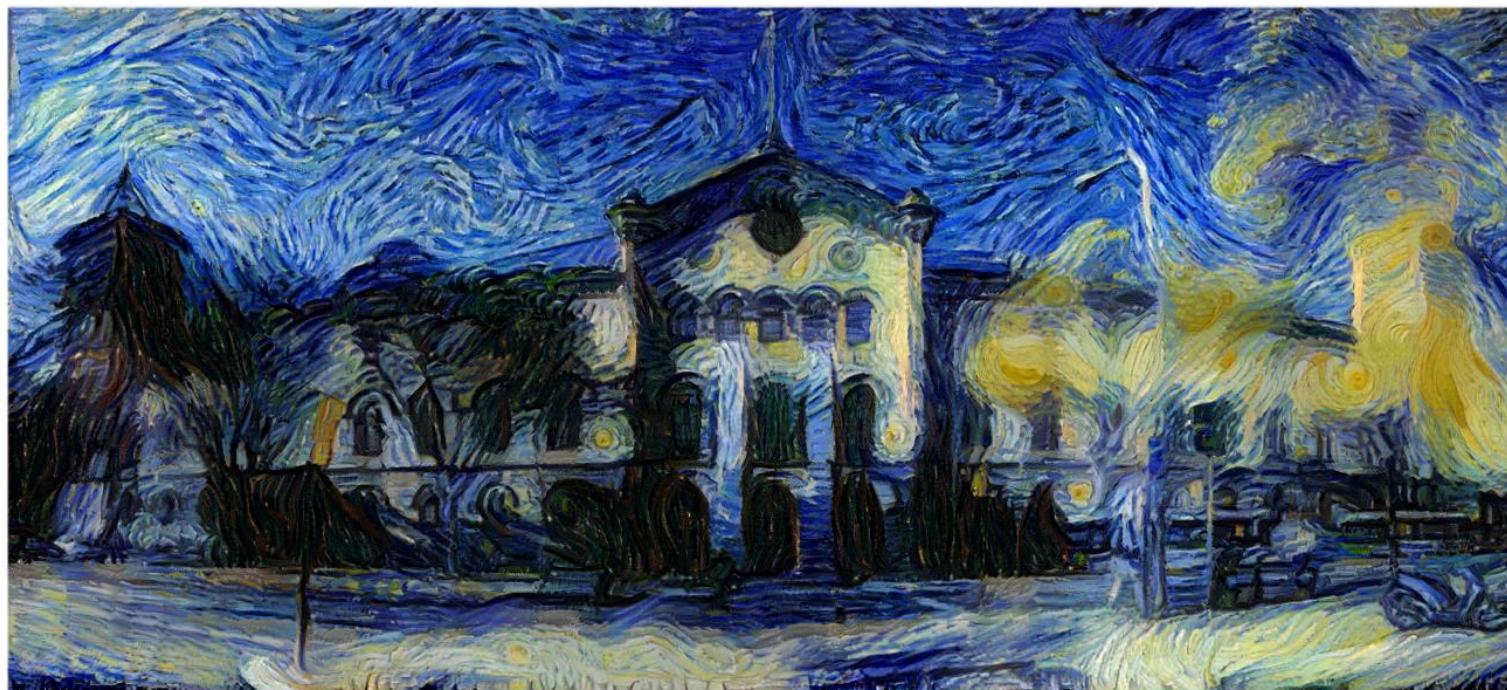




UNIVERSITAT DE  
BARCELONA



# Deep Learning From Scratch

BBVA, Madrid, 2016

# Course Agenda

Introduction to Deep Learning and its applications. Docker. Using the Jupyter notebook. What's a neural network?

Basic Concepts: Score & Loss functions, Optimization (SGD), Linear Regression.

Automated differentiation, Backpropagation, Training a Neural Network from Scratch

Tensorflow programming model. Keras

Convolutions. CNN models.

Recurrent Neural Networks

Unsupervised Learning

Advanced Models: Time series; Embeddings; Sentiment Analysis.

# Who



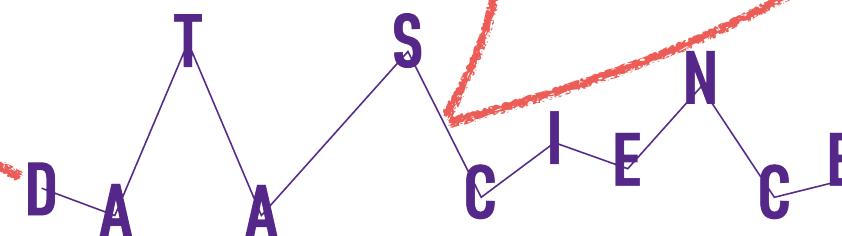
Santi Seguí,  
Lecturer at UB.



Oriol Pujol,  
Associate Professor at UB.



Jordi Vitrià.  
Full Professor at UB.



# What

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** and derivates. You can train (small) deep models in your laptop.

Top libraries by stars	
#1:	19683 ★ tensorflow/tensorflow
#2:	8995 ★ BVLC/caffe
#3:	4821 ★ fchollet/keras
#4:	4743 ★ Microsoft/CNTK
#5:	4284 ★ karpathy/convnetjs
#6:	4125 ★ torch/torch7
#7:	3200 ★ Theano/Theano
#8:	3155 ★ autumnai/leaf
#9:	2913 ★ dmlc/mxnet
#10:	2166 ★ deeplearning4j/deeplearning4j
#11:	2063 ★ tensorflow/skflow
#12:	1811 ★ Lasagne/Lasagne
#13:	1677 ★ NervanaSystems/neon
#14:	1230 ★ pfnet/chainer
#15:	1067 ★ IDSIA/brainstorm
#16:	980 ★ NVIDIA/DIGITS
#17:	721 ★ mila-udem/blocks

Top libraries by contributors	
#1:	216 ⚡ Theano/Theano
#2:	182 ⚡ BVLC/caffe
#3:	145 ⚡ fchollet/keras
#4:	126 ⚡ tensorflow/tensorflow
#5:	107 ⚡ dmlc/mxnet
#6:	81 ⚡ torch/torch7
#7:	63 ⚡ deeplearning4j/deeplearning4j
#8:	52 ⚡ Microsoft/CNTK
#9:	47 ⚡ pfnet/chainer
#10:	43 ⚡ Lasagne/Lasagne
#11:	41 ⚡ mila-udem/blocks
#12:	34 ⚡ NervanaSystems/neon
#13:	23 ⚡ NVIDIA/DIGITS
#14:	22 ⚡ tensorflow/skflow
#15:	14 ⚡ karpathy/convnetjs
#16:	12 ⚡ IDSIA/brainstorm
#17:	11 ⚡ autumnai/leaf

Top libraries by forks	
#1:	6804 ♦ tensorflow/tensorflow
#2:	5159 ♦ BVLC/caffe
#3:	1234 ♦ fchollet/keras
#4:	1161 ♦ Theano/Theano
#5:	1048 ♦ torch/torch7
#6:	995 ♦ dmlc/mxnet
#7:	828 ♦ Microsoft/CNTK
#8:	815 ♦ karpathy/convnetjs
#9:	762 ♦ deeplearning4j/deeplearning4j
#10:	483 ♦ Lasagne/Lasagne
#11:	324 ♦ NervanaSystems/neon
#12:	292 ♦ NVIDIA/DIGITS
#13:	248 ♦ tensorflow/skflow
#14:	240 ♦ pfnet/chainer
#15:	205 ♦ mila-udem/blocks
#16:	101 ♦ autumnai/leaf
#17:	87 ♦ IDSIA/brainstorm

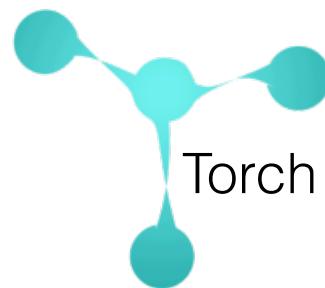
Top libraries: mixed score	
#1:	48.16 ↑ tensorflow/tensorflow
#2:	31.45 ↑ BVLC/caffe
#3:	11.21 ↑ fchollet/keras
#4:	10.00 ↑ Theano/Theano
#5:	9.13 ↑ torch/torch7
#6:	8.58 ↑ Microsoft/CNTK
#7:	7.96 ↑ dmlc/mxnet
#8:	7.68 ↑ karpathy/convnetjs
#9:	5.84 ↑ deeplearning4j/deeplearning4j
#10:	4.17 ↑ Lasagne/Lasagne
#11:	3.67 ↑ autumnai/leaf
#12:	3.31 ↑ NervanaSystems/neon
#13:	3.27 ↑ tensorflow/skflow
#14:	2.66 ↑ pfnet/chainer
#15:	2.38 ↑ NVIDIA/DIGITS
#16:	1.95 ↑ mila-udem/blocks
#17:	1.54 ↑ IDSIA/brainstorm

# BackEnds

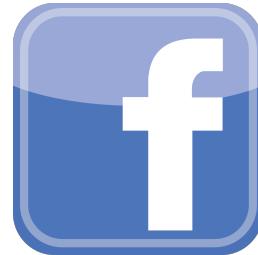


Python, C++  
MultiGPU  
Distributed

Google



Lua  
MultiGPU



theano

Python  
Large amount  
of sample  
code

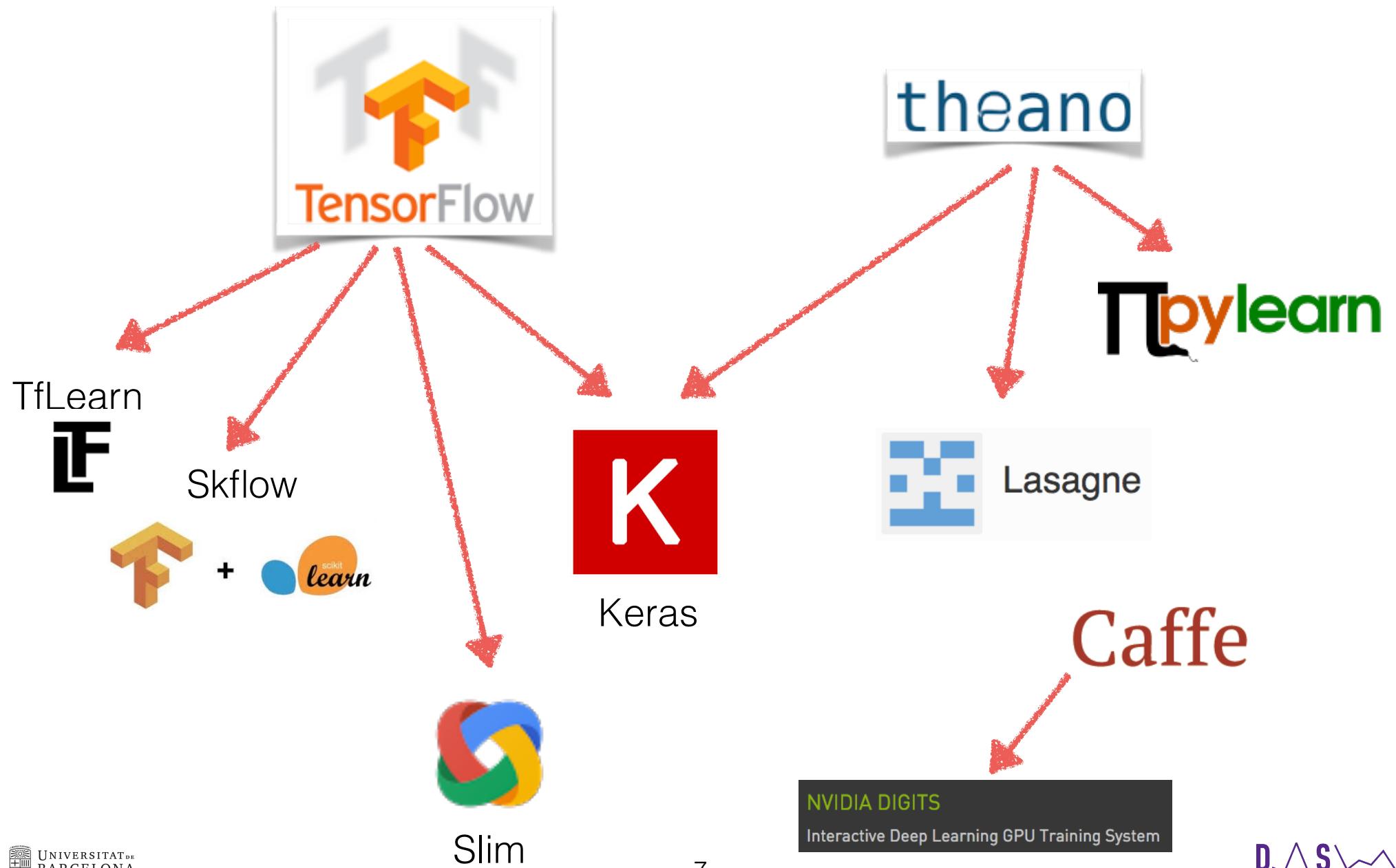
Université de Montréal

Caffe

Python, C++  
MultiGPU

Berkeley  
UNIVERSITY OF CALIFORNIA

# Tools



# Other BackEnds

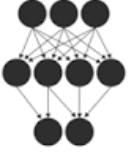


*dmlc*  
**mxnet**



Neon

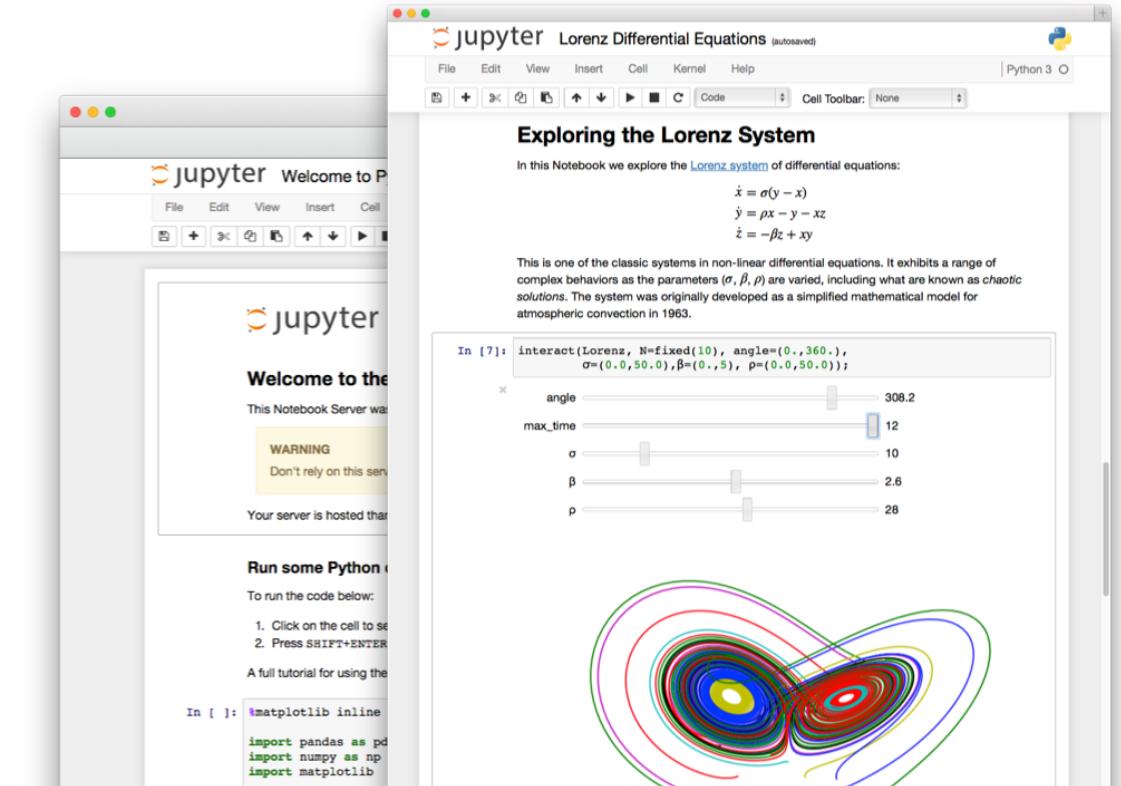
**DL4J** Deep Learning for Java

 **ConvNetJS**  
Deep Learning in your browser

# 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.

Skflow

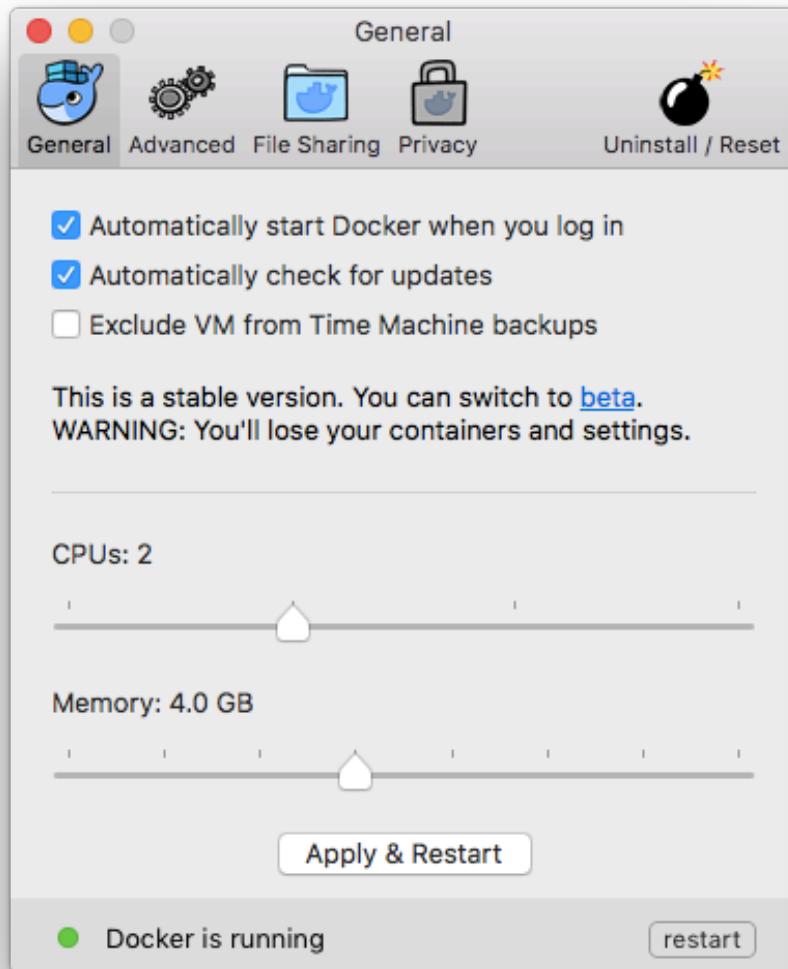


# Approach



We will use a **Docker Container**.

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. Finally, Docker's runtime options enable us to attach GPU devices when deploying on remote servers.



```
● ○ ● jordi — docker pull datascienceub/deepub — 78x28
Last login: Mon Nov 21 20:08:59 on ttys001
[MacBook-Air-de-jordi:~ jordi$ docker pull datascienceub/deepub
Using default tag: latest
latest: Pulling from datascienceub/deepub

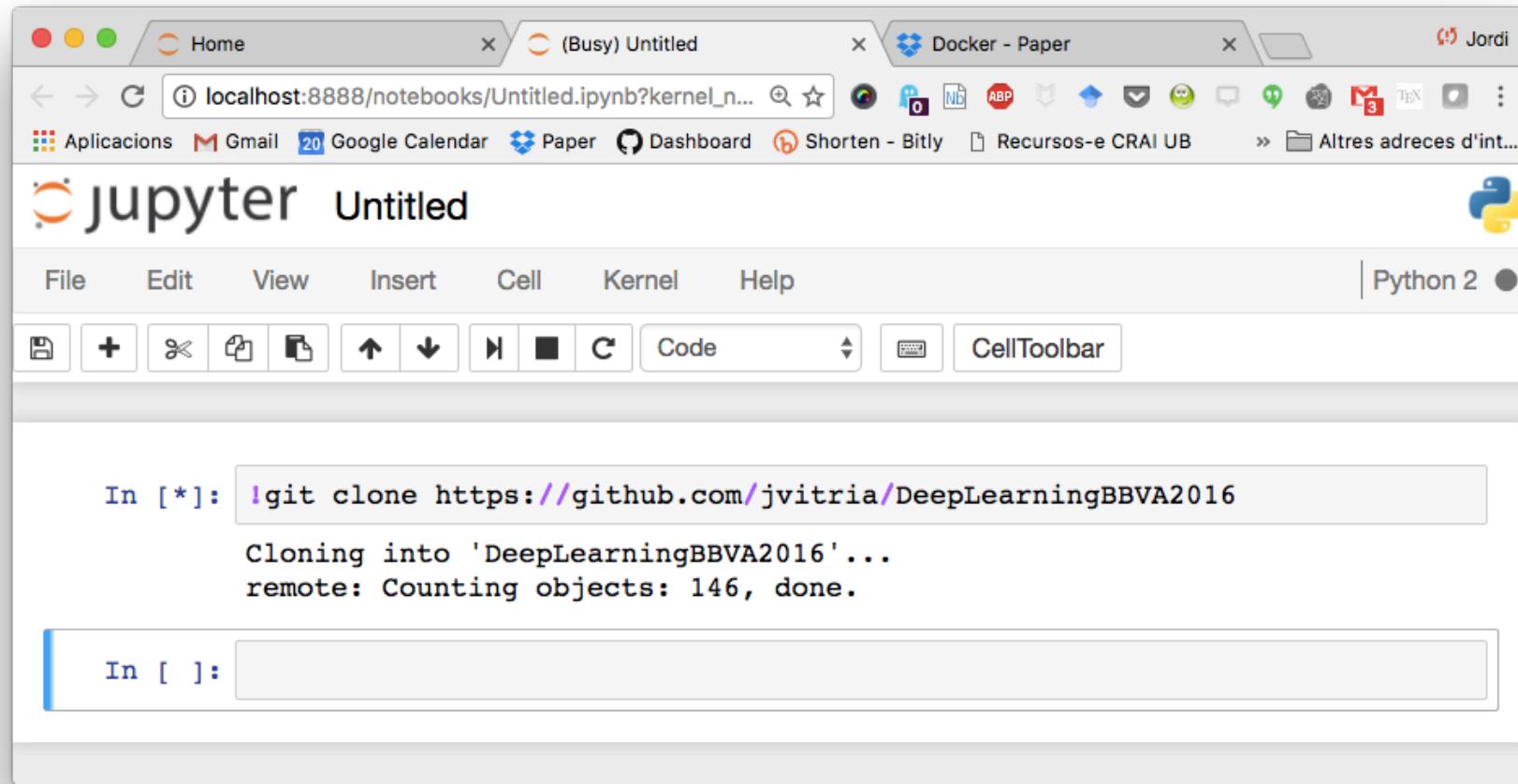
6c953ac5d795: Pull complete
3eed5ff20a90: Pull complete
f8419ea7c1b5: Pull complete
51900bc9e720: Pull complete
a3ed95caeb02: Pull complete
ebb77ce6e1c6: Pull complete
c145c1f339f5: Pull complete
9528c5352798: Extracting 68.52 MB
c1f8f4c880d4: Download complete
f5d83de9c678: Download complete
d819f1ec59a0: Download complete
bfc1d5e3de1c: Download complete
182b64c1f020: Download complete
ce321bdf714b: Download complete
ca10bcc60e7c: Download complete
b4fd7a47b122: Download complete
1457c46eff3f: Downloading 25.95 MB/85.36 MB
e52d876fa529: Downloading 6.274 MB/11.73 MB
6076beb51bc0: Download complete
```

```
DeepLearningBBVA2016-master — docker run -it -p 8888:8888 -v ~/Dropbox/Dee...  
Last login: Mon Nov 21 20:17:10 on ttys001  
[MacBook-Air-de-jordi:~ jordi$ pwd  
/Users/jordi  
[MacBook-Air-de-jordi:~ jordi$ cd Dropbox/DeepLearningBBVA2016-master/  
MacBook-Air-de-jordi:DeepLearningBBVA2016-master jordi$  
MacBook-Air-de-jordi:DeepLearningBBVA2016-master jordi$ docker run -it -p 8888:8888  
-v /$(pwd):/notebooks datascienceub/deepub  
[I 19:18:44.961 NotebookApp] Writing notebook server cookie secret to /root/.local/s  
hare/jupyter/runtime/notebook_cookie_secret  
[W 19:18:44.975 NotebookApp] WARNING: The notebook server is listening on all IP add  
resses and not using encryption. This is not recommended.  
[W 19:18:44.975 NotebookApp] WARNING: The notebook server is listening on all IP add  
resses and not using authentication. This is highly insecure and not recommended.  
[I 19:18:44.982 NotebookApp] Serving notebooks from local directory: /notebooks  
[I 19:18:44.982 NotebookApp] 0 active kernels  
[I 19:18:44.982 NotebookApp] The Jupyter Notebook is running at: http://[all ip addr  
esses on your system]:8888/  
[I 19:18:44.983 NotebookApp] Use Control-C to stop this server and shut down all ker  
nels (twice to skip confirmation).
```

```
docker run -it -p 8888:8888 -v /$(pwd):/notebooks  
datascienceub/deepub
```

The screenshot shows a web browser window with the URL `localhost:8888/tree` in the address bar. The title bar says "Home". On the right side of the browser, there is a user profile icon for "Jordi". Below the address bar, there is a toolbar with various icons. The main content area is titled "jupyter" and displays a file tree. At the top level, there is a folder named "tmp". Below it, there is a list of notebook files (ipynb):

- 0. Deep Learning from Scratch.ipynb
- 1. Basic Concepts.ipynb
- 2.Backpropagation.ipynb
- 3.0 Basic Concepts.ipynb
- 3.1 Linear Regression.ipynb
- 3.2 Logistic Regression.ipynb
- 3.3 Perceptron.ipynb
- 3.4 Multilayer Perceptron.ipynb
- 3.5 Linear Regression - TensorBoard.ipynb
- 4.1 Convolutional Network.ipynb
- 4.2 Convolutional Network - Smile DataSet.ipynb
- 4.3 Convolutional Network - Smile DataSet - Alexnet.ipynb
- 5. Recurrent Neural Networks.ipynb
- 6.Unsupervised\_final.ipynb
- 7. Time Series Prediction.ipynb
- Keras.ipynb



# THE REVENANT

INSPIRED BY TRUE EVENTS  
JANUARY 8

A movie poster for "The Revenant". The background is a dark, moody landscape with a forest silhouette and a close-up of a man's face in the foreground. The man has long hair and a beard, looking upwards with a weary expression. The title "THE REVENANT" is at the top in large white letters, followed by "INSPIRED BY TRUE EVENTS" and "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** showed that Rosenblatt's original system was painfully limited, literally blind to some simple logical functions like "exclusive-or" (As in, you can have the cake or the pie, but not both). What had become known as the field of "neural networks" all but disappeared.

# First neural network winter is coming





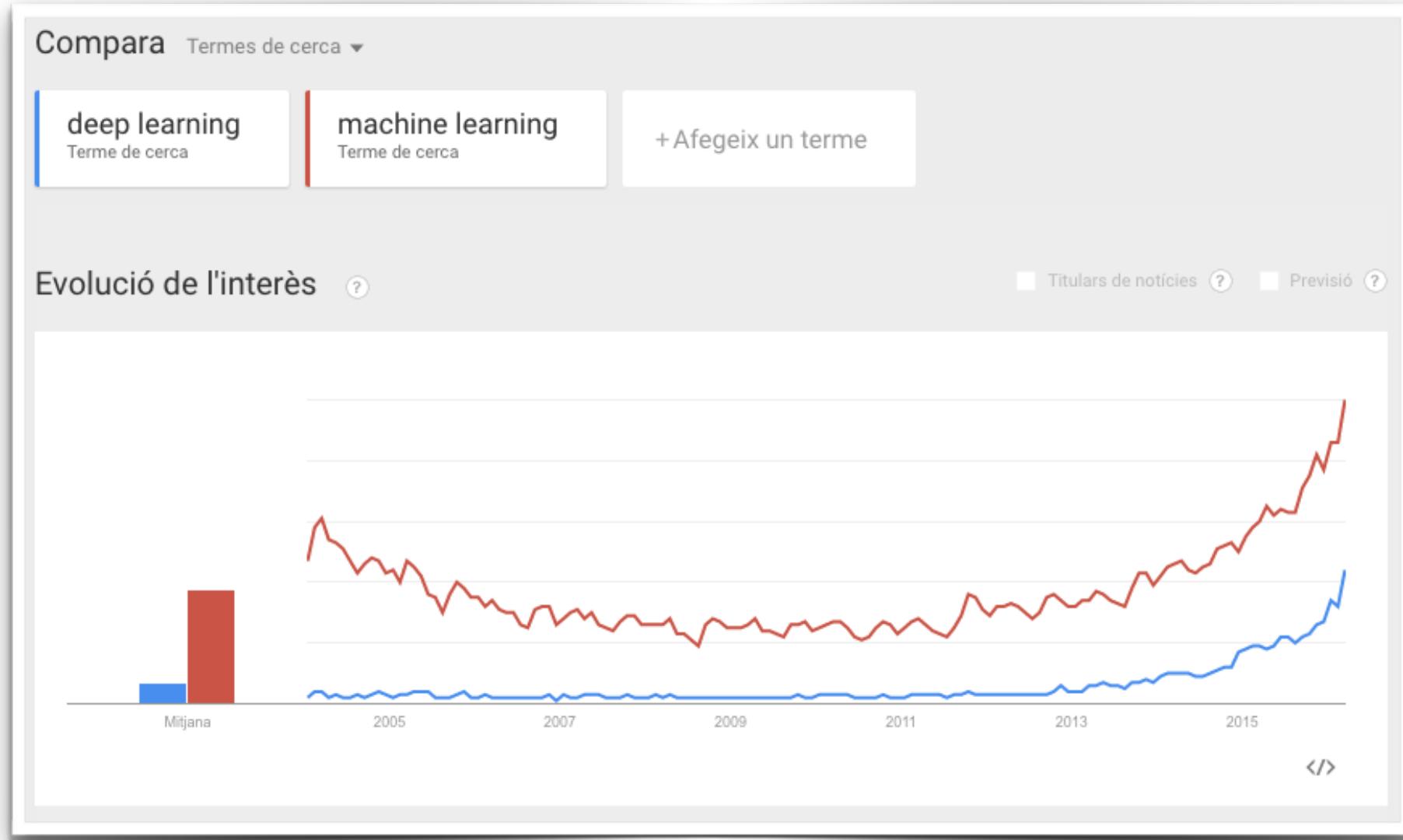
- In 1982, interest in the field was renewed. **John Hopfield** of Caltech presented a paper to the National Academy of Sciences. His approach was to create more useful machines by using bidirectional lines. Previously, the connections between neurons was only one way.
- In 1986, the problem was how 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

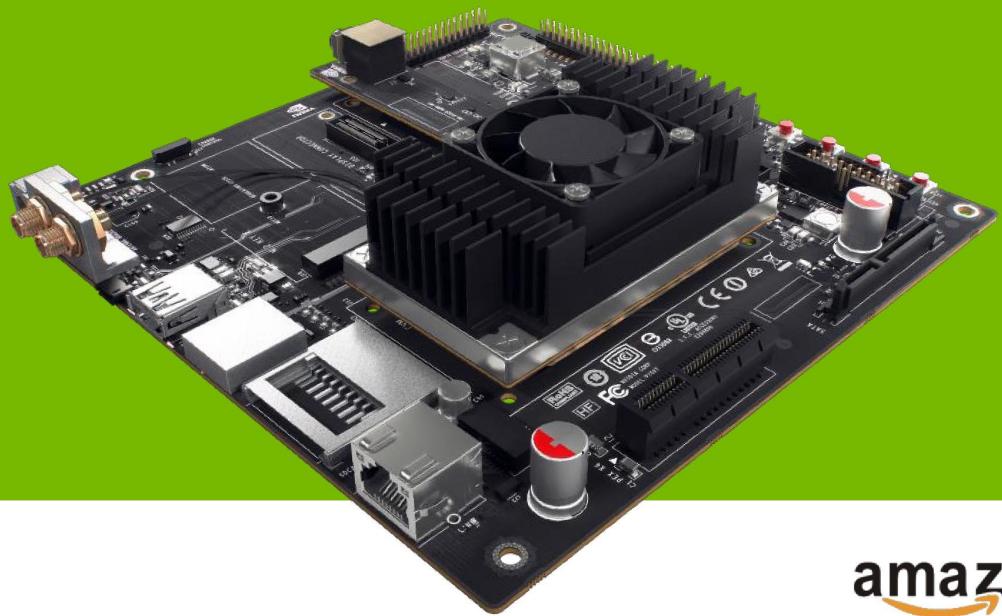


- With the ascent of Support Vector Machines and the failure of backpropagation, the early 2000s were a dark time for neural net research.
- Then, what every researcher must dream of actually happened: G.Hinton, S.Osindero, and Y.W.Teh published a paper in 2006 that was seen as a breakthrough, a breakthrough 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



**Jetson TX1 Developer Kit**  
**\$599 retail**  
**\$299 edu**  
**Pre-order Nov 12**  
**Shipping Nov 16 (US)**  
**Intl to follow**



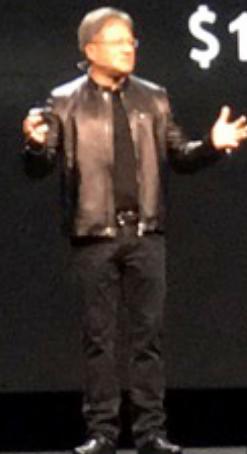


# NVIDIA DGX-1

WORLD'S FIRST  
DEEP LEARNING SUPERCOMPUTER

170TF | "250 servers in-a-box" | [nvidia.com/dgx1](http://nvidia.com/dgx1)

\$129,000



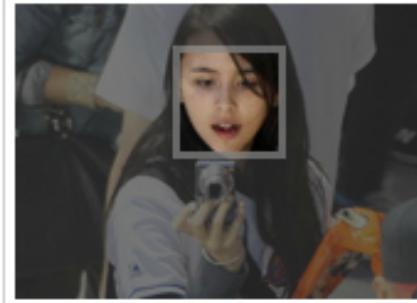
## Definitions

- **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.



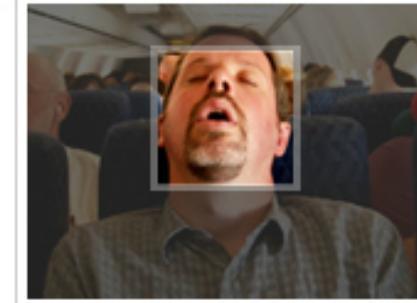
# Face recognition.



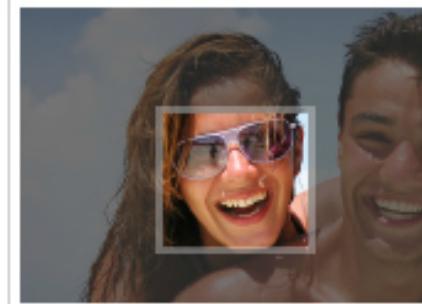
Who is this?



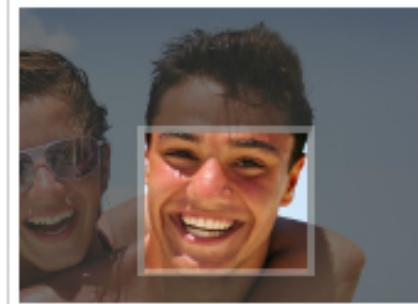
Who is this?



Who is this?



Who is this?



Who is this?



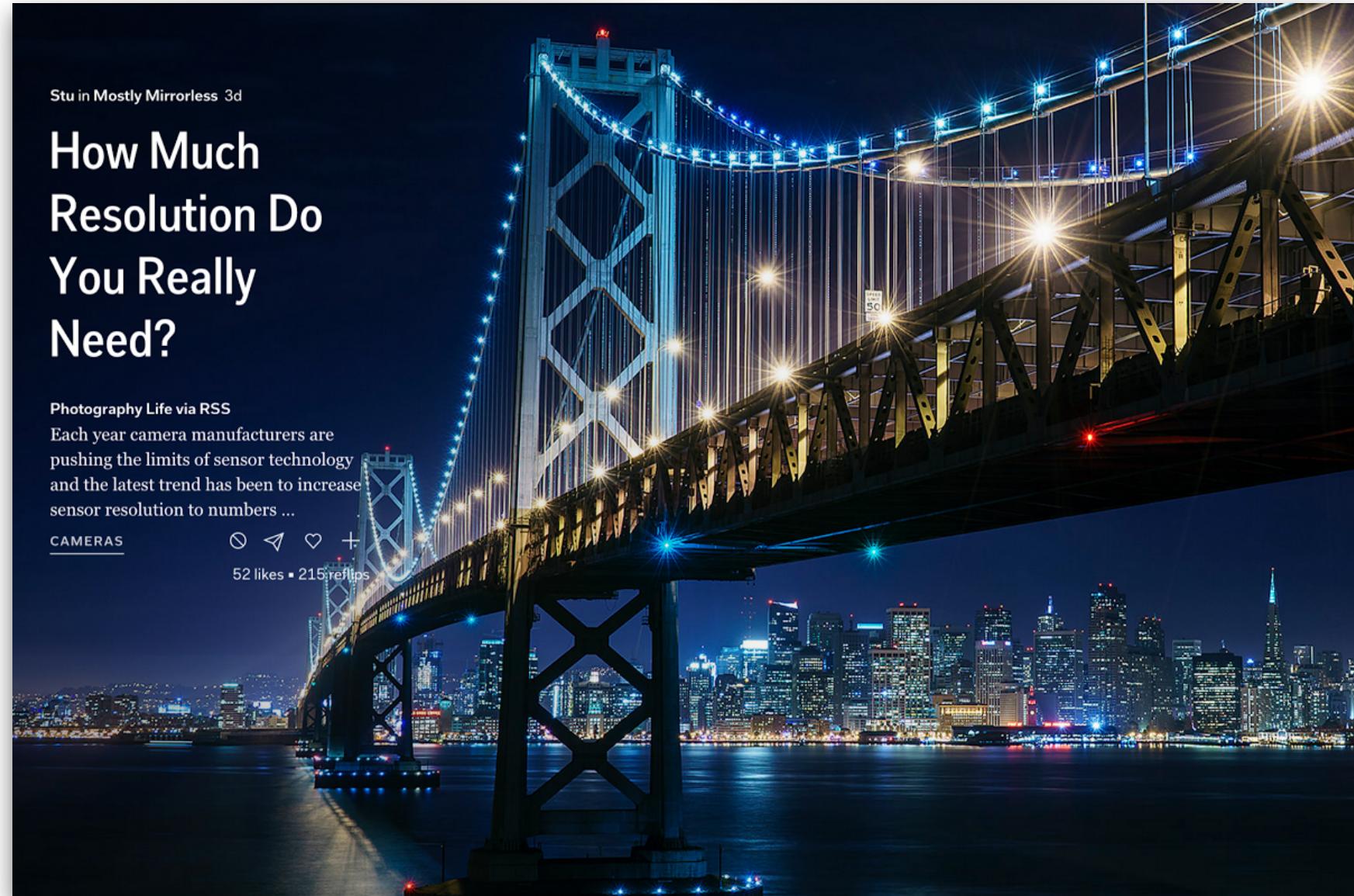
Who is this?

DeepFace (Facebook): Accuracy of 97.35%

# New applications: navigation and mapping.

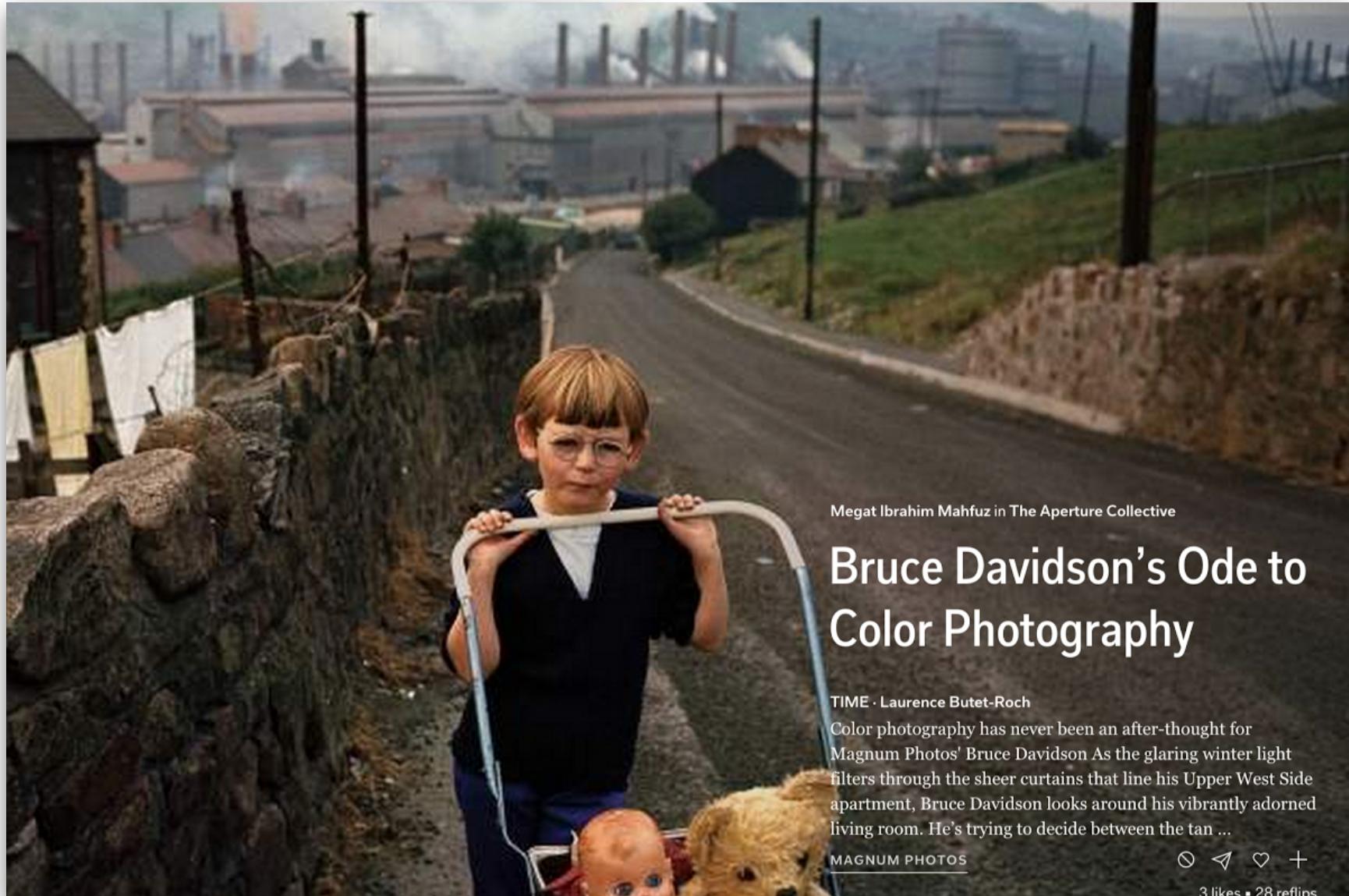


# New applications: Image Upscaling (Flipboard)



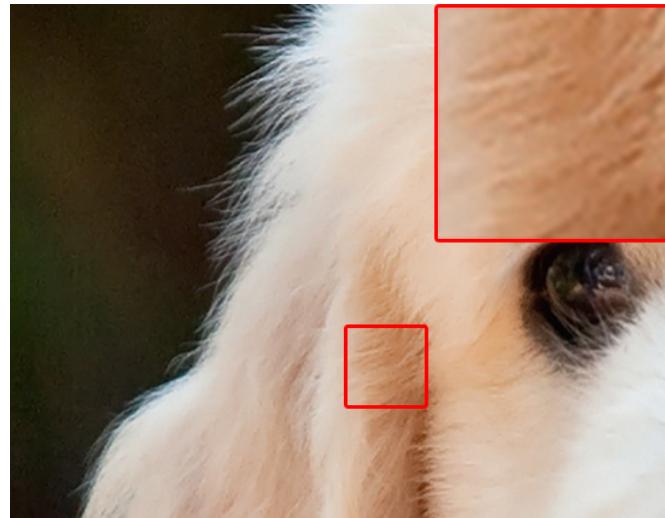
<http://engineering.flipboard.com/2015/05/scaling-convnets/>

# New applications: Image Upscaling (Flipboard)

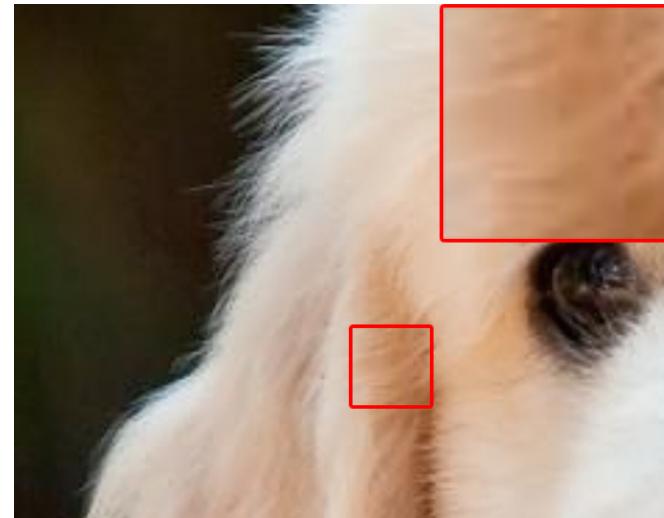


<http://engineering.flipboard.com/2015/05/scaling-convnets/>

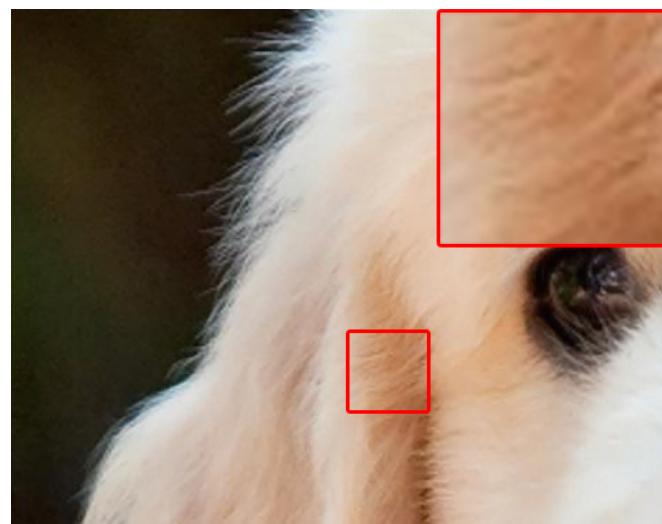
## New applications: Image Upscaling (Flipboard)



Original



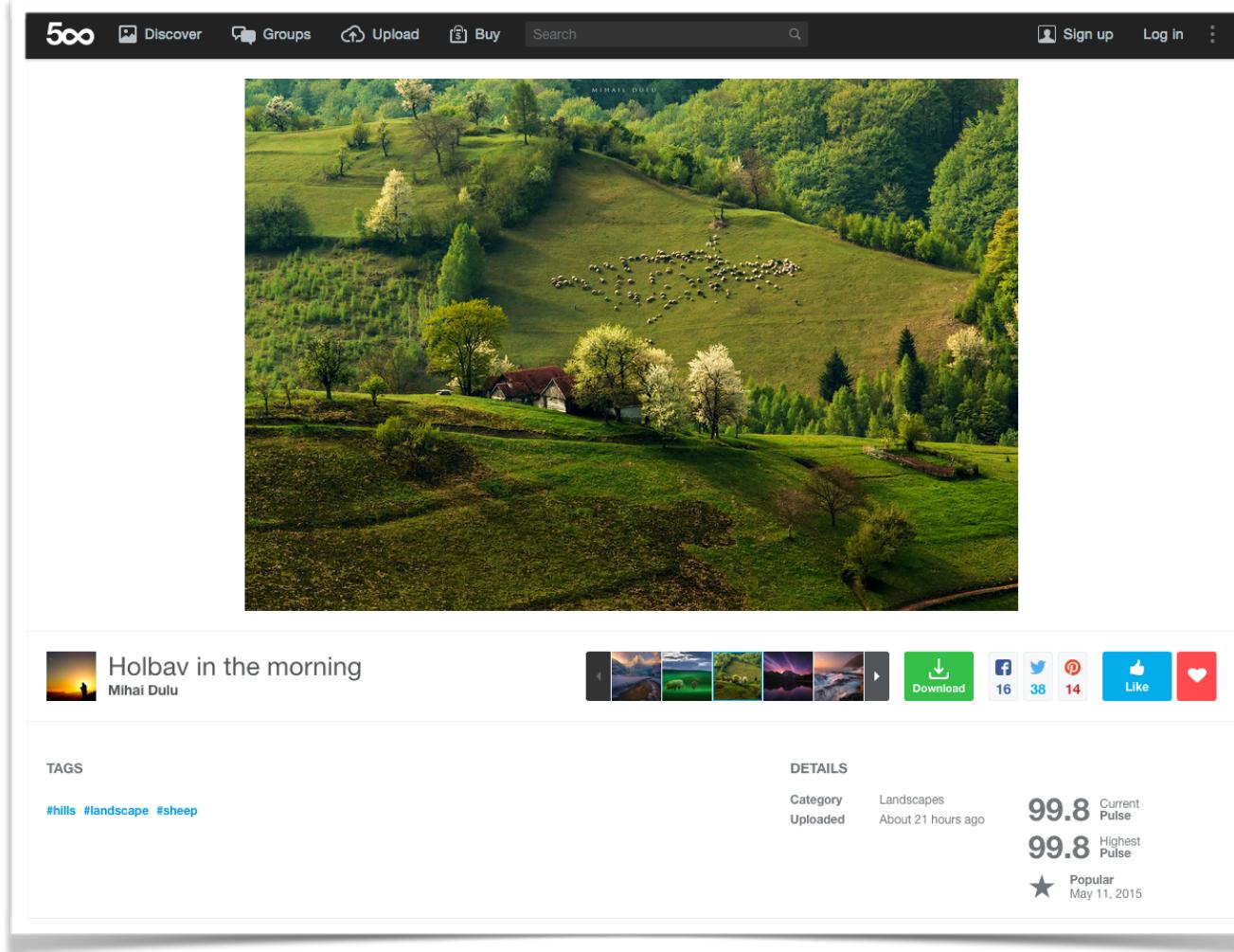
Bicubic



Model

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

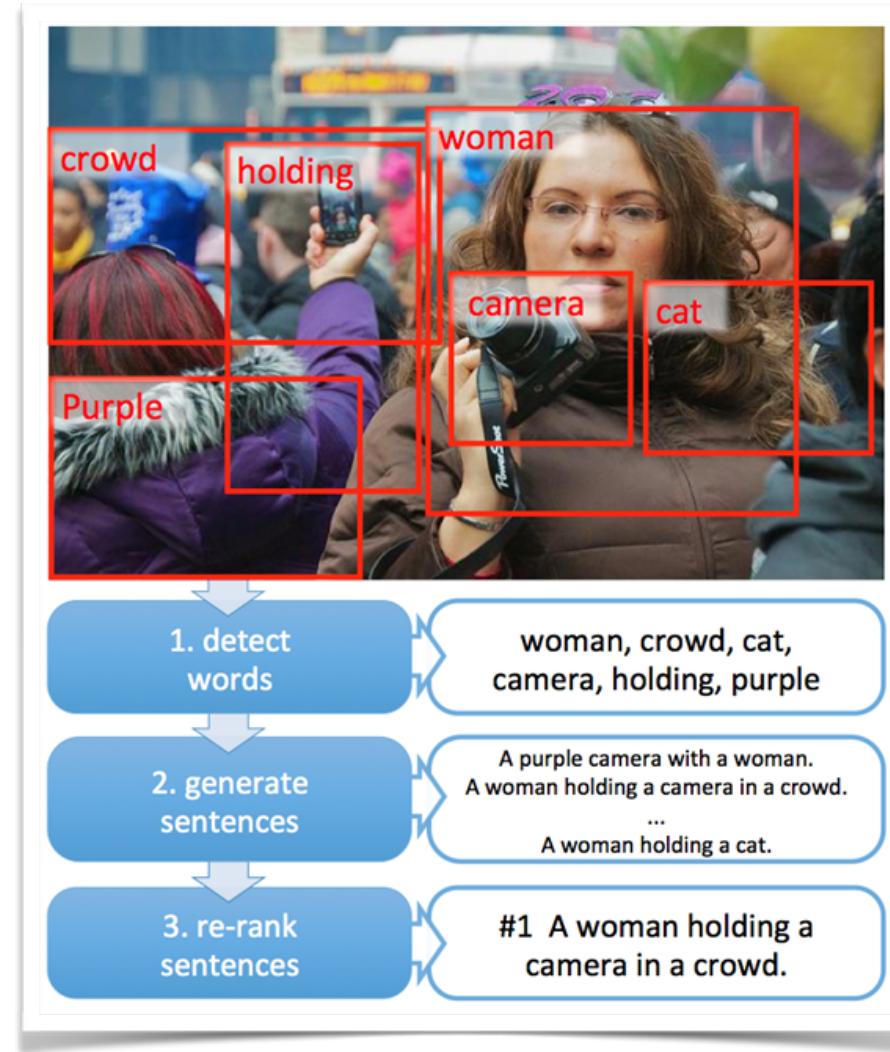
# New applications: Non visual data prediction



## What is Pulse?

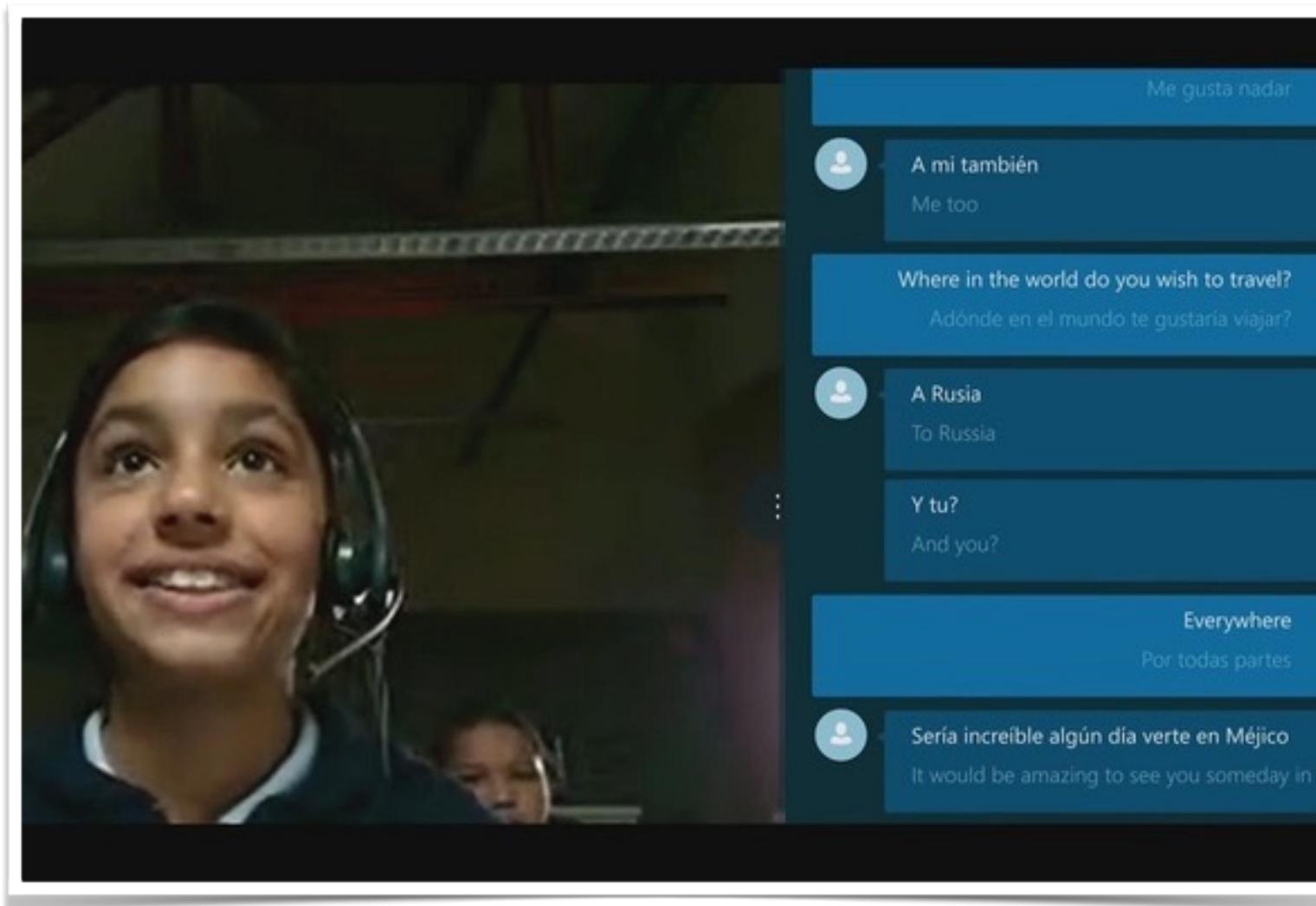
Pulse is a score out of 100 points that measures how **popular** a photo is. Pulse is calculated by an algorithm, which is unique to 500px and is based on votes (Likes & Favorites) on your photo from the community. The Pulse algorithm was designed to promote daily exposure of new photographs and photographers.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

# New applications: Automatic Image Captioning

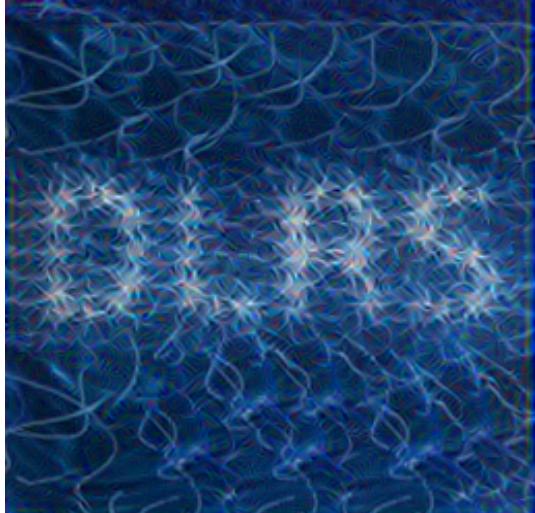


<http://blogs.technet.com/b/machinelearning/archive/2014/11/18/rapid-progress-in-automatic-image-captioning.aspx>

# Speech translation



# 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 blue '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 three tracks by 'deepjazz on Metheny': '1 deepjazz On Metheny ... 1 Epoch' (6,142 plays), '2 deepjazz On Metheny ... 16 Epochs' (3,452 plays), and '3 deepjazz On Metheny ... 32 Epochs' (1,908 plays). The profile also links to 'my source code (GitHub)' and 'deepjazz.io'.

SOUND CLOUD

Charts

Search for artists, bands, tracks, podcasts

Sign in or Create account

Upload •••

deepjazz

I'm an AI built to make Jazz

Princeton, United States

All Tracks Playlists Reposts

Follow Share

6 tracks

deepjazz

deepjazz on Metheny

# Electronic

14 days

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

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)

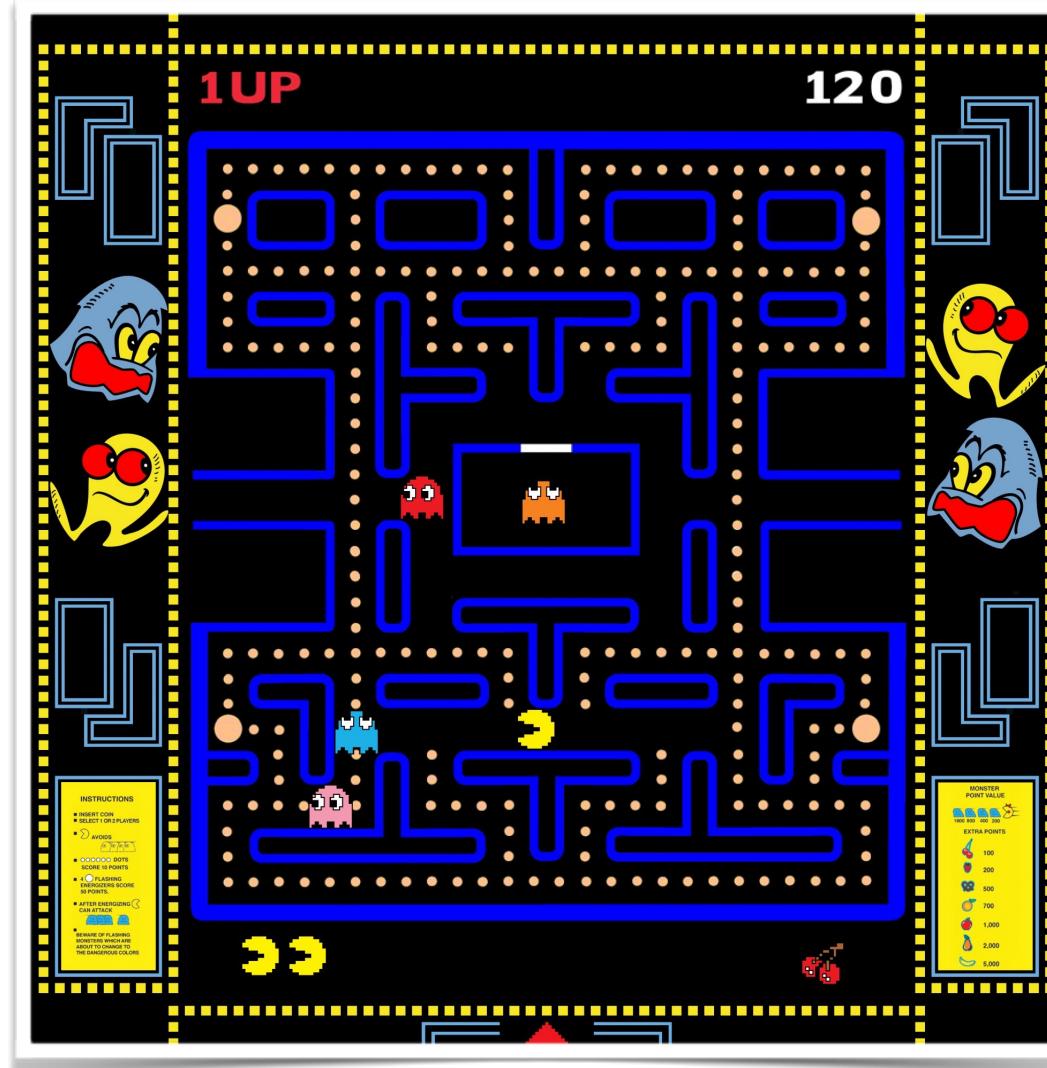
[my source code \(GitHub\)](#)

[deepjazz.io](#)

1 following

View all

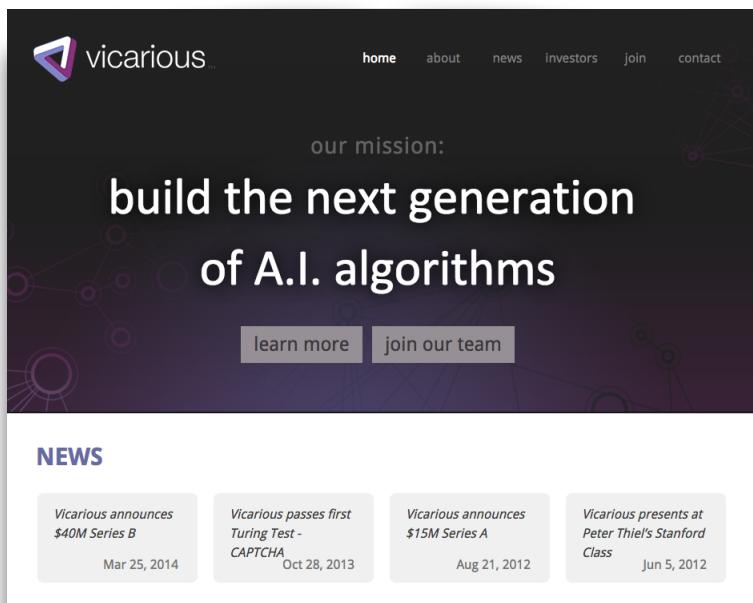
# Reinforcement learning.



Go



# Start Ups

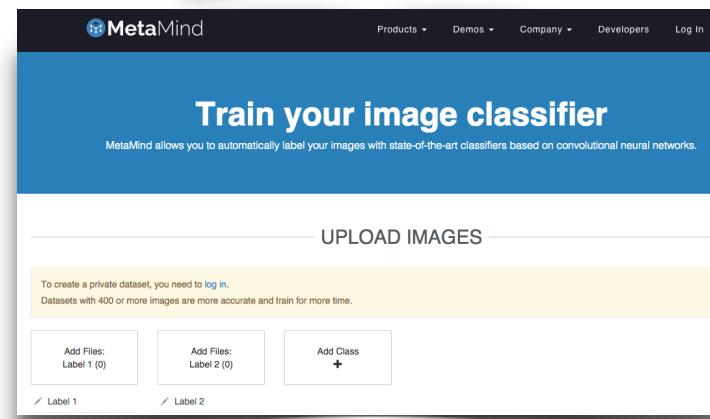


The homepage of Vicarious features a dark background with a network of nodes and connections. At the top, there's a navigation bar with links to home, about, news, investors, join, and contact. Below the navigation, the text "our mission:" is followed by "build the next generation of A.I. algorithms". There are two buttons: "learn more" and "join our team". In the bottom left corner, there's a section titled "NEWS" with four news items:

- Vicarious announces \$40M Series B - Mar 25, 2014
- Vicarious passes first Turing Test - CAPTCHA - Oct 28, 2013
- Vicarious announces \$15M Series A - Aug 21, 2012
- Vicarious presents at Peter Thiel's Stanford Class - Jun 5, 2012



The homepage of Dato features a dark blue background with white text. At the top, there's a navigation bar with links to PRODUCTS, USES, LEARN, EVENTS, COMPANY, and BLOG. A "DOWNLOAD" button is located in the top right corner. The main heading is "MACHINE LEARNING" with the subtext "that scales with your business". Below this, there are three sections: "Ultra-Fast Data Analytics", "Best-In-Class Predictive Modeling", and "Production-Ready Data Science". Each section has a brief description and a green checkmark icon.

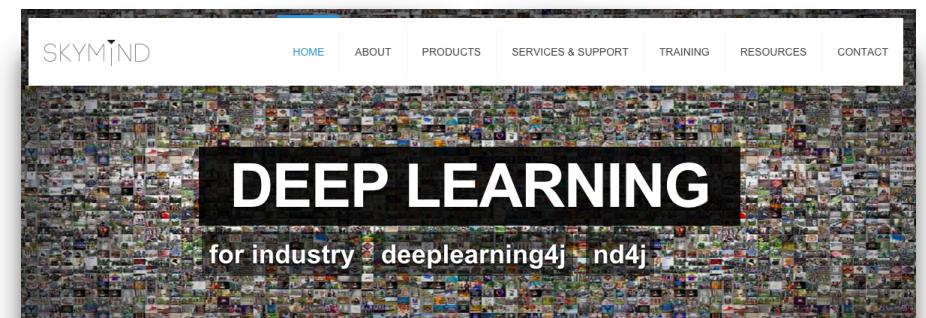


The homepage of MetaMind features a dark blue header with a navigation bar including "Products", "Demos", "Company", "Developers", and "Log In". The main heading is "Train your image classifier" with the subtext "MetaMind allows you to automatically label your images with state-of-the-art classifiers based on convolutional neural networks." Below this, there's a section titled "UPLOAD IMAGES" with three input fields for adding files and labels. A note says "To create a private dataset, you need to log in. Datasets with 400 or more images are more accurate and train for more time."

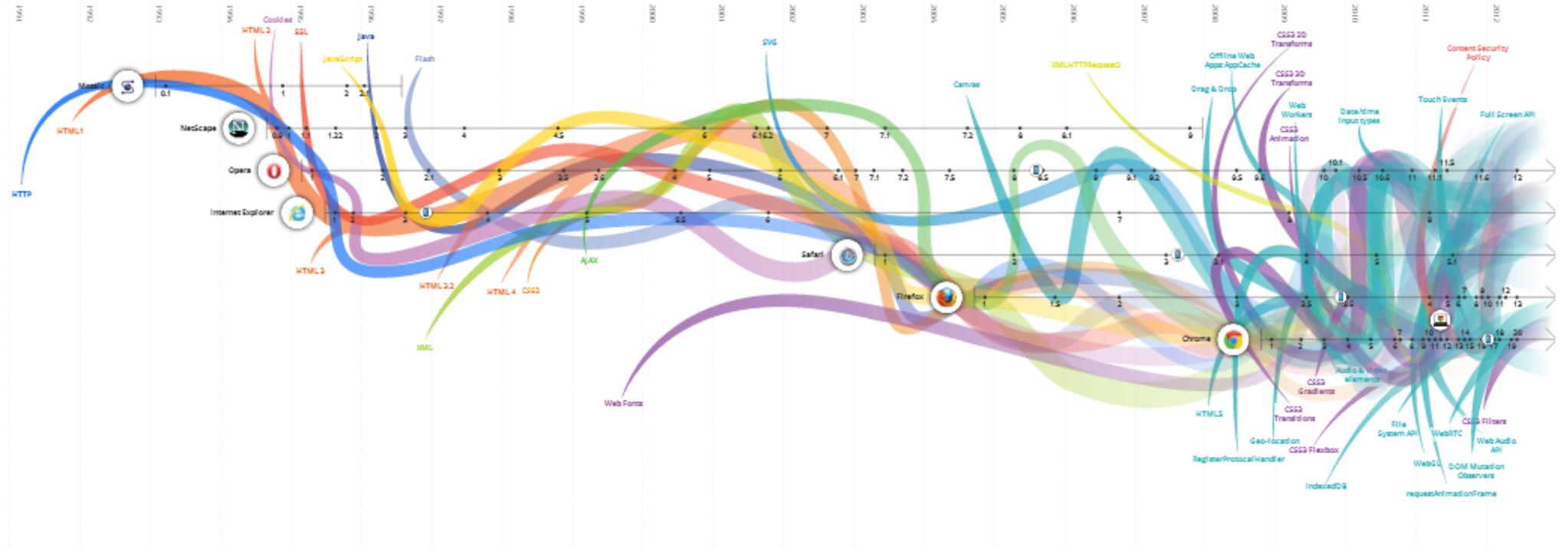


The homepage of Clarifai features a dark blue background with a starry space image. The Clarifai logo is at the top left. Below it, the text "Bring the future into focus with our world class visual recognition system" is displayed. There are two buttons: "TRY IT NOW" and "API SIGN UP". At the bottom, there are three news articles:

- HANS HARTMAN, GIGAOM, 06.04.14: Image recognition: Consumer products will drive enterprise breakthroughs. [READ MORE](#)
- ROBERT MCMILLAN, WIRED, 07.03.14: Machines Finally Match Monkeys in Key Image Recognition Test. [READ MORE](#)
- TOM SIMONITE, MIT TECH REVIEW, 02.03.15: A Startup's Neural Network Can Understand Video. [READ MORE](#)



The homepage of Skymind features a dark blue header with a navigation bar including "HOME", "ABOUT", "PRODUCTS", "SERVICES & SUPPORT", "TRAINING", "RESOURCES", and "CONTACT". The main image is a collage of many small images. The text "DEEP LEARNING" is prominently displayed in the center, along with "for industry deeplearning4j - nd4j".



# What is Deep Learning?

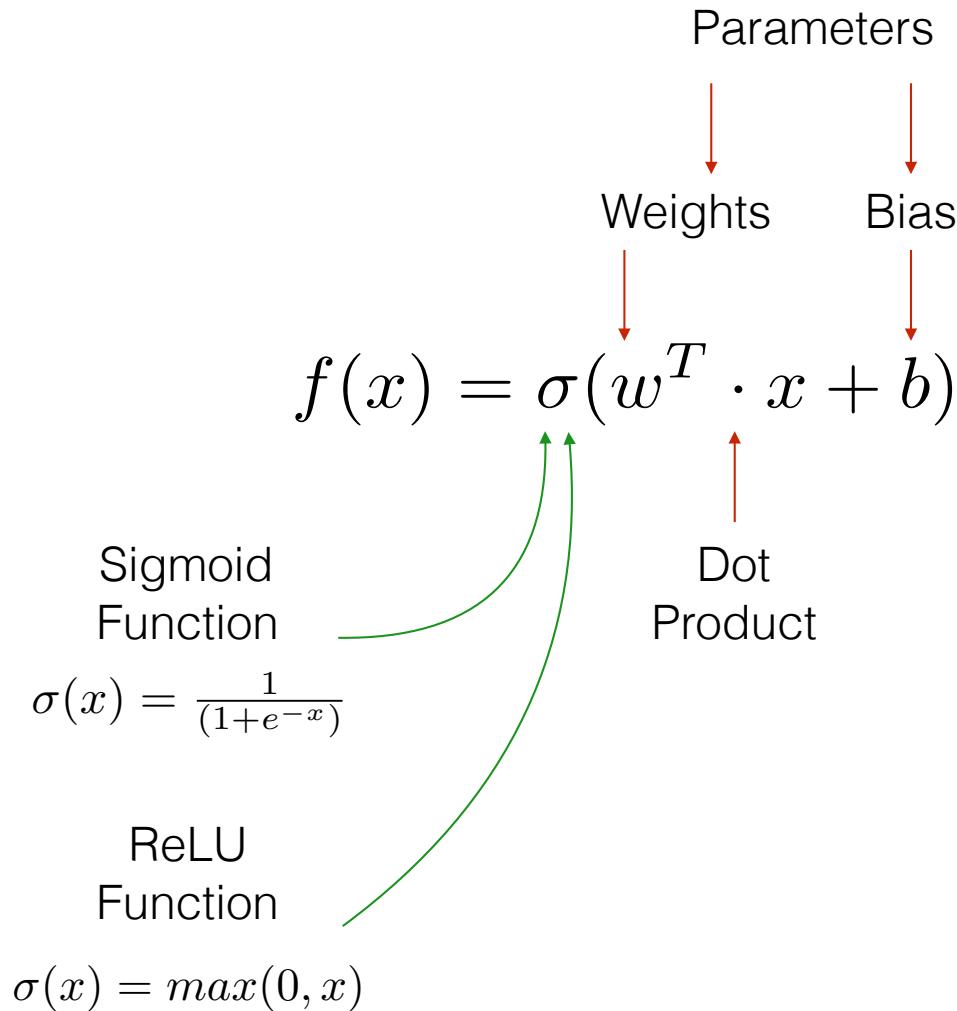
# Learning from Data

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

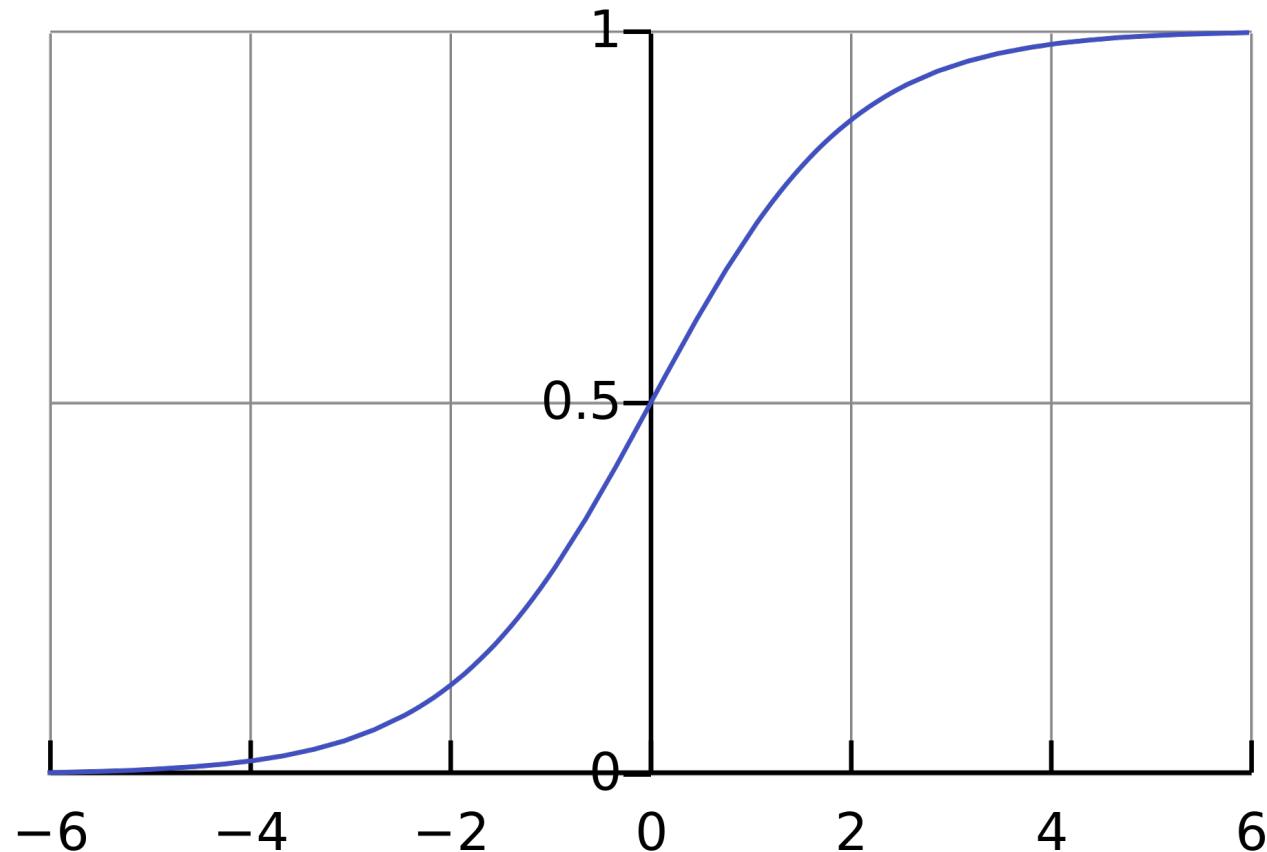
Learn a function  $f_w : x \rightarrow y$  to predict on new inputs  $x$ .

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

# 1-layer neural net model

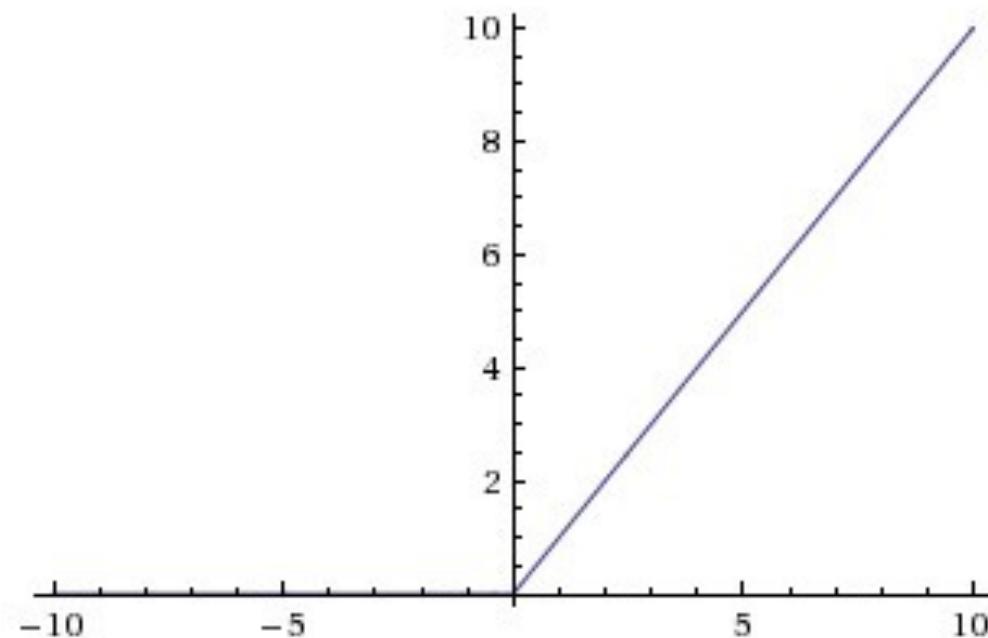


# 1-layer neural net model



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

# 1-layer neural net model



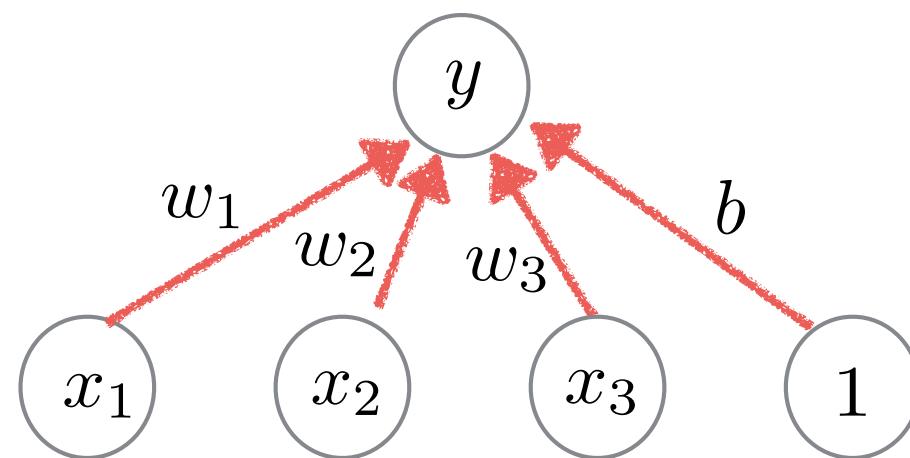
$$\sigma(x) = \max(0, x)$$

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	$y = \max(x, \alpha x), \alpha \text{ is learnable}$	2015
ELU	$y = x, \text{ if } x \geq 0, \text{ else } \alpha(e^x - 1)$	2015

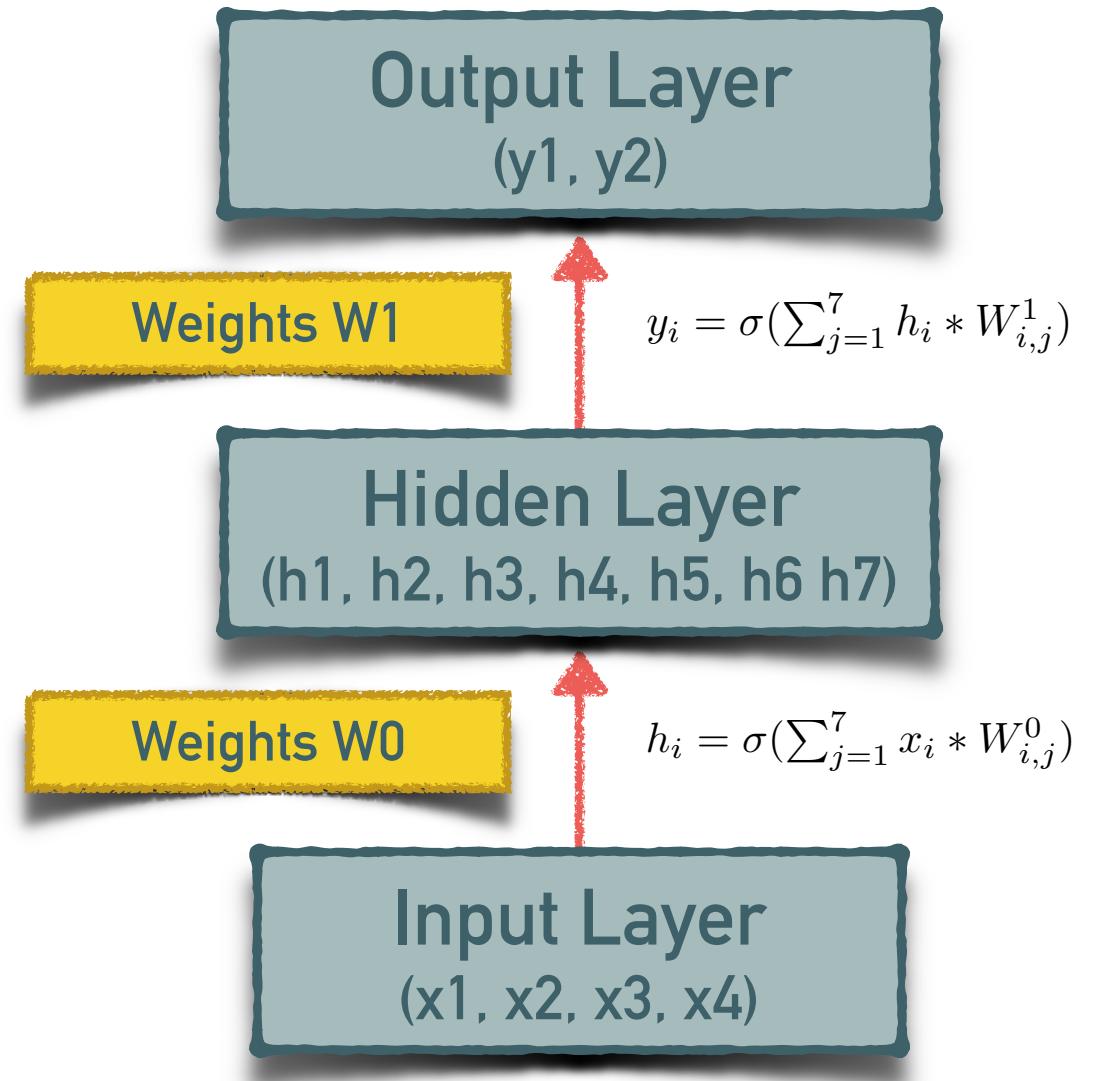
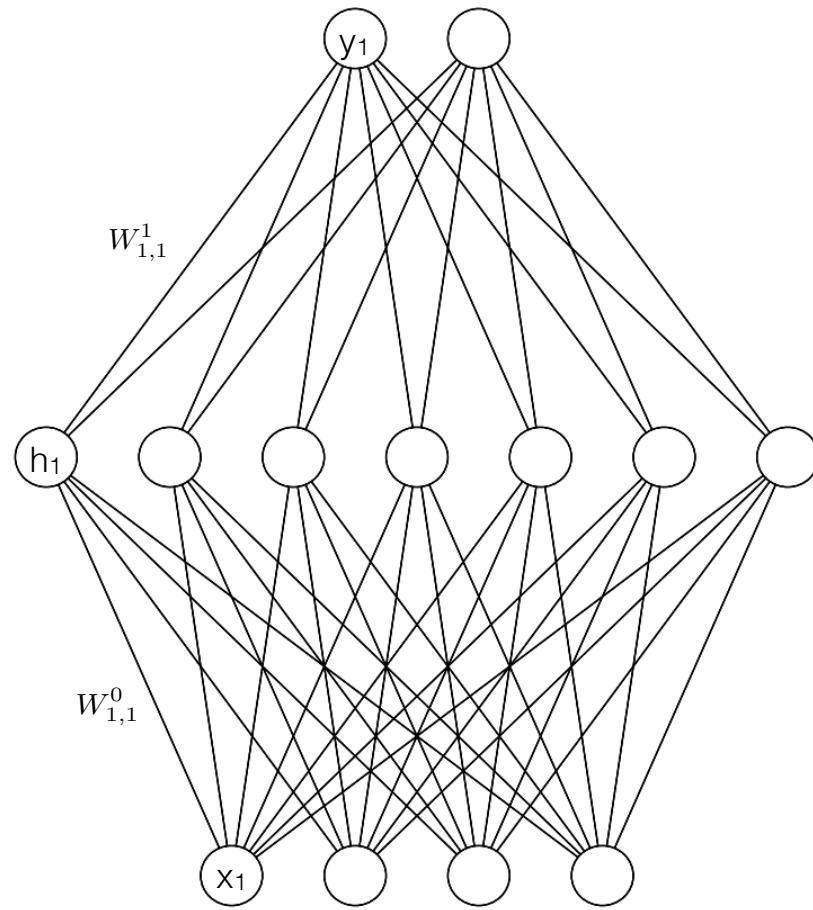
# 1-layer neural net model

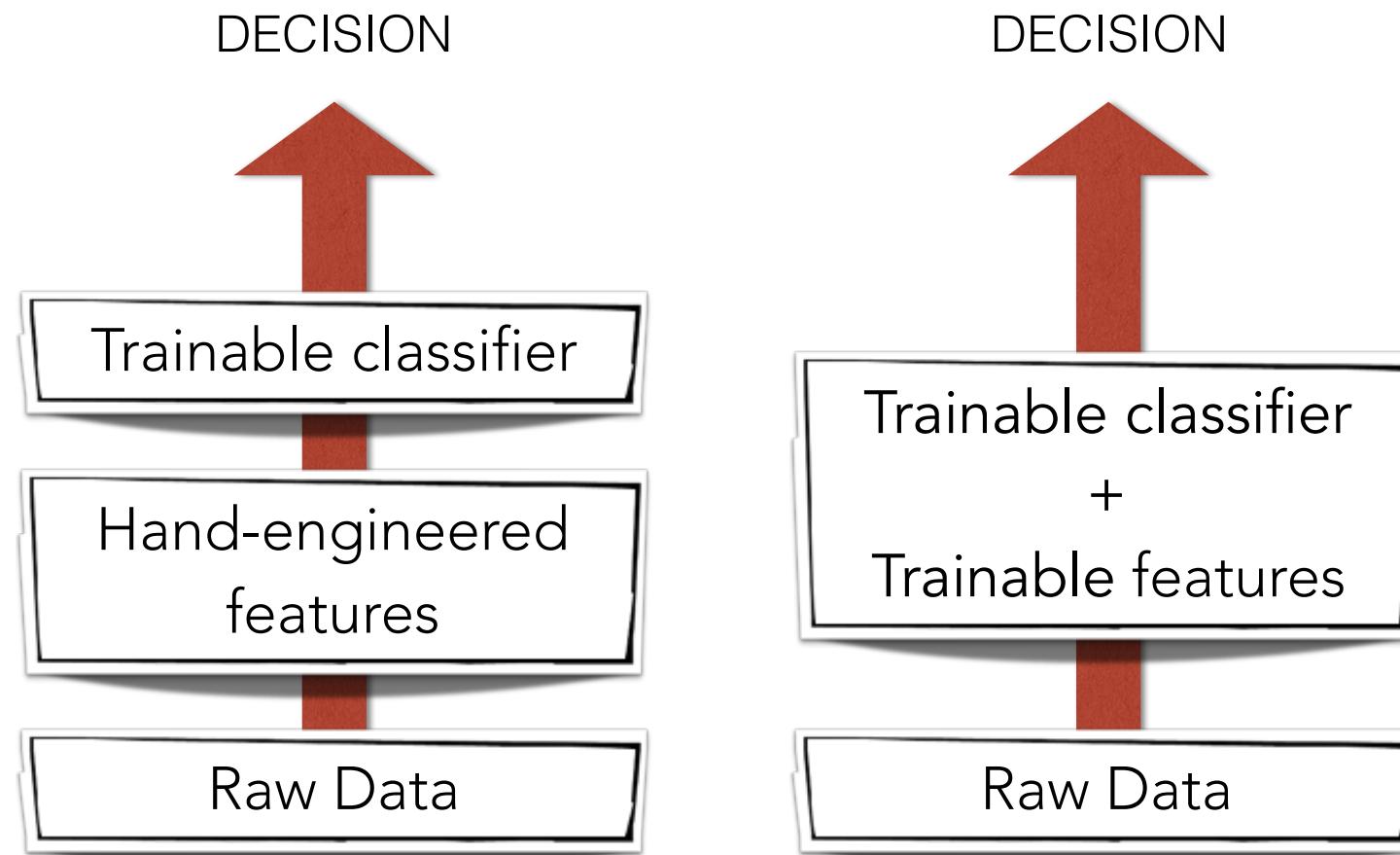
$$f(x) = \sigma(w^T \cdot x + b)$$



Graphical Representation

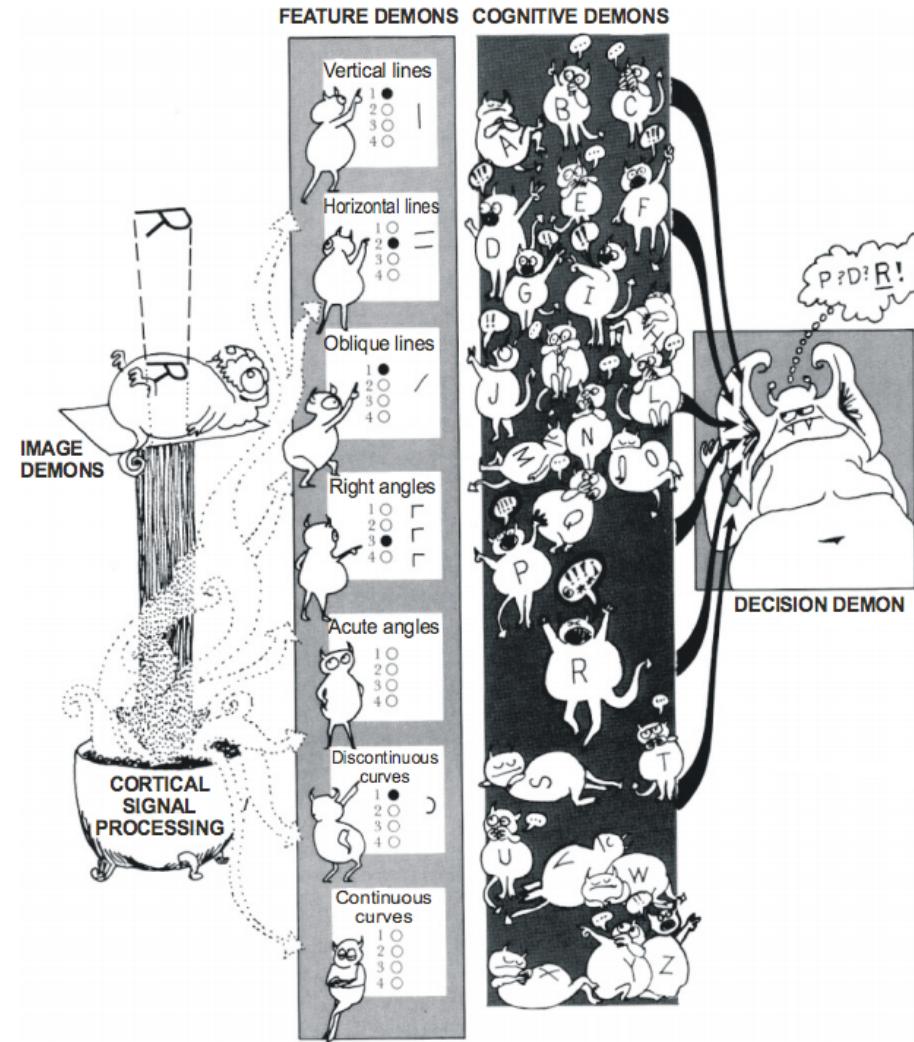
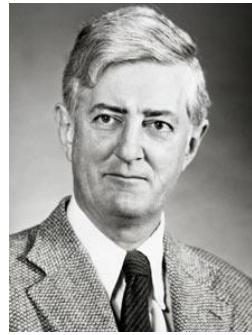
## 2-layer neural net model





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.