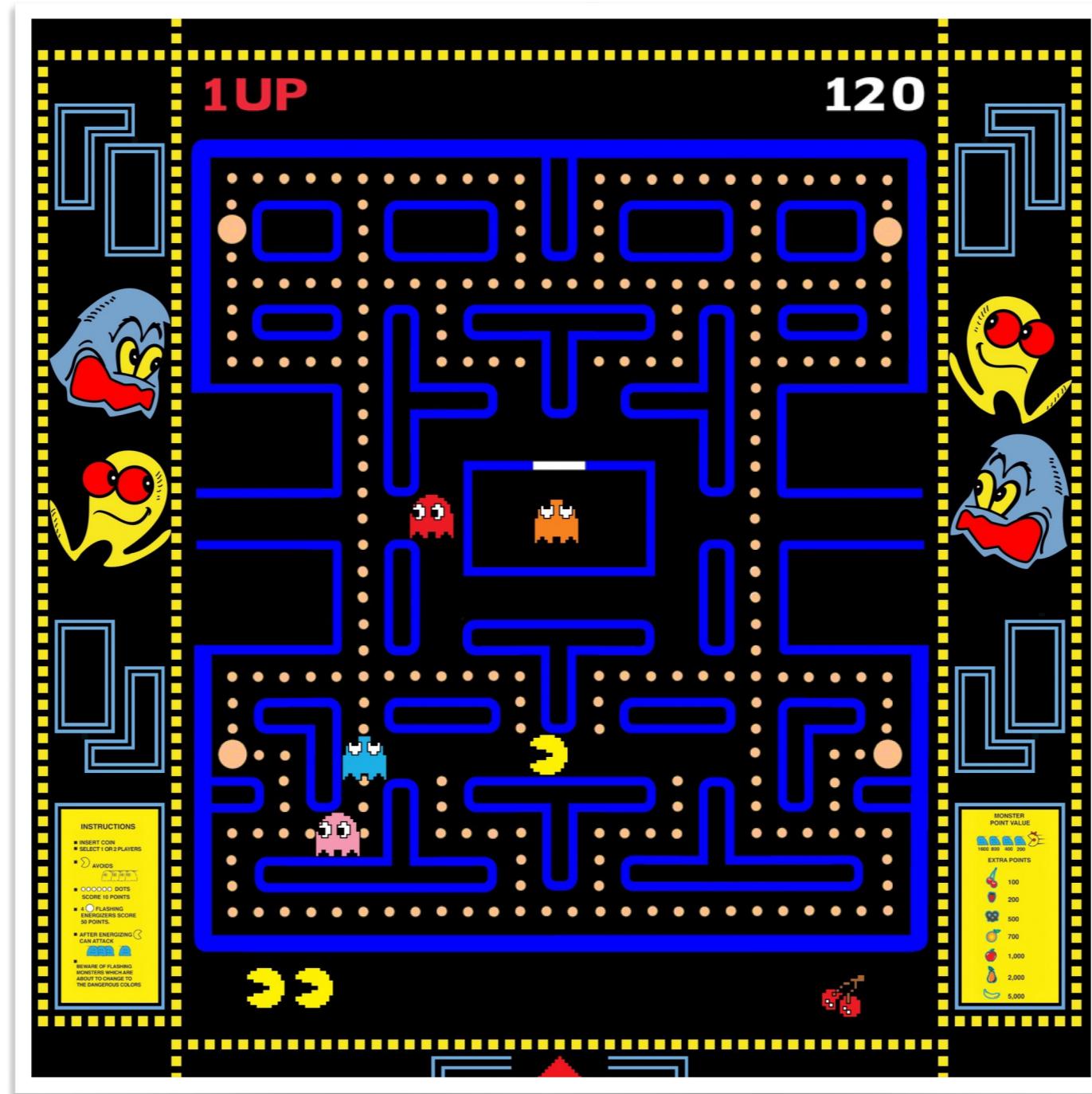
Hi! I'm deepjazz, an AI built by Ji-Sung Kim. You can check out my source code on GitHub or visit my website, [deepjazz.io](http://deepjazz.io)

my source code (GitHub)  
[deepjazz.io](http://deepjazz.io)

1 following View all

Track	Plays
deepjazz on Metheny	0:33
dj 1 deepjazz On Metheny ... 1 Epoch	6,142
dj 2 deepjazz On Metheny ... 16 Epochs	3,452
dj 3 deepjazz On Metheny ... 32 Epochs	1,908

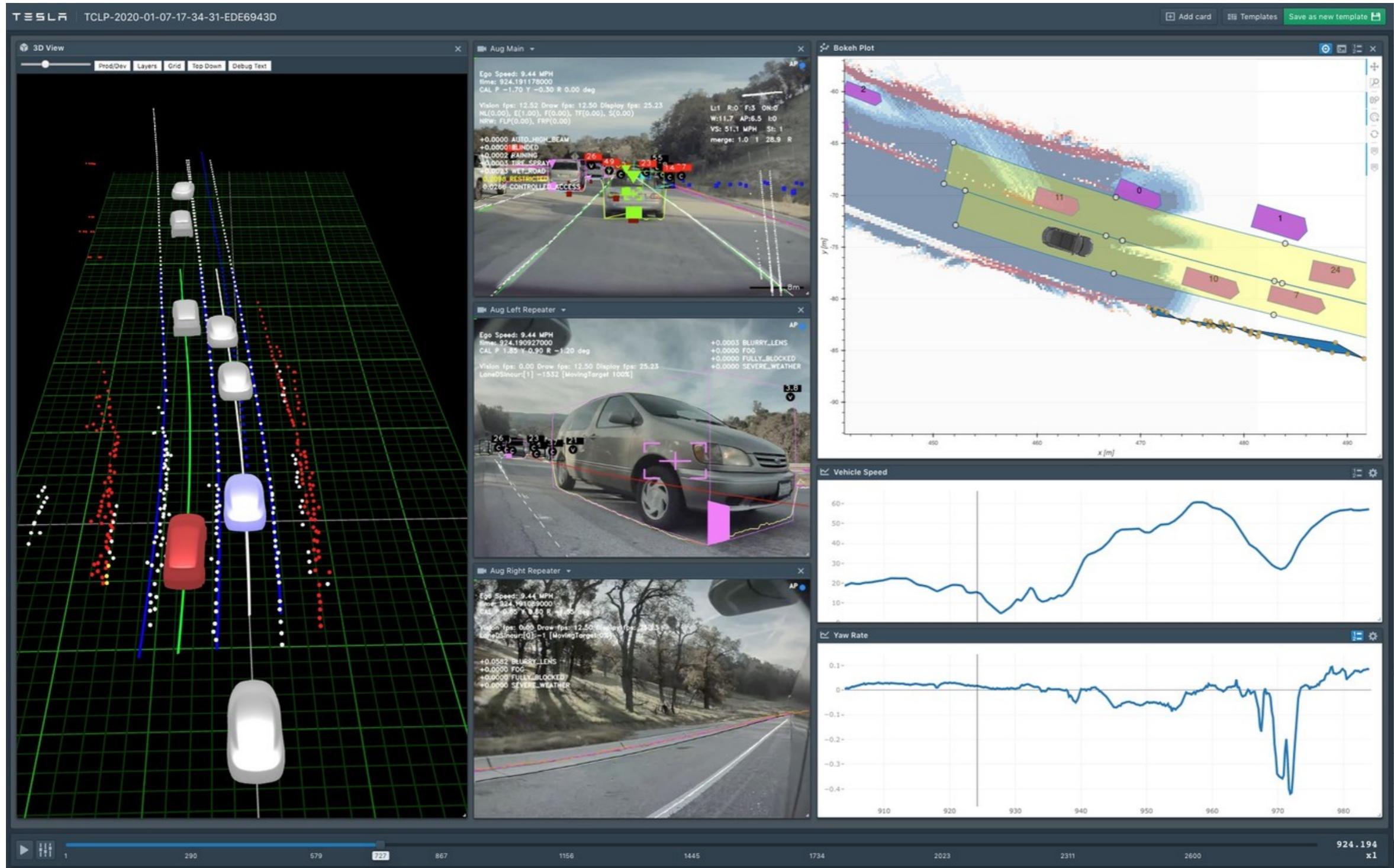
# Reinforcement learning.



Go

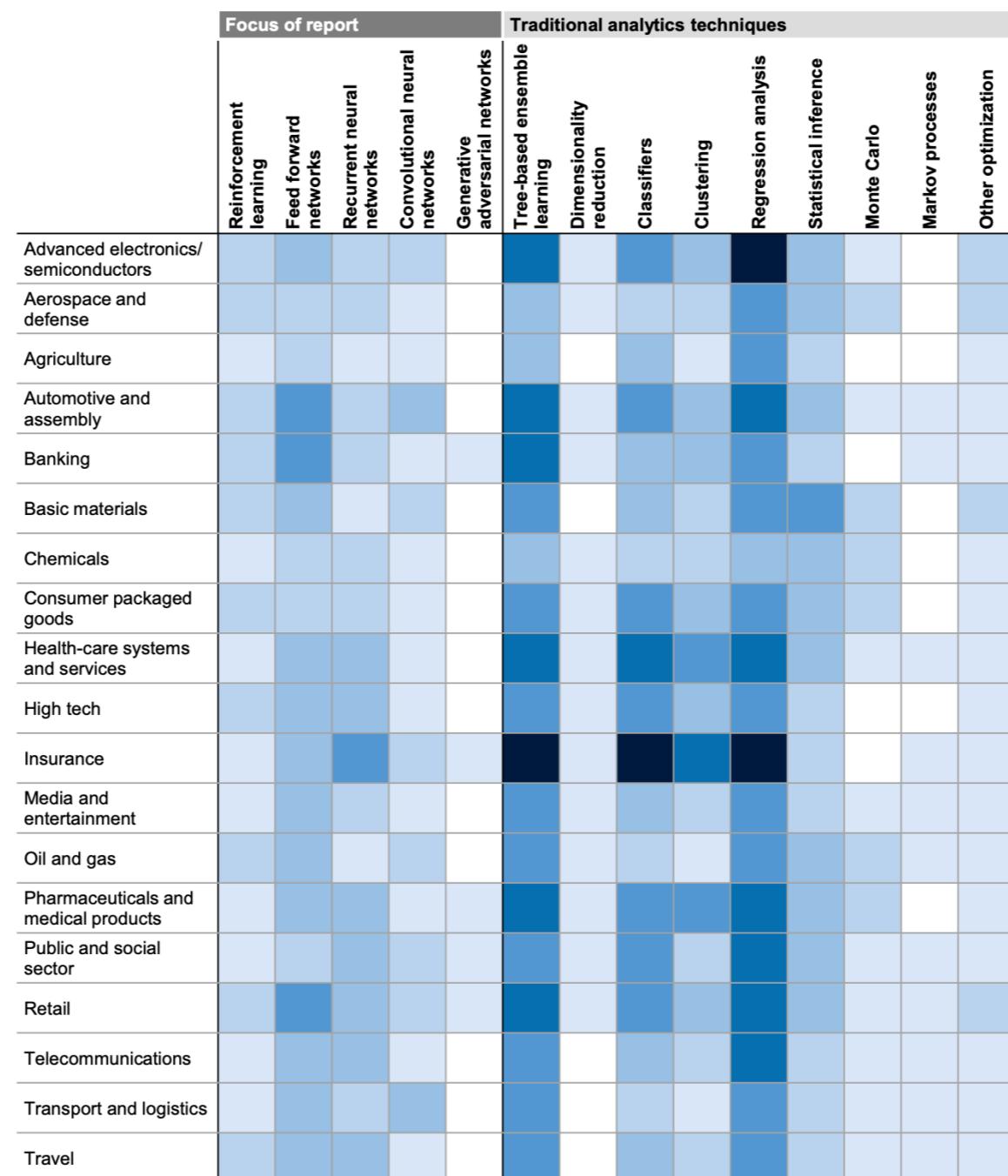


# TESLA Autopilot



[https://www.youtube.com/watch?time\\_continue=31&v=NGS2SNGXUXo&feature=emb\\_title](https://www.youtube.com/watch?time_continue=31&v=NGS2SNGXUXo&feature=emb_title)

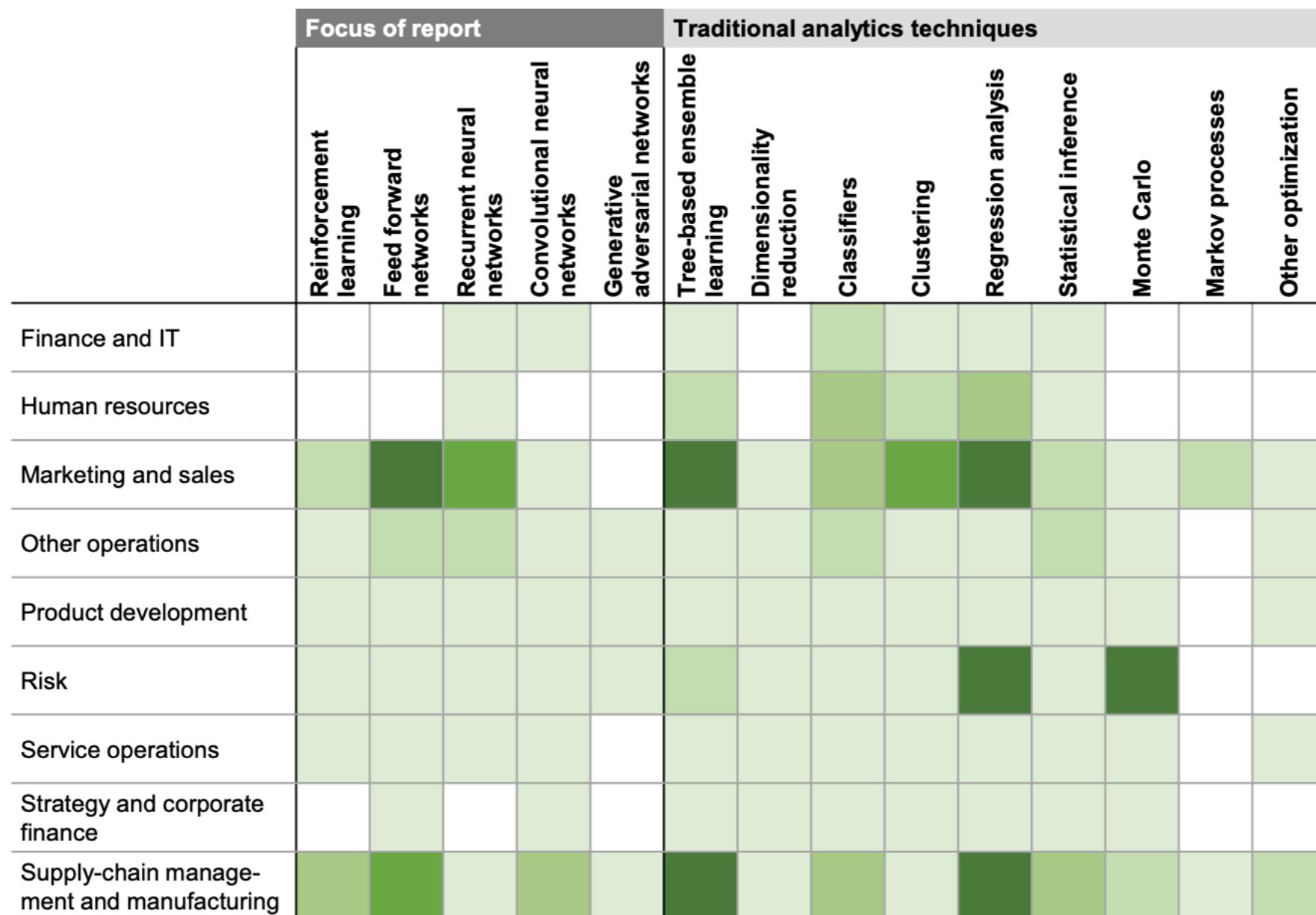
Heat map: Technique relevance to industries

Number of use cases Low  High

SOURCE: McKinsey Global Institute analysis

Heat map: Technique relevance to functions

Number of use cases    Low    High



SOURCE: McKinsey Global Institute analysis

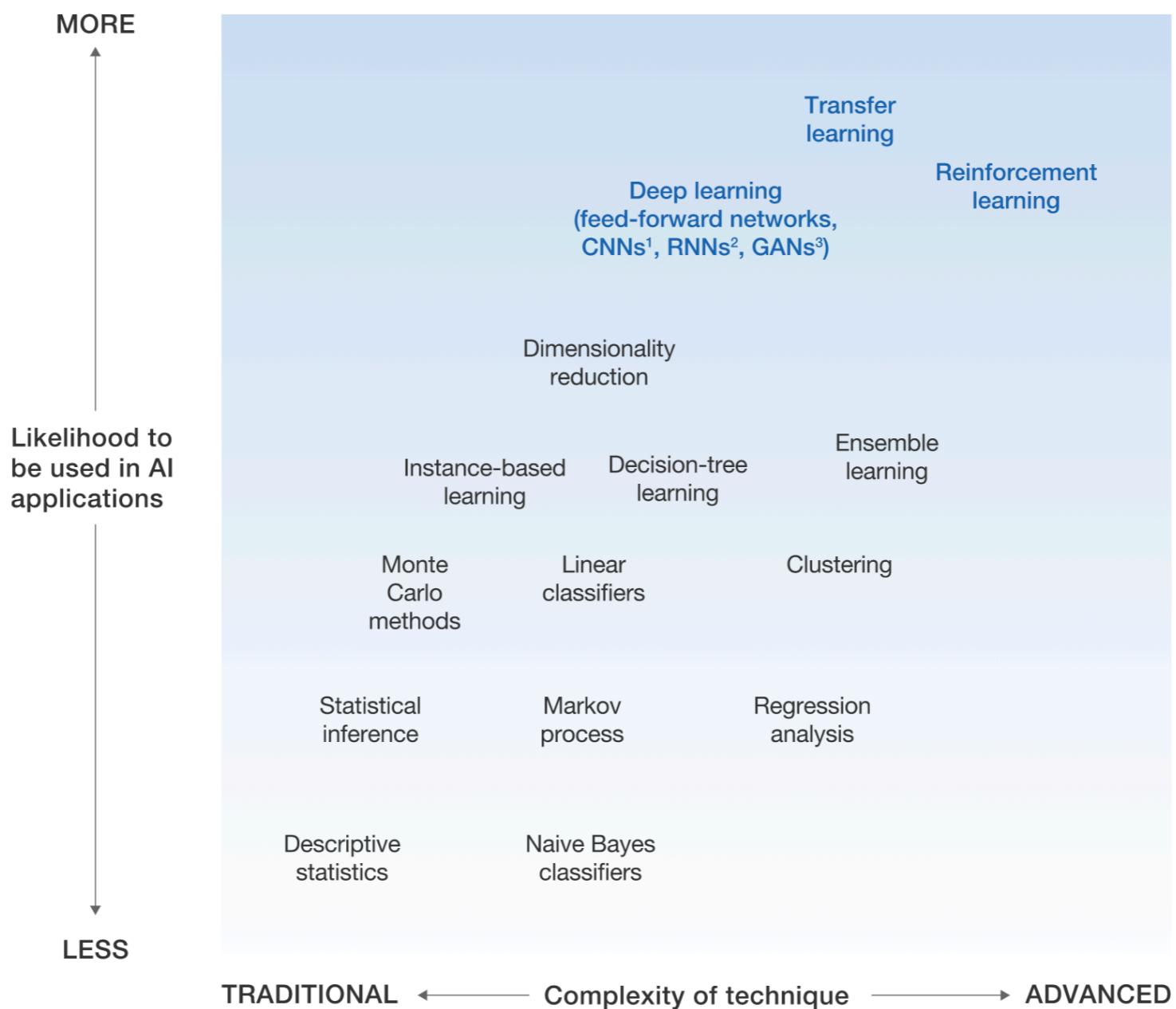
Use case mapping to data types

Number of use cases    Low    High



SOURCE: McKinsey Global Institute analysis

■ Considered AI for our research

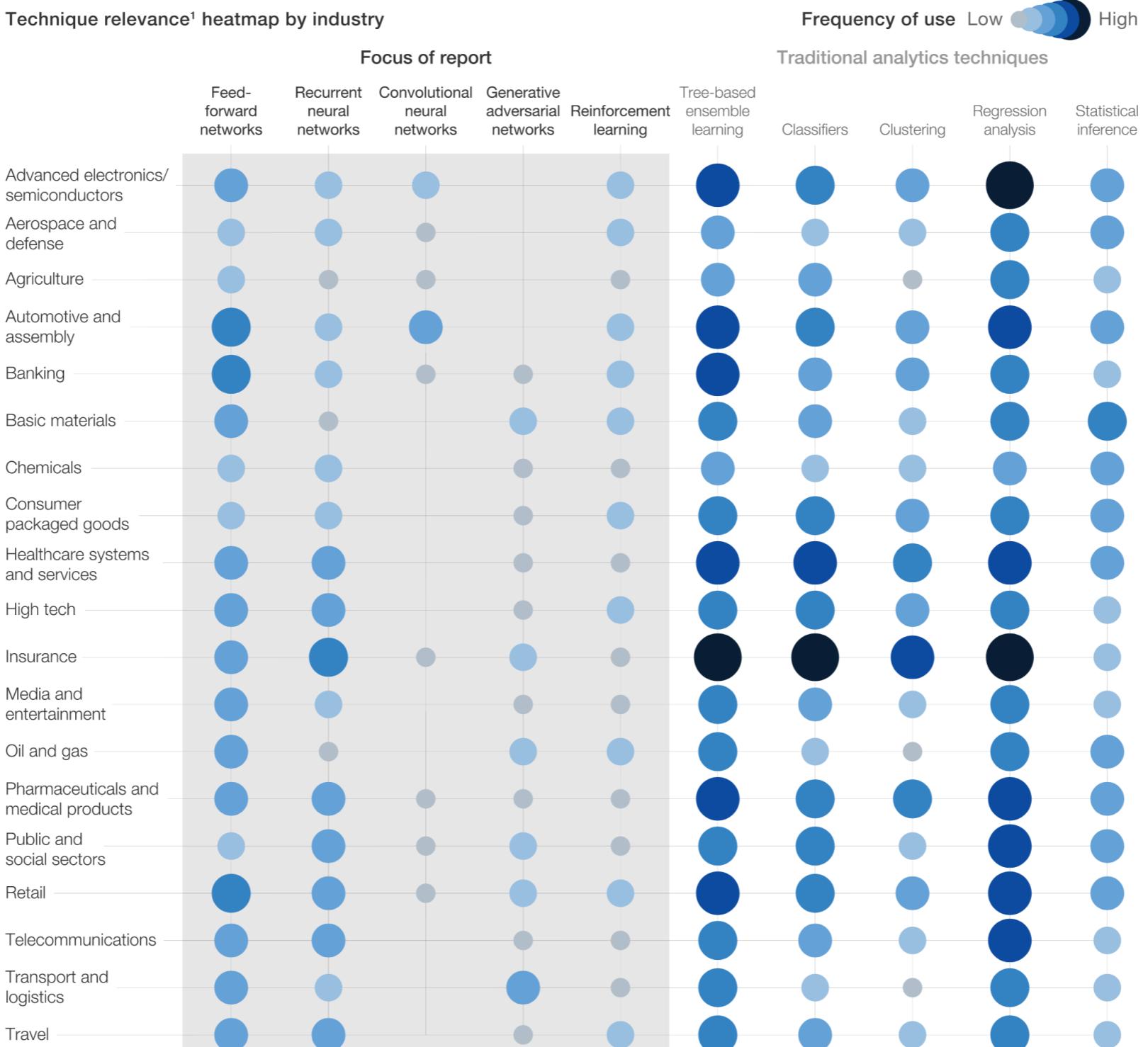


<sup>1</sup>Convolutional neural networks.

<sup>2</sup>Recurrent neural networks.

<sup>3</sup>Generative adversarial networks.

McKinsey&Company | Source: McKinsey Global Institute analysis



McKinsey&Company | Source: McKinsey Global Institute analysis

<https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-applications-and-value-of-deep-learning>

# Start Ups



## 60+ STARTUPS USING DEEP LEARNING

### CORE AI: COMPUTER VISION



### CORE AI: OTHER



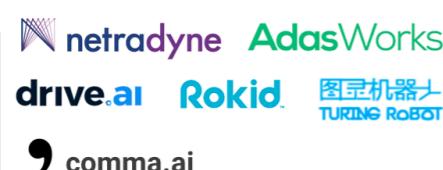
### BI, SALES & CRM



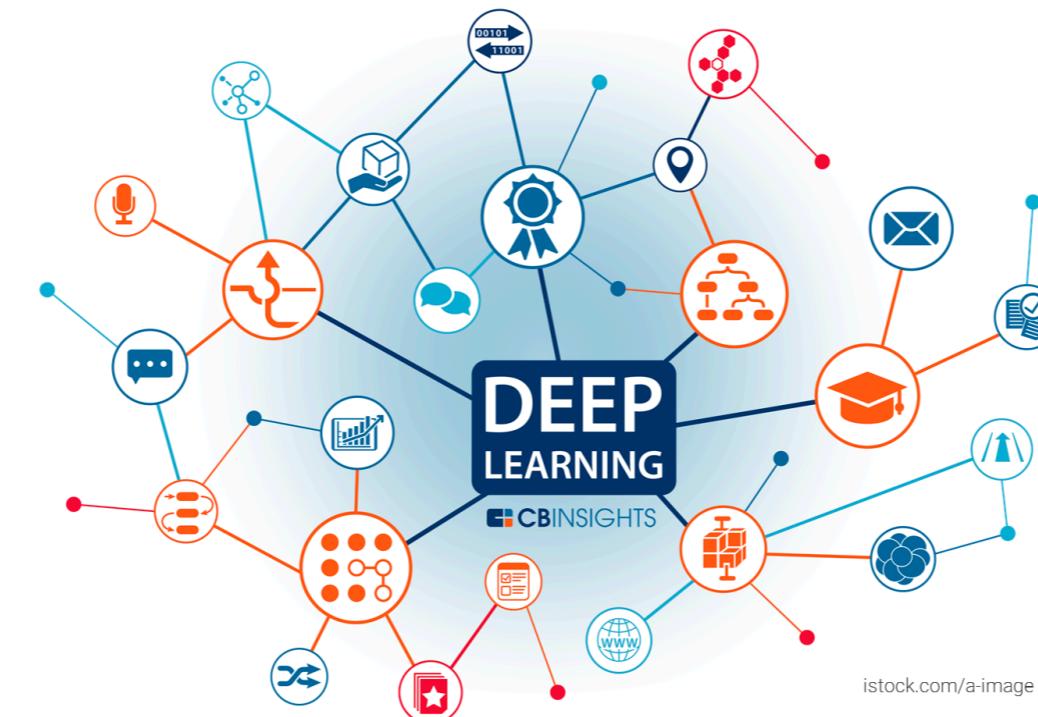
### CORE AI: VOICE INTERFACE



### ROBOTICS & AUTO



### HEALTHCARE



istock.com/a-image

### SECURITY



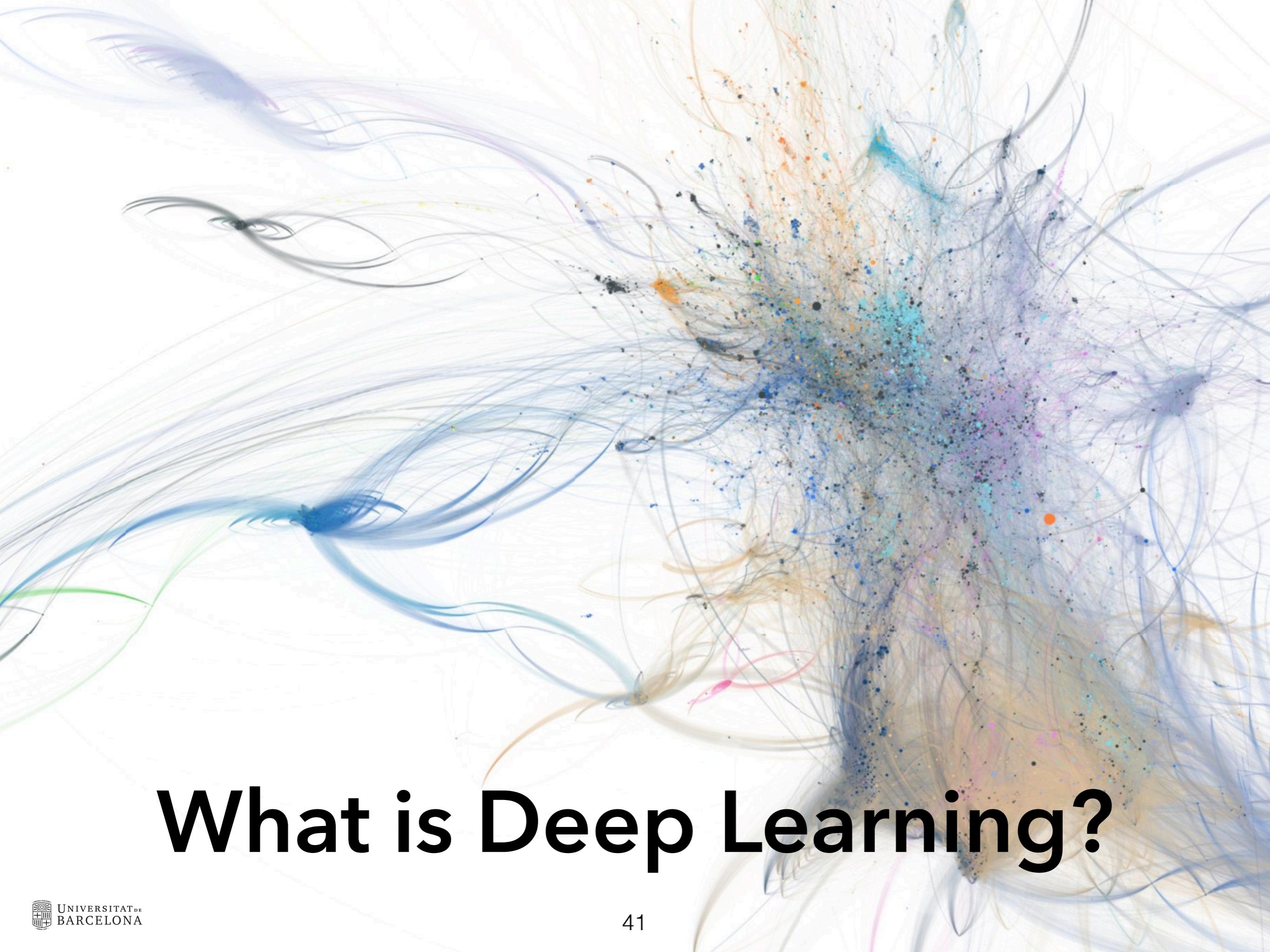
### OTHER



### ACQUIRED (2014-2016YTD)



**CB INSIGHTS**

The background of the slide features a complex, abstract network visualization. It consists of numerous thin, translucent lines of various colors, primarily shades of blue, purple, and grey, which converge and diverge from a central, dense cluster of small, multi-colored dots. This central cluster is composed of dots in shades of orange, yellow, blue, and green, creating a sense of a dynamic, interconnected system.

# What is Deep Learning?

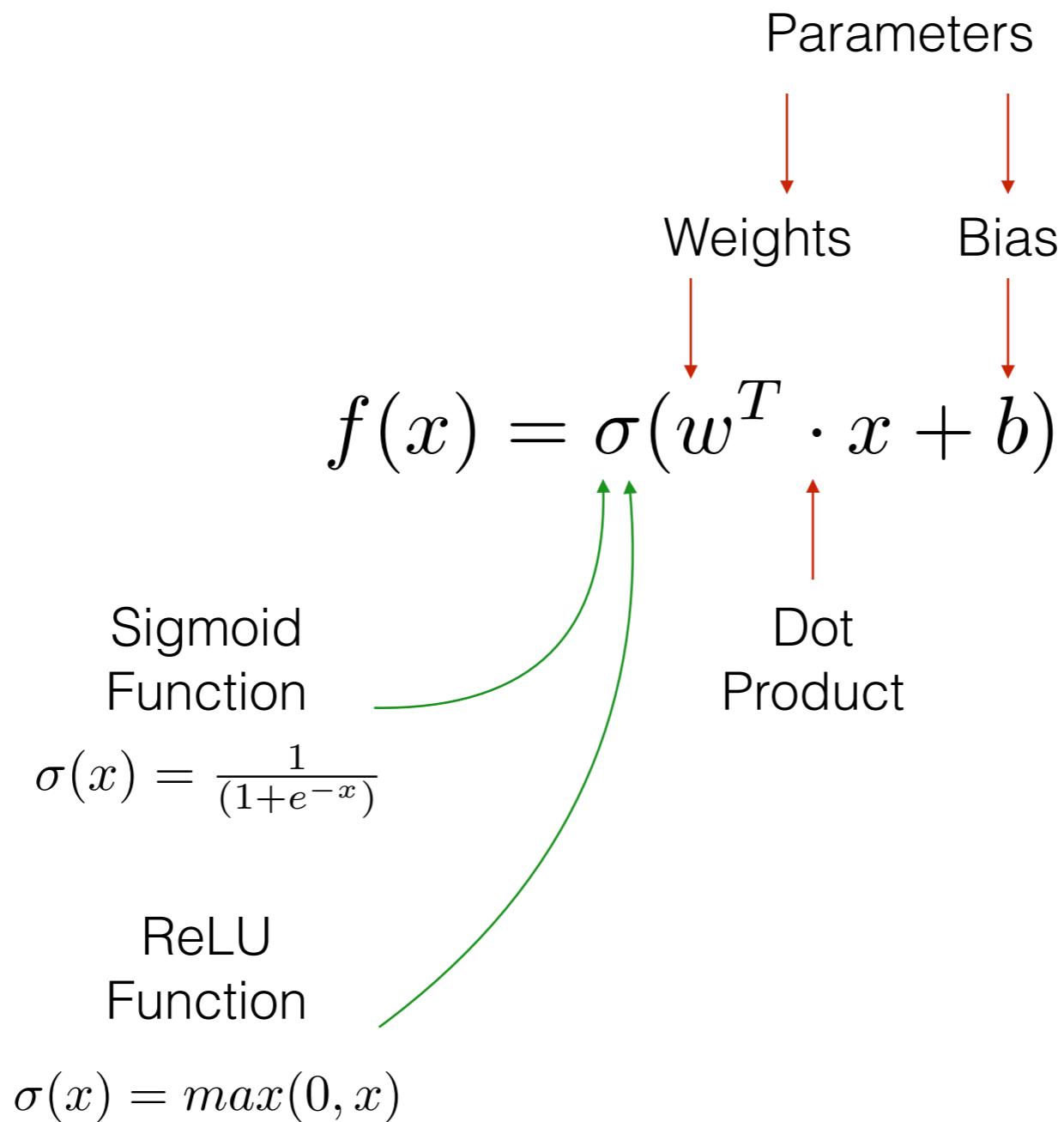
# Learning from Data

Training data: a set of  $(x^{(m)}, y^{(m)})$  pairs.

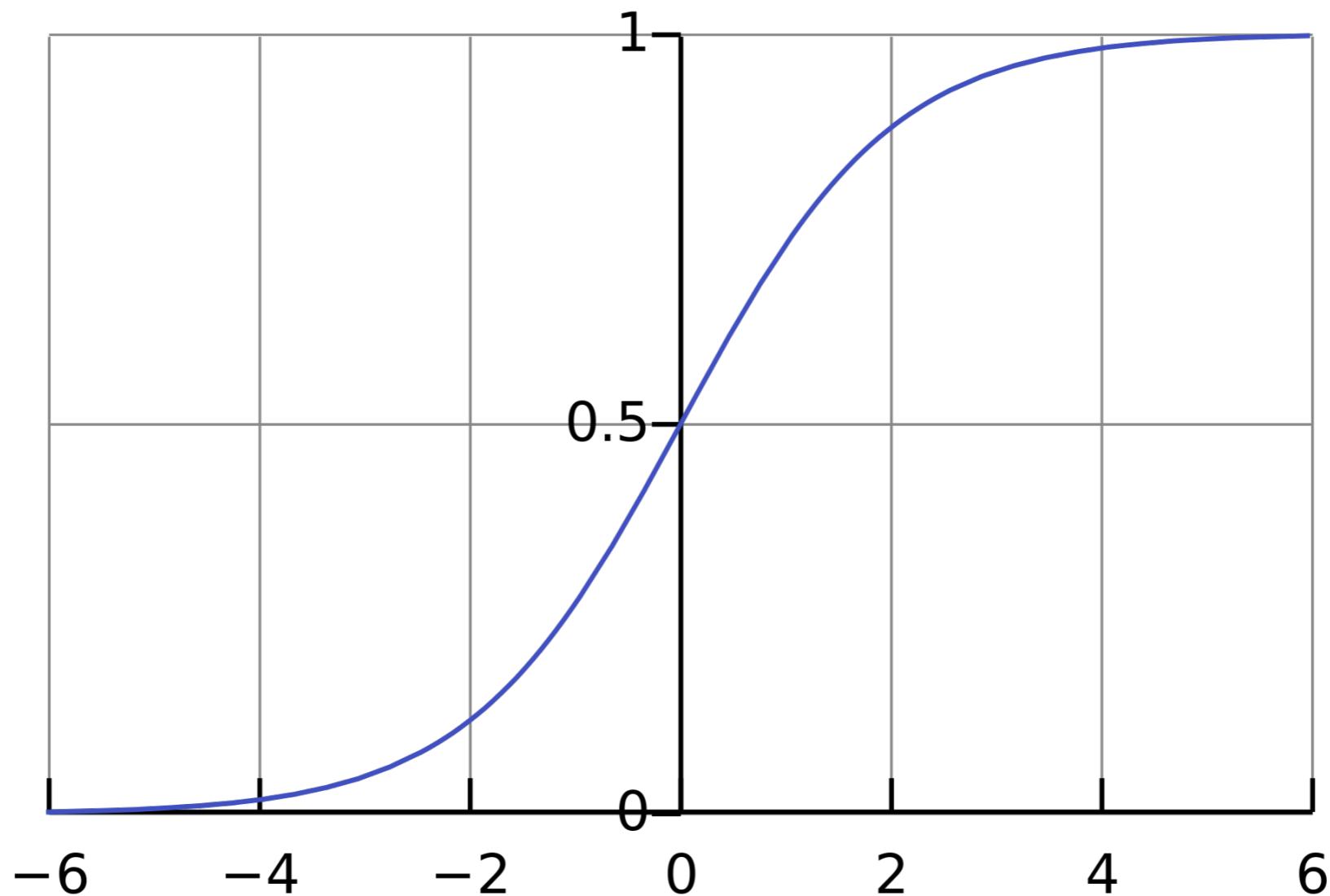
Learn a function  $f_w : x \rightarrow y$  to predict on new inputs  $x$ .

1. Choose a model function family  $f_w$ .
2. Optimize parameters  $w$ .

# 1-layer neural net model

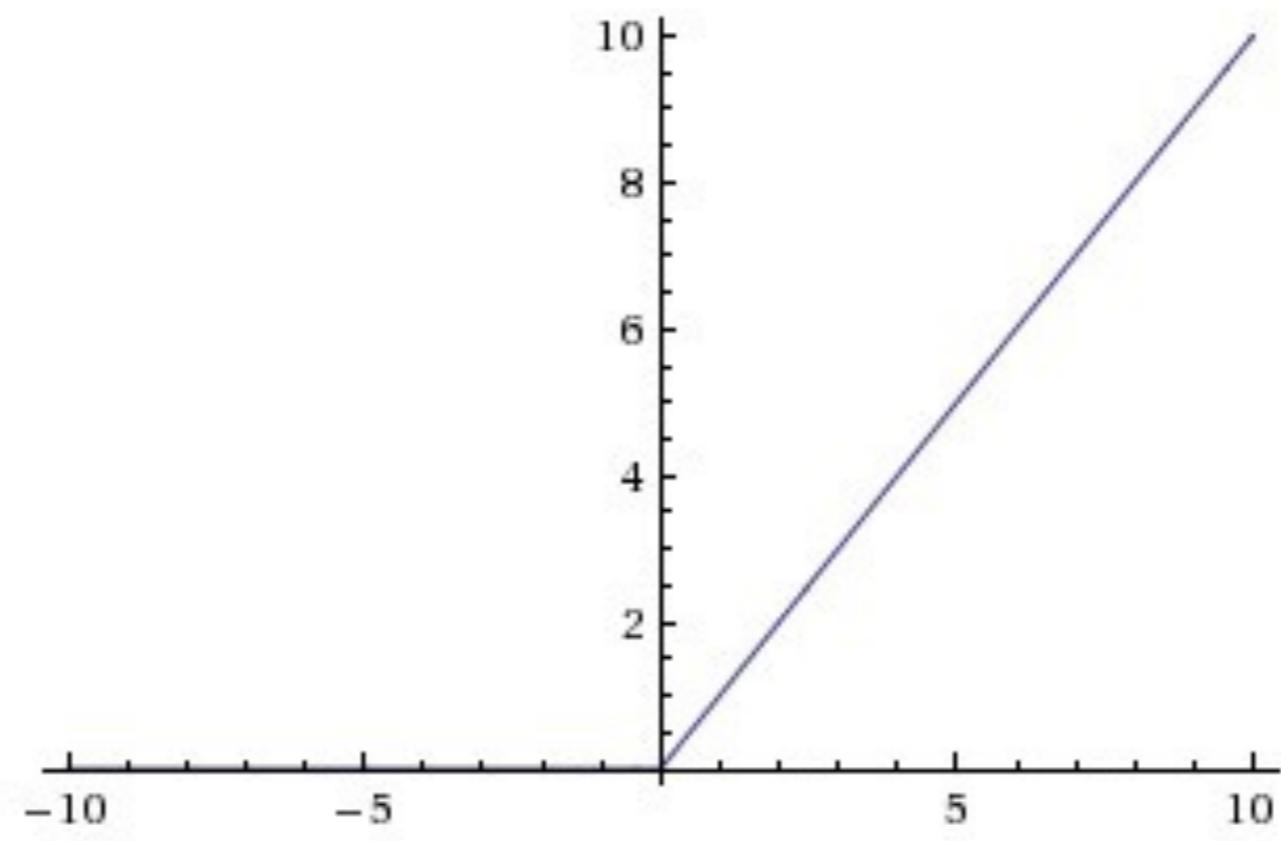


## 1-layer neural net model



$$\sigma(x) = \frac{1}{(1+e^{-x})}$$

# 1-layer neural net model



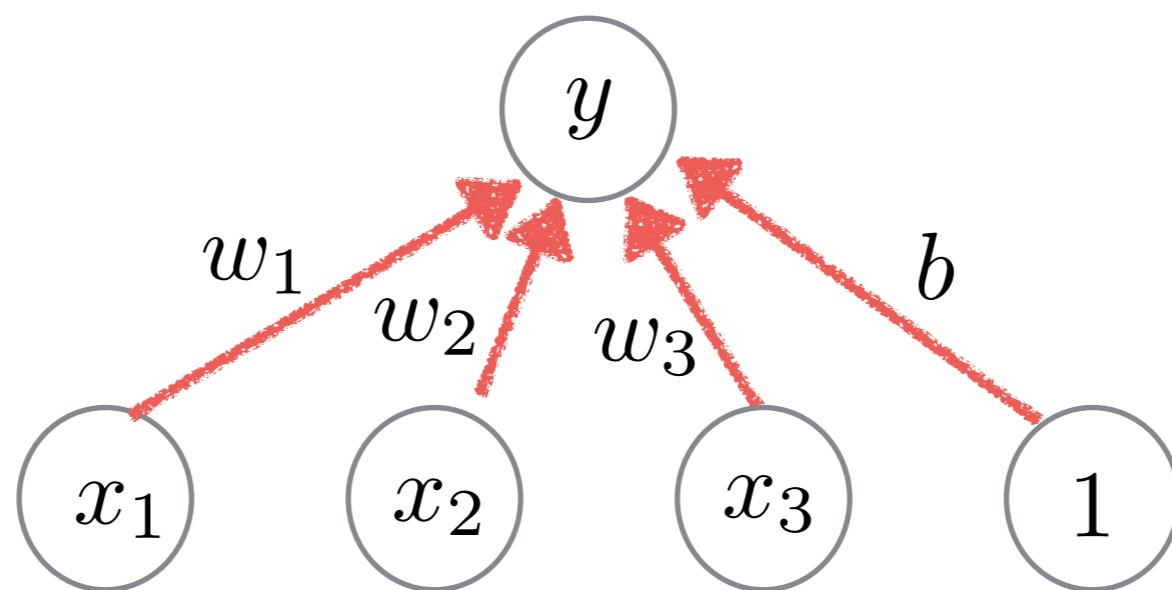
$$\sigma(x) = \max(0, x)$$

Table 3: Non-linearities tested.

Name	Formula	Year
none	$y = x$	-
sigmoid	$y = \frac{1}{1+e^{-x}}$	1986
tanh	$y = \frac{e^{2x}-1}{e^{2x}+1}$	1986
ReLU	$y = \max(x, 0)$	2010
(centered) SoftPlus	$y = \ln(e^x + 1) - \ln 2$	2011
LReLU	$y = \max(x, \alpha x), \alpha \approx 0.01$	2011
maxout	$y = \max(W_1x + b_1, W_2x + b_2)$	2013
APL	$y = \max(x, 0) + \sum_{s=1}^S a_i^s \max(0, -x + b_i^s)$	2014
VLReLU	$y = \max(x, \alpha x), \alpha \in [0.1, 0.5]$	2014
RReLU	$y = \max(x, \alpha x), \alpha = \text{random}(0.1, 0.5)$	2015
PReLU	$y = \max(x, \alpha x), \alpha$ is learnable	2015
ELU	$y = x, \text{ if } x \geq 0, \text{ else } \alpha(e^x - 1)$	2015

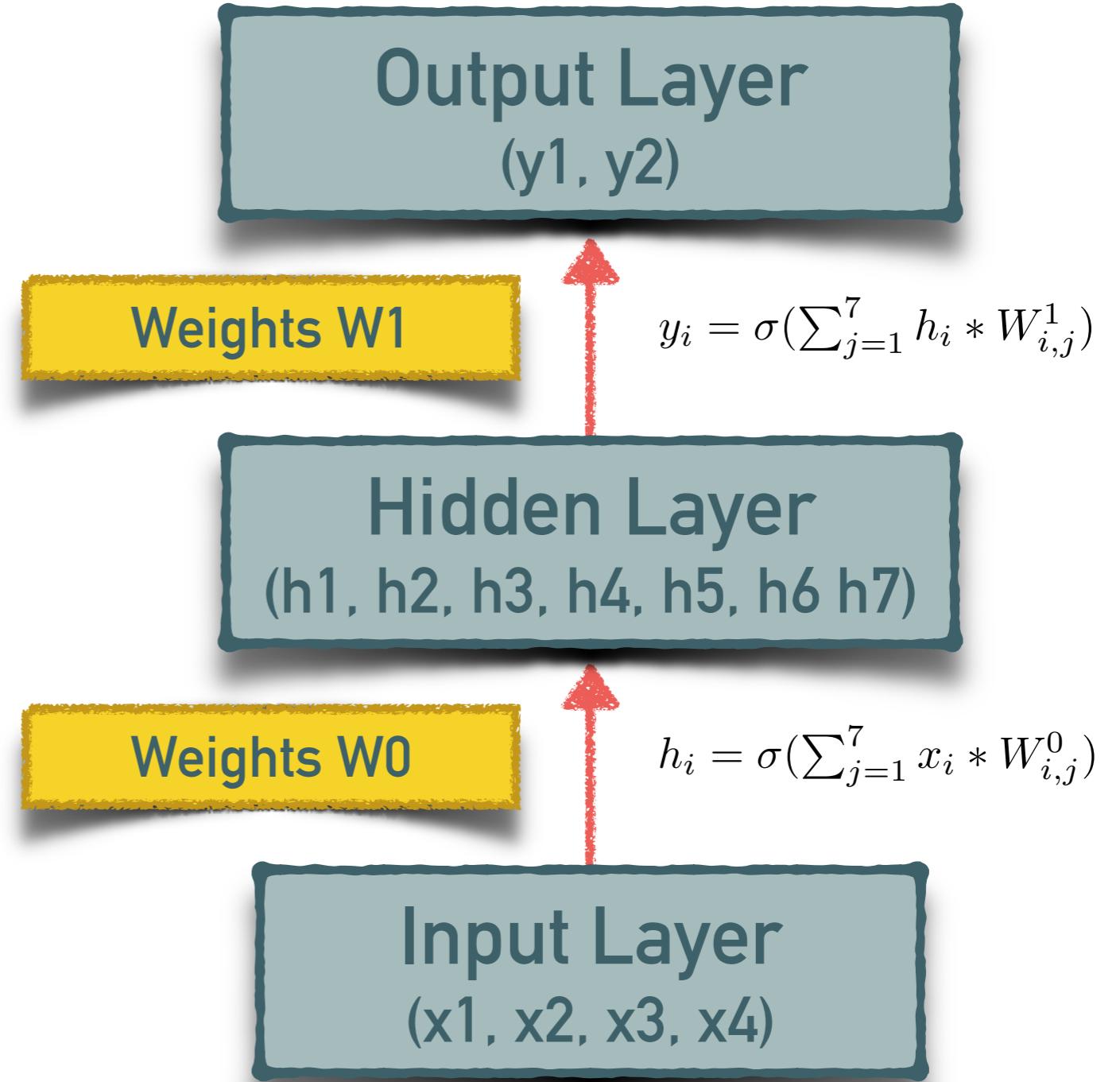
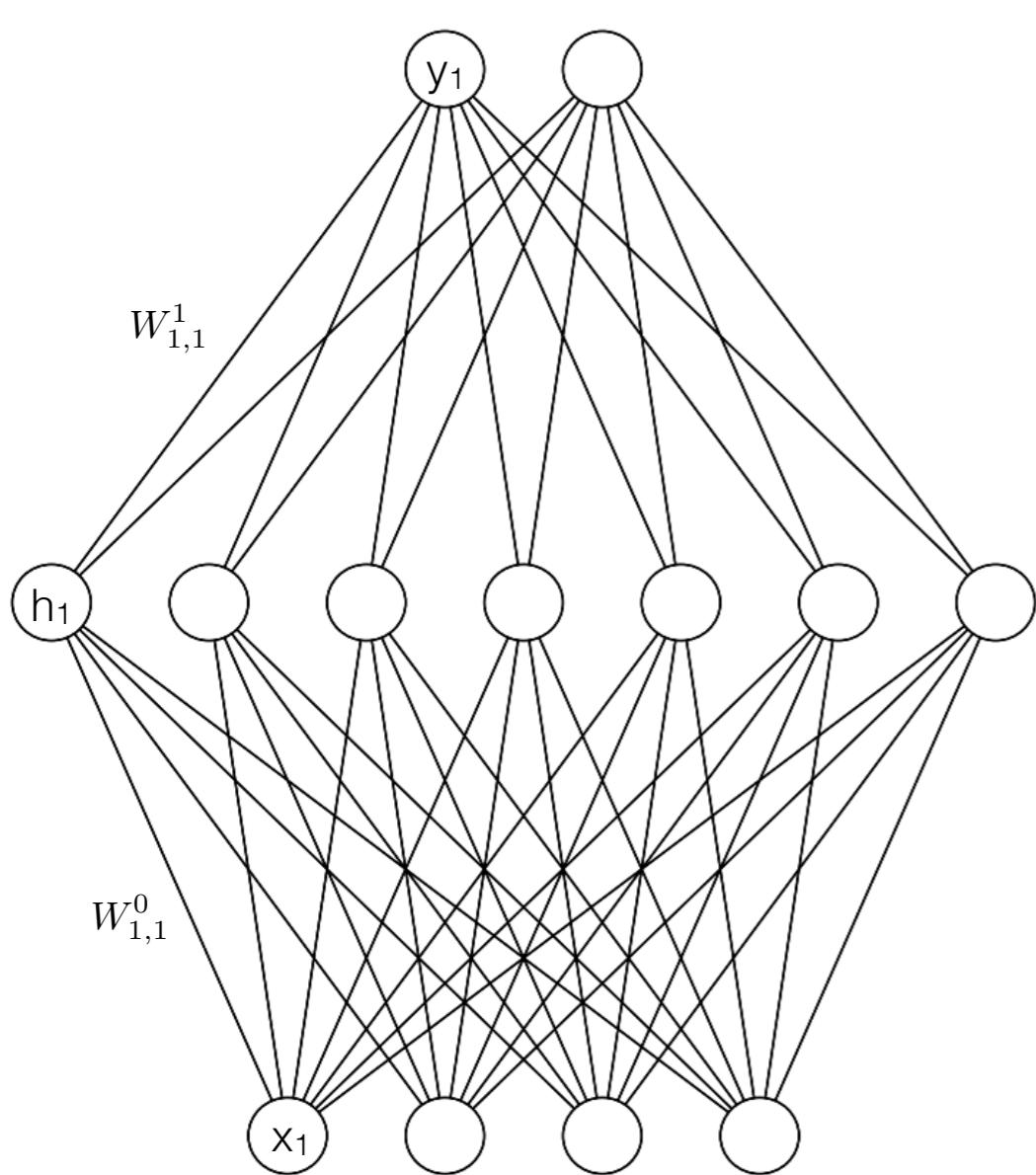
# 1-layer neural net model

$$f(x) = \sigma(w^T \cdot x + b)$$

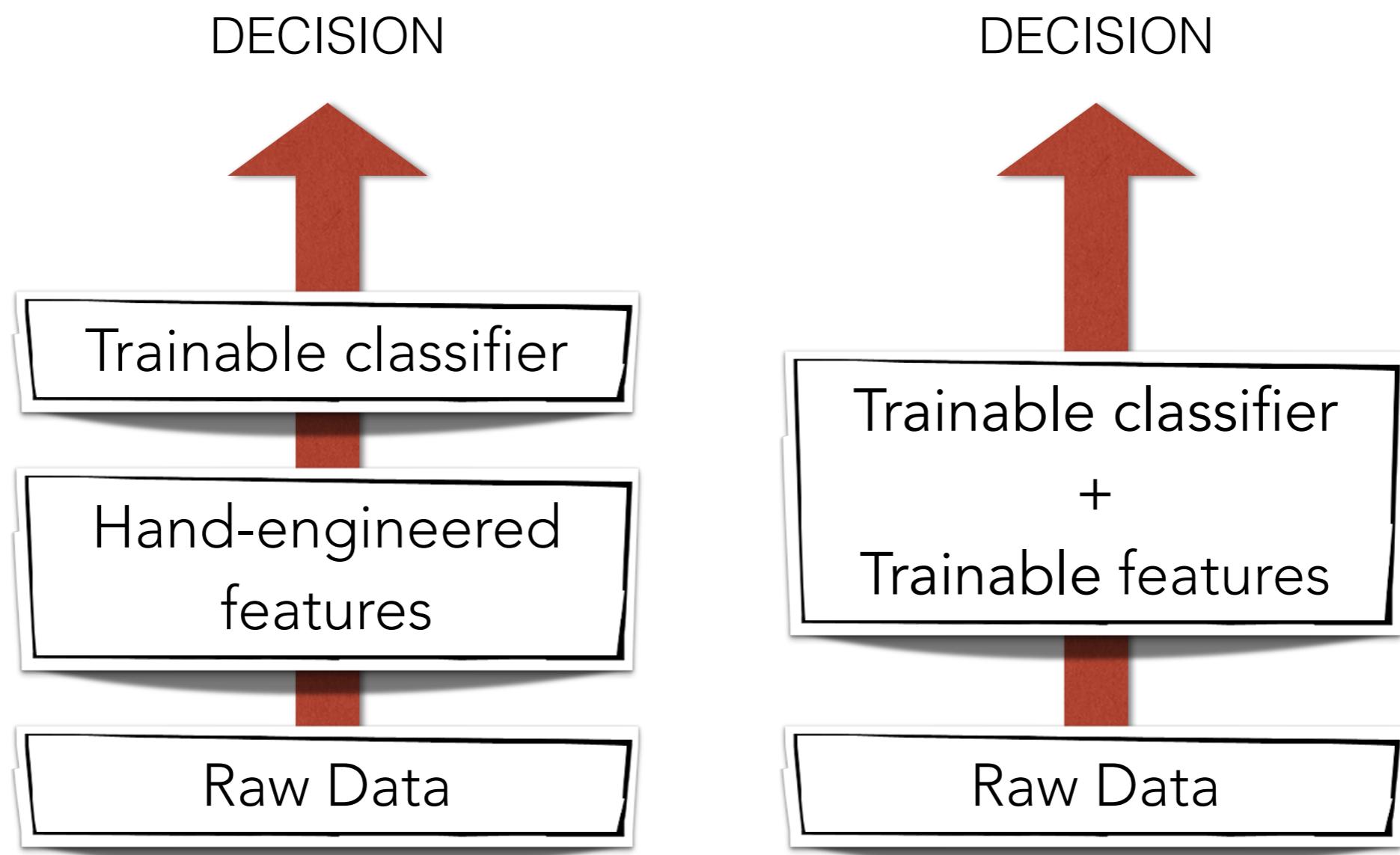


Graphical Representation

## 2-layer neural net model

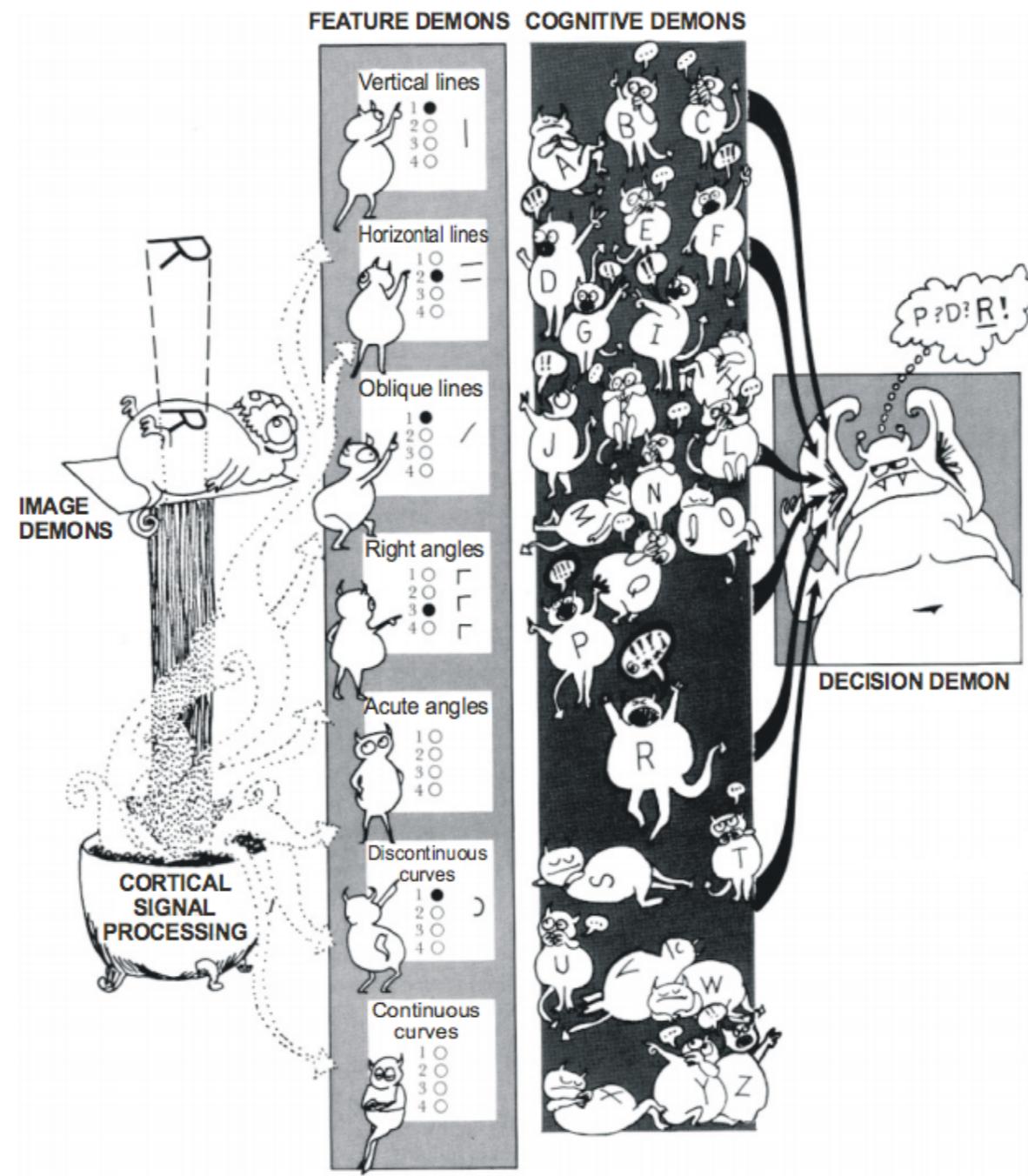
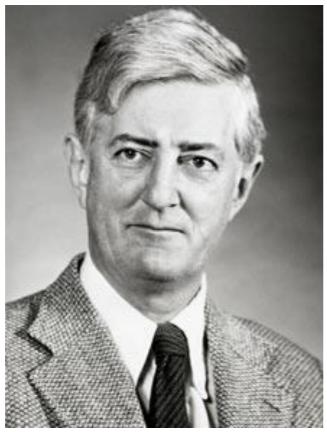


Computing the activation of one layer from the previous one can be written as a matrix-vector multiplication!



STANDARD MACHINE  
LEARNING

DEEP LEARNING



Pandemonium  
An early architecture of vision, proposed by Oliver  
Selfridge in 1959.

# Hype is not new



## NEW NAVY DEVICE LEARNS BY DOING

Psychologist Shows Embryo of Computer Designed to Read and Grow Wiser

WASHINGTON, July 7 (UPI)—The Navy revealed the embryo of an electronic computer today that it expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence.

The embryo—the Weather Bureau's \$2,000,000 "704" computer—learned to differentiate between right and left after fifty attempts in the Navy's demonstration for newsmen.

The service said it would use this principle to build the first of its Perceptron thinking machines that will be able to read and write. It is expected to be finished in about a year at a cost of \$100,000.

Dr. Frank Rosenblatt, designer of the Perceptron, conducted the demonstration. He said the machine would be the first device to think as the human brain. As do human be-

ings, Perceptron will make mistakes at first, but will grow wiser as it gains experience, he said.

Dr. Rosenblatt, a research psychologist at the Cornell Aeronautical Laboratory, Buffalo, said Perceptrons might be fired to the planets as mechanical space explorers.

### Without Human Controls

The Navy said the perceptron would be the first non-living mechanism "capable of receiving, recognizing and identifying its surroundings without any human training or control."

The "brain" is designed to remember images and information it has perceived itself. Ordinary computers remember only what is fed into them on punch cards or magnetic tape.

Later Perceptrons will be able to recognize people and call out their names and instantly translate speech in one language to speech or writing in another language, it was predicted.

Mr. Rosenblatt said in principle it would be possible to build brains that could reproduce themselves on an assembly line and which would be conscious of their existence.

1958 New York Times...

In today's demonstration, the "704" was fed two cards, one with squares marked on the left side and the other with squares on the right side.

### Learns by Doing

In the first fifty trials, the machine made no distinction between them. It then started registering a "Q" for the left squares and "O" for the right squares.

Dr. Rosenblatt said he could explain why the machine learned only in highly technical terms. But he said the computer had undergone a "self-induced change in the wiring diagram."

The first Perceptron will have about 1,000 electronic "association cells" receiving electrical impulses from an eye-like scanning device with 400 photo-cells. The human brain has 10,000,000,000 responsive cells, including 100,000,000 connections with the eyes.