Deep Learning

01-Introduction

Marco Piastra

This presentation can be downloaded at:
http://vision.unipv.it/DL

Prologue:
"Deep Learning"?

Deep Learning in the Artificial Intelligence Cosmos

Artificial Intelligence

automated reproduction of human cognitive activities

Machine Learning

*automated extraction of generalized knowledge
from data and experience*

Deep Learning

*automated extraction of generalized knowledge
from data and experience …*

… using multi-layered **representations**

The revolution in AI has been profound, it definitely surprised me, even though I was sitting right there.

Sergey Brin
Google co-founder

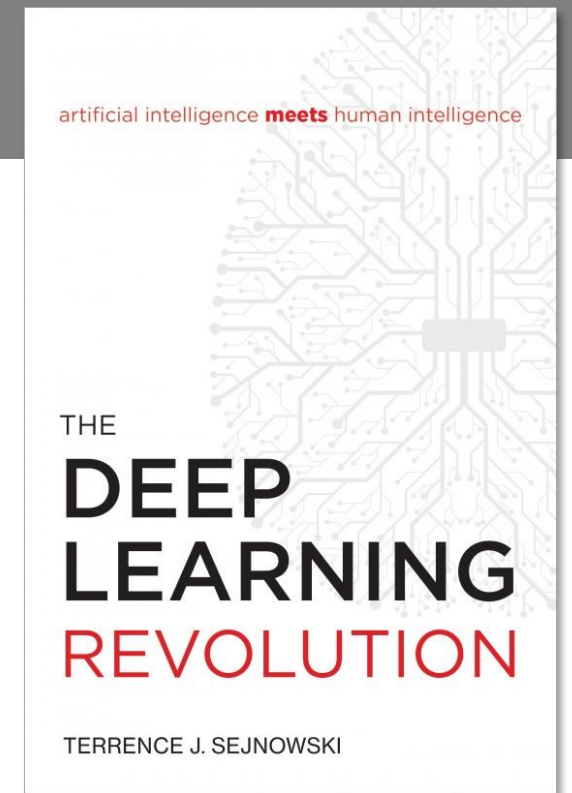■ **Sergey Brin** [Google Co-Founder, January 2017]

"I didn't pay attention to it [i.e. Artificial Intelligence] at all, to be perfectly honest."

*"Having been trained as a computer scientist in the 90s,
everybody knew that AI didn't work.
People tried it, they tried neural nets and none of it worked."*

[Quote and image from https://www.weforum.org/agenda/2017/01/google-sergey-brin-i-didn-t-see-ai-coming/]

The Deep Learning Revolution

artificial intelligence **meets** human intelligence

THE
DEEP
LEARNING
REVOLUTION

TERRENCE J. SEJNOWSKI

- **Terrence J. Sejnowski** [President of the Neural Information Processing (NIPS) Foundation, October 2018]

  "How deep learning—from Google Translate to driverless cars to personal cognitive assistants—is changing our lives and transforming every sector of the economy."
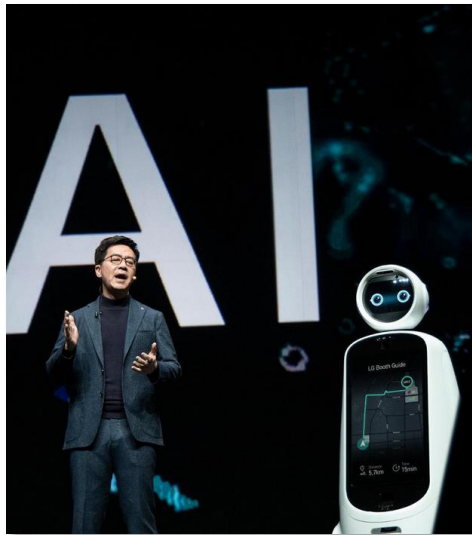
  "AI is now awakening and transforming our world.
  Driving these breakthroughs is the deep learning revolution […]"

[Quote and image from https://mitpress.mit.edu/books/deep-learning-revolution]

Artificial Intelligence Overhype

Forbes

10,693 views | Feb 15, 2019, 02:15pm

3 Reasons AI Is Way Overhyped

Peter Cohan Contributor ⓘ
Markets

1. Many CEOs Are Being Scared Into Caring Too Much About AI

2. There Are Very Few Examples of High Payoff AI Applications

3. Very Few Companies Can Afford or Find Good Uses For AI

[Quote from https://www.forbes.com/sites/petercohan/2019/02/15/3-reasons-ai-is-way-overhyped/#3d3fef8c5a6a/

I'm an AI researcher, and here's what scares me about AI

Rachel Thomas · Follow
Jan 30 · 11 min read

AI is being increasingly used to make important decisions. Many AI experts (including Jeff Dean, head of AI at Google, and Andrew Ng, founder of Coursera and deeplearning.ai) say that warnings about sentient robots are overblown, but other harms are not getting enough attention. I agree. I am an AI researcher, and I'm worried about some of the societal impacts that we're already seeing. In particular, these 5 things scare me about AI:

1. Algorithms are often implemented without ways to address mistakes.

2. AI makes it easier to not feel responsible.

3. AI encodes & magnifies bias.

4. Optimizing metrics above all else leads to negative outcomes.

5. There is no accountability for big tech companies.

[Quote from https://medium.com/@racheltho/im-an-ai-researcher-and-here-is-what-scares-me-about-ai-909a406e4a71]

AI's current hype and hysteria could set the technology back by decades

July 24, 2019 10.11am BST

AI isn't as scary as we imagine. AndreyZH/Shutterstock

The reality of AI is currently very different, particularly when you look at the threat of automation. Back in 2013, researchers estimated that, in the following ten to 20 years, 47% of jobs in the US could be automated. Six years later, instead of a trend towards mass joblessness, we're in fact seeing US unemployment at a historic low.

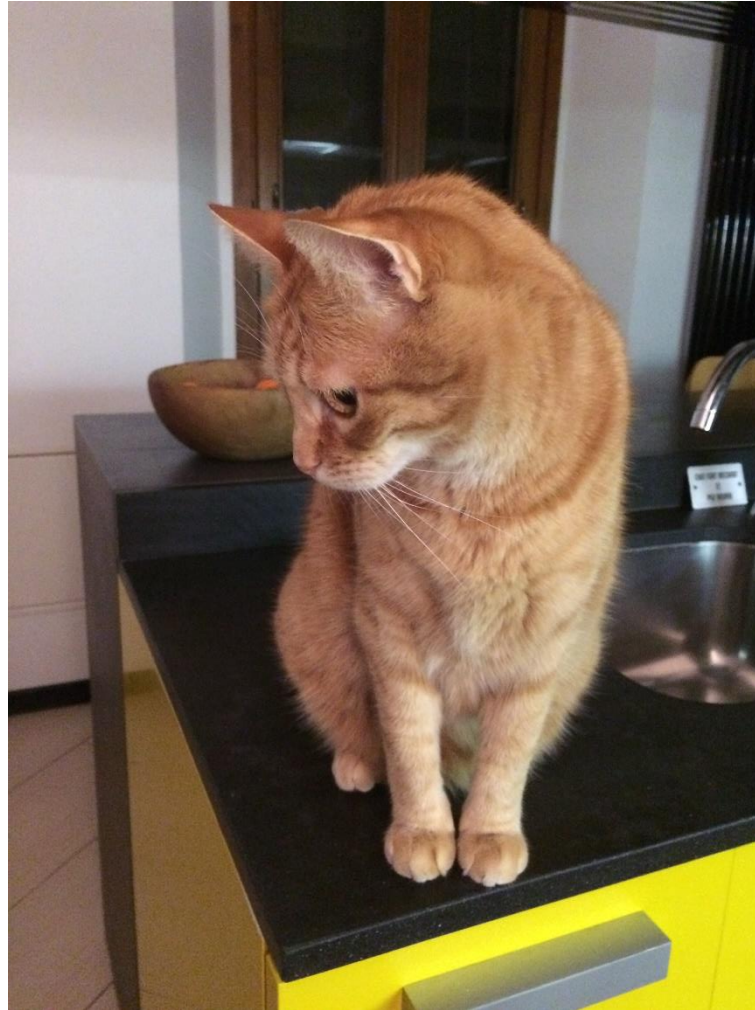Current AI is good at finding patterns in large datasets, and not much else.

[Quote from https://theconversation.com/ais-current-hype-and-hysteria-could-set-the-technology-back-by-decades-120514]

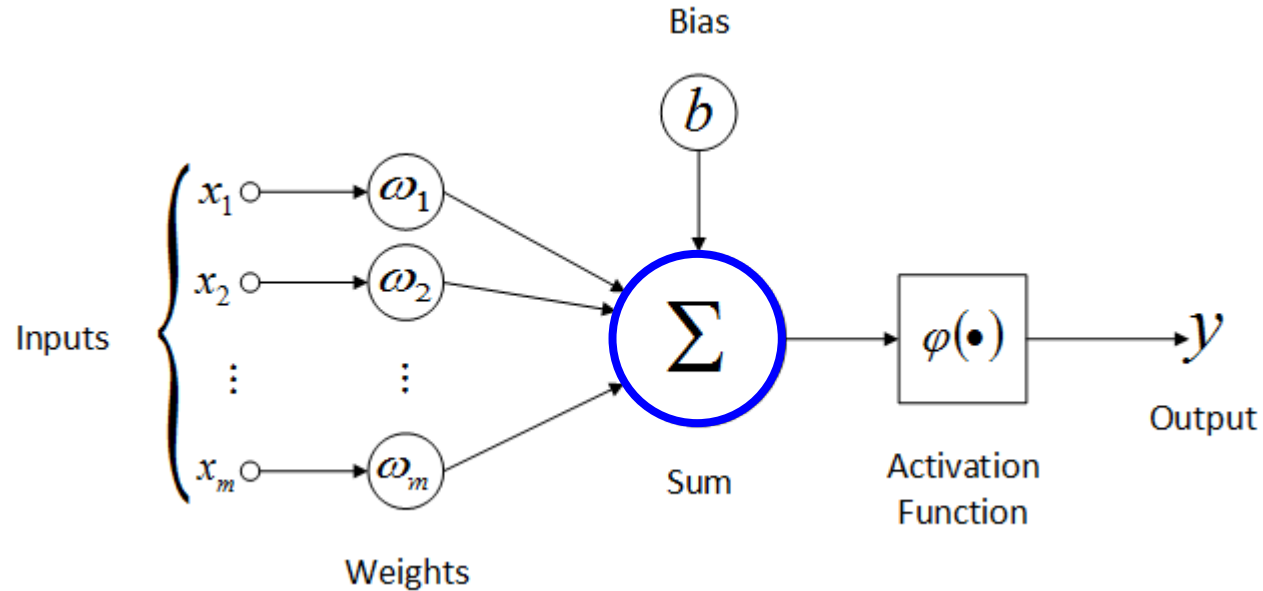# Finding Patterns
in Large Datasets

One Giant Leap for Mankind

*Is there a cat
in this picture?*

[this is *my* cat, Rabarbaro]

Artificial Neural Networks: The Origins

[Rumelhart, D.E., J.L. McClelland 1986]

- **Basic assumption**

  Mental phenomena can be described
  by interconnected networks of simple and often uniform units

  "In our view, people are smarter than today's computers because the brain employs a basic computational architecture that is more suited to deal with a central aspect of the natural information processing tasks that people are so good at."
  [Rumelhart, D.E., J.L. McClelland and the PDP Research Group (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*]
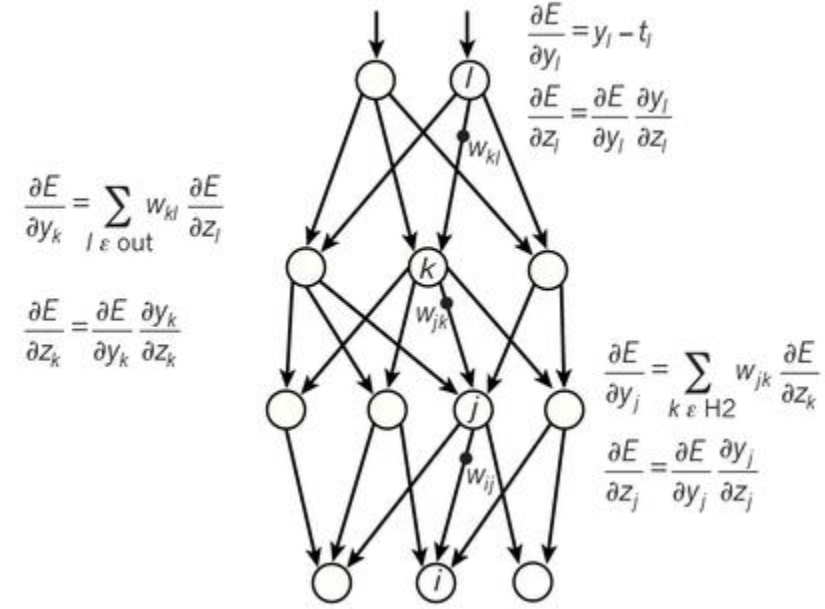
Artificial Neural Network

c

Output units

Hidden units H2

Hidden units H1

Input units
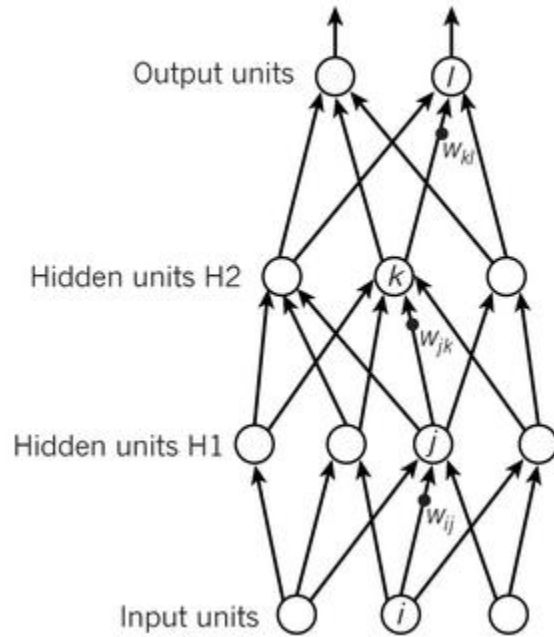
$$y_l = f(z_l)$$
$$z_l = \sum_{k \,\varepsilon\, H2} w_{kl}\, y_k$$

$$y_k = f(z_k)$$
$$z_k = \sum_{j \,\varepsilon\, H1} w_{jk}\, y_j$$

$$y_j = f(z_j)$$
$$z_j = \sum_{i \,\varepsilon\, Input} w_{ij}\, x_i$$

d

Compare outputs with correct
answer to get error derivatives

$$\frac{\partial E}{\partial y_l} = y_l - t_l$$
$$\frac{\partial E}{\partial z_l} = \frac{\partial E}{\partial y_l}\frac{\partial y_l}{\partial z_l}$$

$$\frac{\partial E}{\partial y_k} = \sum_{l \,\varepsilon\, out} w_{kl}\frac{\partial E}{\partial z_l}$$

$$\frac{\partial E}{\partial z_k} = \frac{\partial E}{\partial y_k}\frac{\partial y_k}{\partial z_k}$$

$$\frac{\partial E}{\partial y_j} = \sum_{k \,\varepsilon\, H2} w_{jk}\frac{\partial E}{\partial z_k}$$
$$\frac{\partial E}{\partial z_j} = \frac{\partial E}{\partial y_j}\frac{\partial y_j}{\partial z_j}$$

[figure from LeCun, Bengio, Hinton, *Deep Learning*, Nature 521, 2015]

Function approximation

Basically, this is what a 'classical' artificial neural network does

Supervised learning

The parameters (i.e. *weights*) are "learnt" from a dataset of inputs and expected outputs pairs

Incremental optimization — a.k.a. "*backward propagation*"

Weights are progressively corrected to reduce *the difference* between actual and expected outputs

Artificial Neural Networks

- **From *shallow* to *deep* networks**

  A feed-forward neural network with one hidden layer

  $$\tilde{y} = \boldsymbol{w} \cdot g(\boldsymbol{W}^{[1]}\boldsymbol{x} + \boldsymbol{b}^{[1]}) + b$$

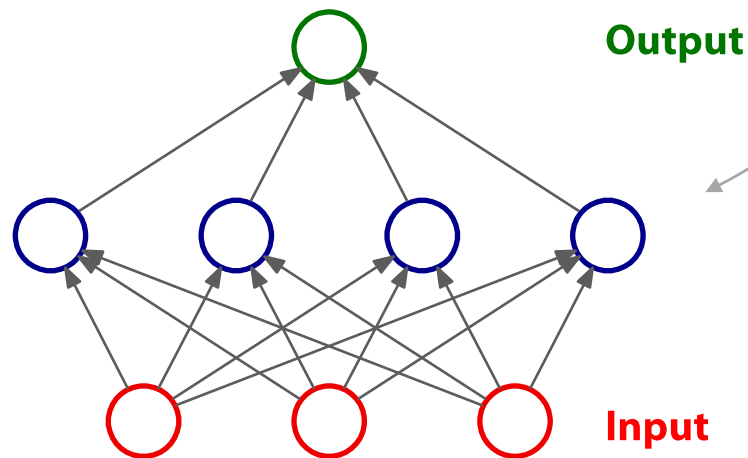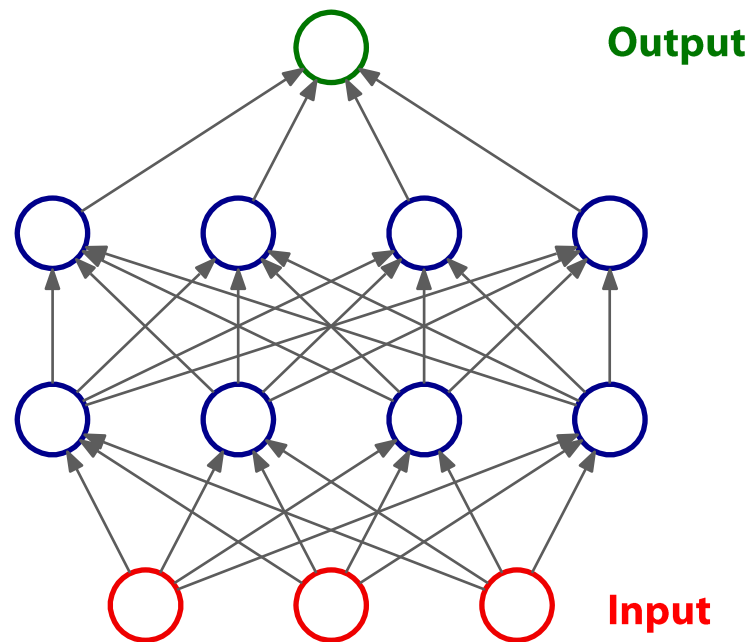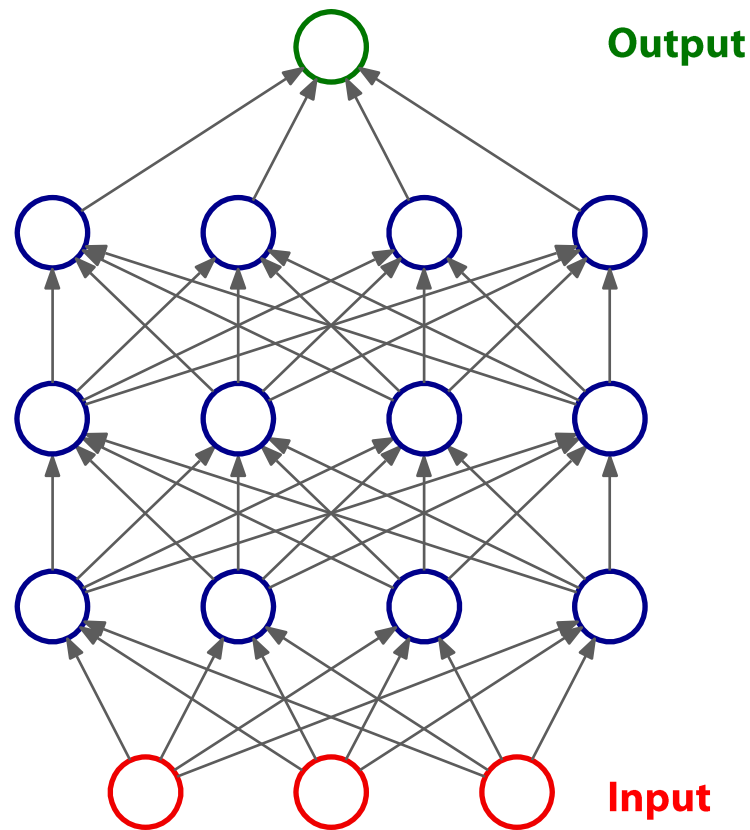 *Deep Learning systems (e.g. TensorFlow, PyTorch) use this representation*

  It can approximate any target function

  $$y = f^*(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^d$$

  (given enough units and proper parameters)

  The two representations are equivalent

  Output

  Input

Artificial Neural Networks

- **From *shallow* to *deep* networks**

  A feed-forward neural network with two hidden layers

  $$\tilde{y} = \boldsymbol{w} \cdot g(\boldsymbol{W}^{[2]} g(\boldsymbol{W}^{[1]} \boldsymbol{x} + \boldsymbol{b}^{[1]}) + \boldsymbol{b}^{[2]}) + b$$

- **From *shallow* to *deep* networks**

  A feed-forward neural network with three hidden layers

  $$\tilde{y} = \boldsymbol{w} \cdot g(\boldsymbol{W}^{[3]} g(\boldsymbol{W}^{[2]} g(\boldsymbol{W}^{[1]} \boldsymbol{x} + \boldsymbol{b}^{[1]}) + \boldsymbol{b}^{[2]}) + \boldsymbol{b}^{[3]}) + b$$
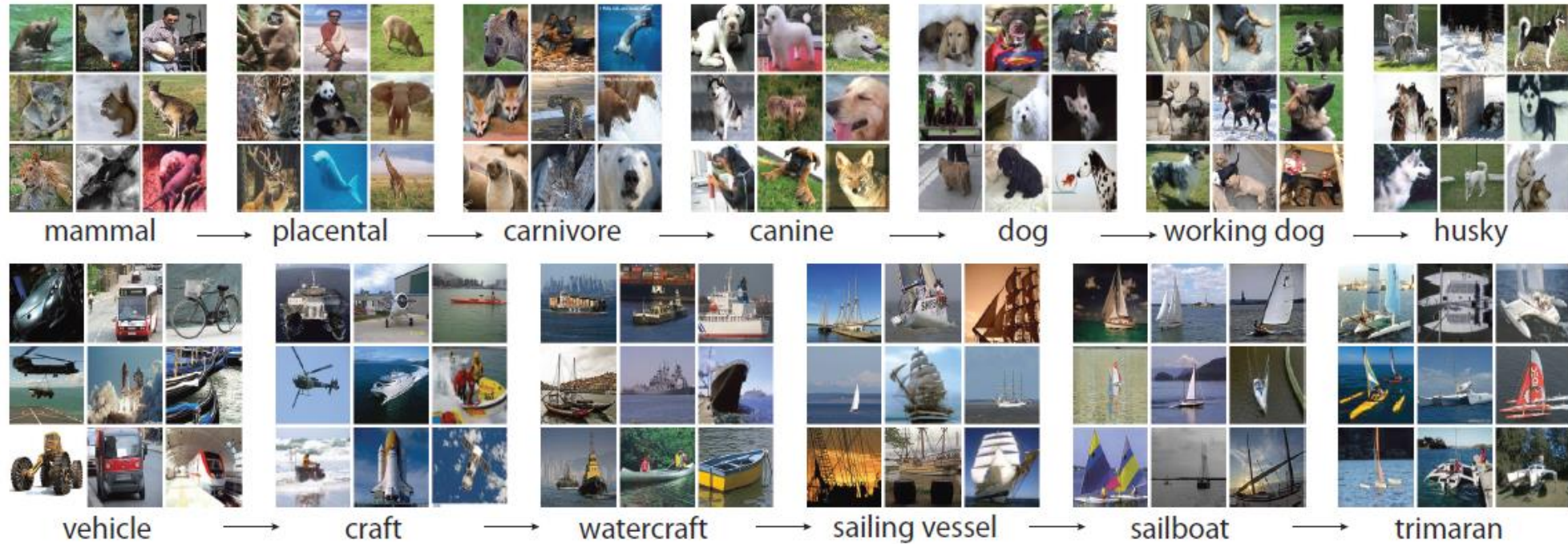
Well, it's just a <u>function</u> anyway ...

- The ImageNet Large Scale Visual Recognition Challenge
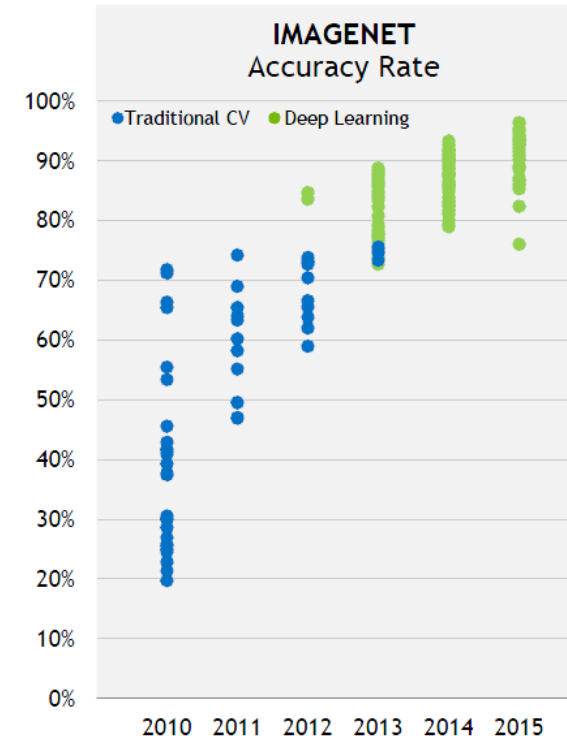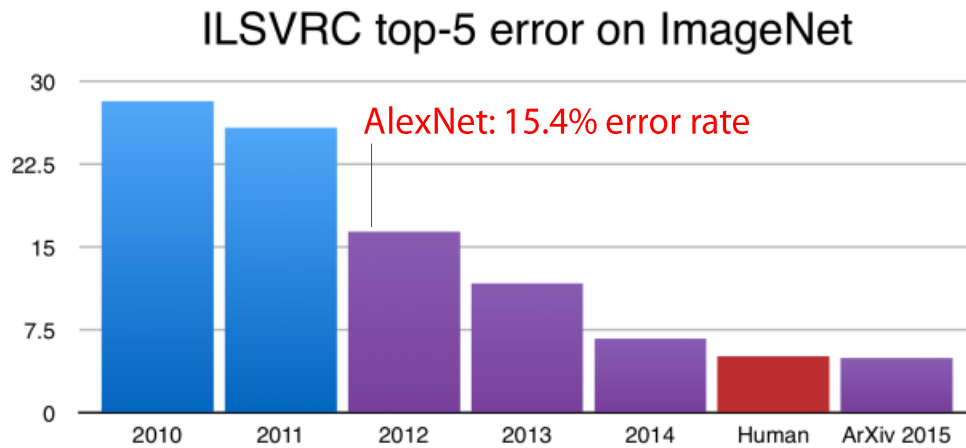

1,461,406 full resolution images

Complex and multiple textual annotation,
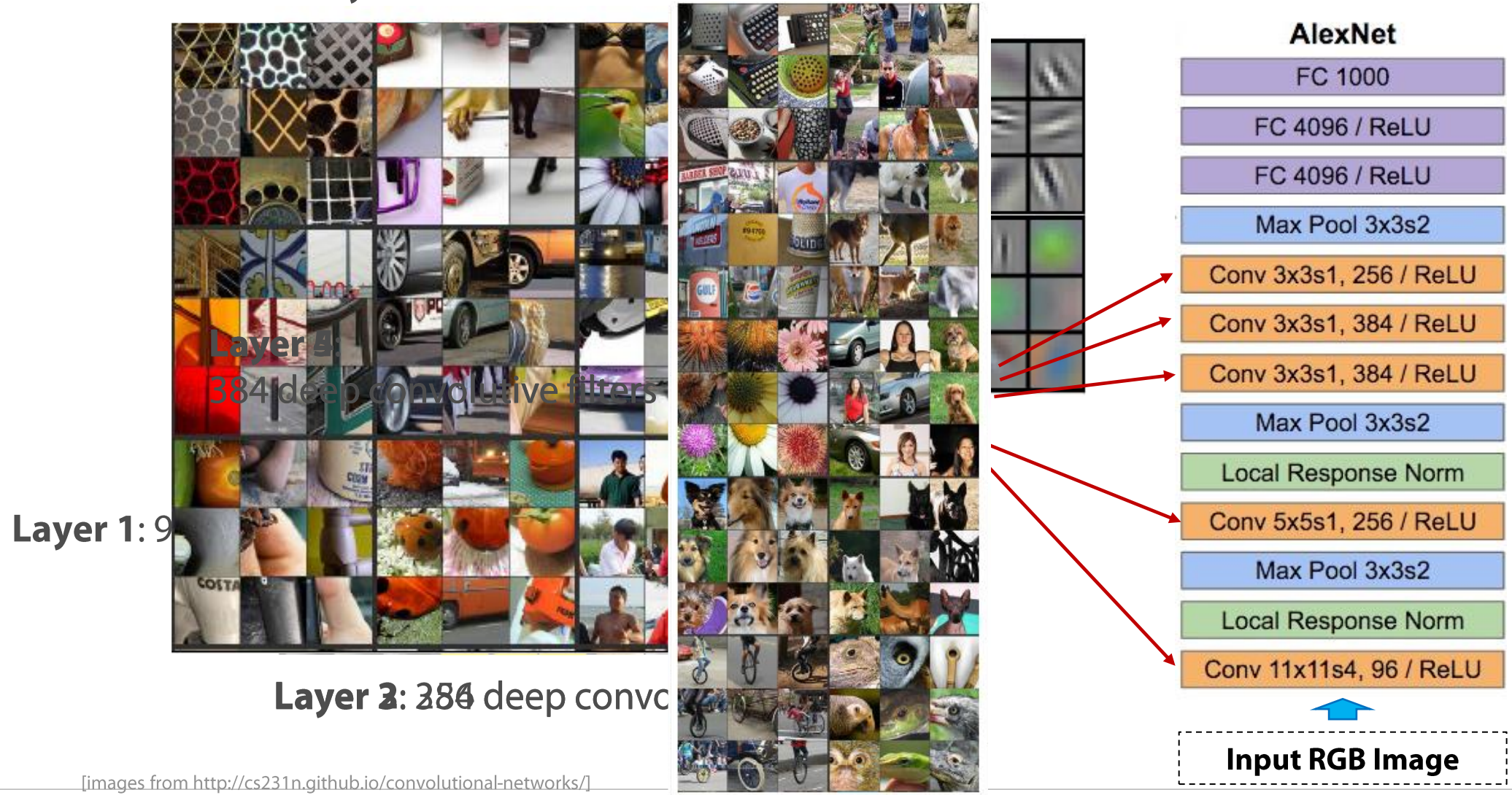hierarchy of 1000 object classes along several dimensions

The image classification challenge was run annually from 2010 to 2017

[figures from www.nvidia.com]

■ The ImageNet Large Scale Visual Recognition Challenge

ILSVRC top-5 error on ImageNet

AlexNet: 15.4% error rate

IMAGENET
Accuracy Rate

1,461,406 full resolution images

Complex and multiple textual annotation,
hierarchy of 1000 object classes along several dimensions

The image classification challenge was run annually from 2010 to 2017

[figures from www.nvidia.com]

Deep Convolutional Neural Networks (DCNN)

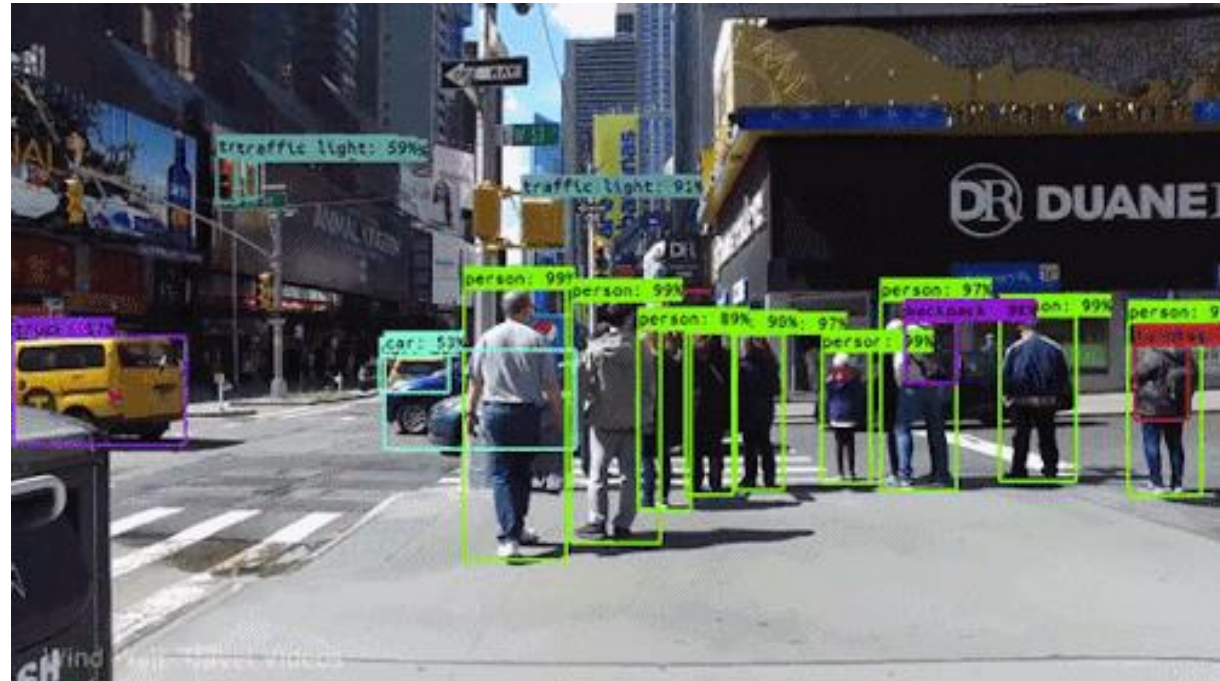- **AlexNet** [Krizhevsky, Sutskever & Hinton, 2012]

AlexNet

| FC 1000 |
| FC 4096 / ReLU |
| FC 4096 / ReLU |
| Max Pool 3x3s2 |
| Conv 3x3s1, 256 / ReLU |
| Conv 3x3s1, 384 / ReLU |
| Conv 3x3s1, 384 / ReLU |
| Max Pool 3x3s2 |
| Local Response Norm |
| Conv 5x5s1, 256 / ReLU |
| Max Pool 3x3s2 |
| Local Response Norm |
| Conv 11x11s4, 96 / ReLU |

Input RGB Image

Layer 4
384 deep convolutive filters

Layer 1: 9

Layer 2: 384 deep convo

Layer 3: 384

[images from http://cs231n.github.io/convolutional-networks/]

- *Deep Convolutional Neural Networks have evolved since then ...*

Now these system can identify objects and persons from videos, in real time

NOTE:
According to the recent
EU Proposal for a Regulation
*about AI, **remote biometric***
***identification** (RBI)*
in public places will require
a special authorization

Image from: https://sgu.ac.id/id/computer-vision-artificial-intelligence-why-is-it-important/

- *Deep Convolutional Neural Networks have evolved since then …*

They can perform a complete scene analysis, from videos, in real time

At present, DCNN work on a frame-by-frame basis

Image from: https://sgu.ac.id/id/computer-vision-artificial-intelligence-why-is-it-important/

Well, it's just a <u>function</u> anyway (contd.)…

- **Mastering the game of Go with deep neural networks and tree search** [2016, D. Silver, et al. (22 authors), http://www.nature.com/nature/journal/v529/n7587/full/nature16961.html]

There are more possible positions in Go than there are atoms in the universe

Sophisticated machine-learning techniques

Strategy selection via Monte Carlo Tree Search (MCTS)

Deep neural networks (trained on human matches) for both guidance and learning

Adversarial self-training:
playing again itself and improving via reinforcement learning

Super-human player?

On March 2016, AlphaGo won 4-1 against the legendary Lee Sedol,
the top Go player in the world over the past decade

- **A Deep Neural Network learns a <u>policy</u>**
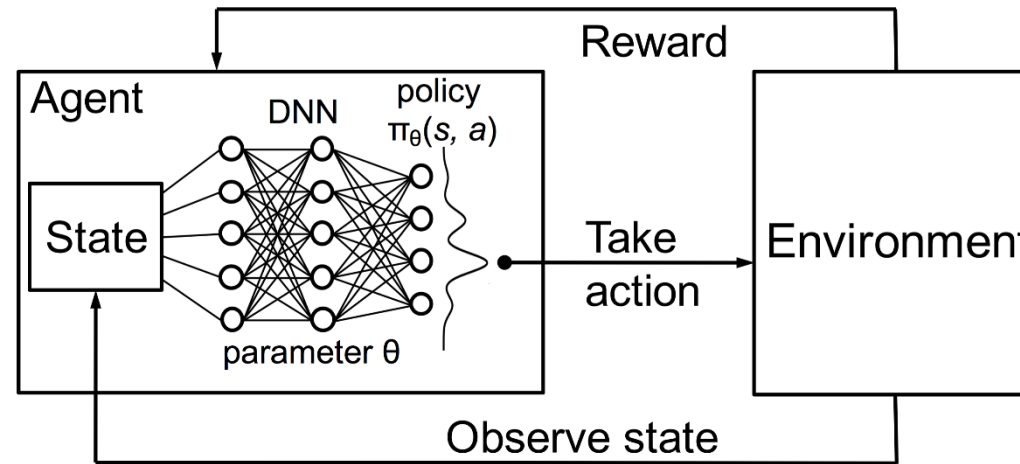
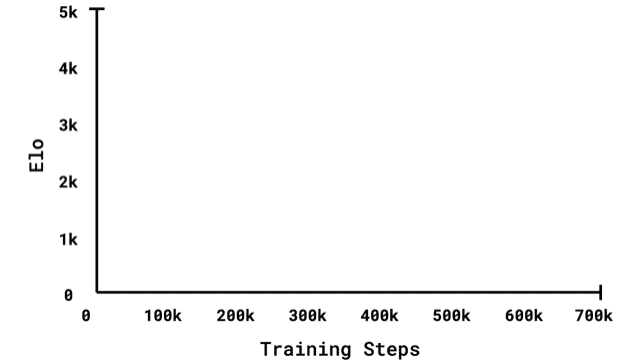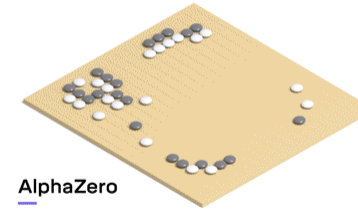The agent interacts with an environment (it could be a copy of itself)

It selects an **action** in each **state** and receives a **reward** (possibly deferred) as a function of the results obtained

The DRL system optimizes its <u>policy</u>

Image from: https://deepmind.com/blog/article/alphazero-shedding-new-light-grand-games-chess-shogi-and-go

*AlphaGo is heavily reliant
on the experience of human players*

- **AlphaZero learns by itself**
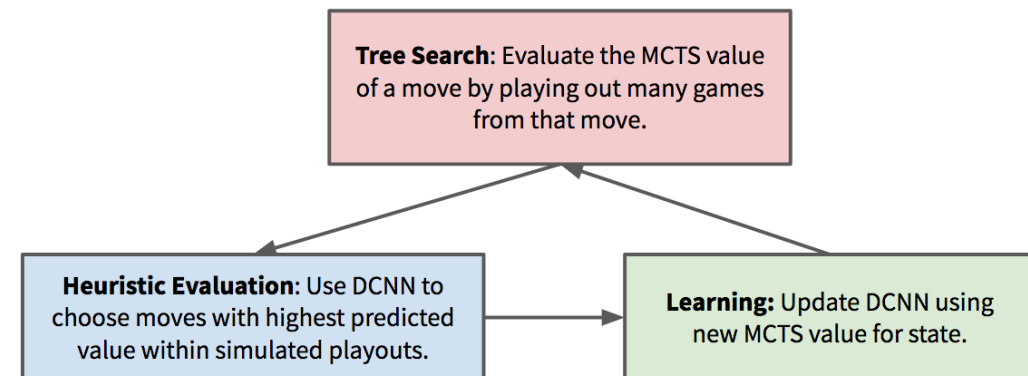 [2018, D. Silver, et al. (13 authors), https://science.sciencemag.org/content/362/6419/1140.full]

Basic Knowledge Only
It just knows the basic rules of the games

Learning via Self-Play
It plays against a (frozen) copy of itself

MCTS and DCNN in a closed loop

Tree Search: Evaluate the MCTS value of a move by playing out many games from that move.

Heuristic Evaluation: Use DCNN to choose moves with highest predicted value within simulated playouts.

Learning: Update DCNN using new MCTS value for state.

Image from: https://nikcheerla.github.io/deeplearningschool/2018/01/01/AlphaZero-Explained/

Image from: https://deepmind.com/blog/article/alphazero-shedding-new-light-grand-games-chess-shogi-and-go

Amount of Search per Decision

Human Grandmaster
100's of moves

AlphaZero
10,000's of moves

State-of-the-Art Chess Engines
10,000,000's of moves

- **AlphaZero uses much less 'brute force' search**
    - When playing, the search process is driven by its neural network
        - *It acts like a memory of past experiences*
    - While training, it learns through a huge amount of self-playing
        - *But it is a faster learner than Alpha Go*

AlphaFold (2020)

▪ **Predicting how proteins fold**

Several Neural Networks
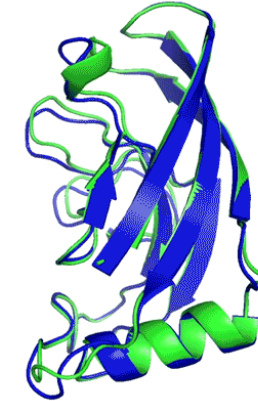
Trained separately, working together

A big step ahead

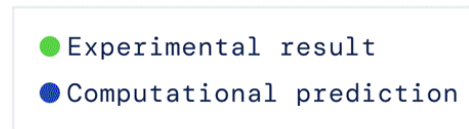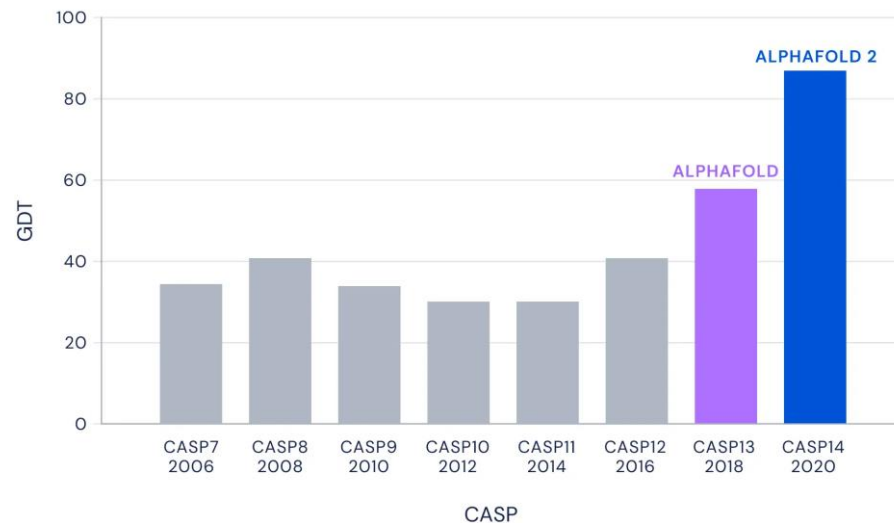*Substantial improvement
over the best actual systems yet*
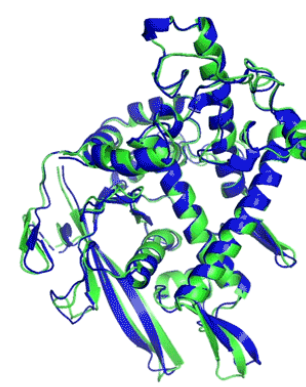
T1037 / 6vr4
90.7 GDT
(RNA polymerase domain)

T1049 / 6y4f
93.3 GDT
(adhesin tip)

● Experimental result
● Computational prediction

It's open source

In the version used for the CASP13 challenge

https://github.com/deepmind/deepmind-research/tree/master/alphafold_casp13

GPT-3 (2020)

Support The Guardian

Available for everyone, funded by readers

Contribute → Subscribe →

⊖ My account ⌄ The Guardian

Opinion Artificial intelligence (AI)

A robot wrote this entire article. Are you scared yet, human?
GPT-3

We asked GPT-3, OpenAI's powerful new language generator, to write an essay for us from scratch. The assignment? To convince us robots come in peace

- For more about GPT-3 and how this essay was written and edited, please read our editor's note below

Tue 8 Sep 2020 09.45 BST

70,298 | 1,188

GPT-3 (2020)

Fortier fed GPT-3 a strange prompt: "Below is a transcript from an interview where Barack Obama explained why he was banned from Golden Corral for life." The system then filled in the rest of the interview, running with the concept that Obama had been banned from an all-you-can-eat buffet.

Obama: Yes. It's true. I am no longer allowed in Golden Corral.

**Interviewer**: Is this because of your extensive shrimp-n-crab legs policy?

Obama: Absolutely.

Interviewer: What is your extensive shrimp-n-crab legs policy?

Obama: Oh, well, in brief, they were offering an all-you-can-eat shrimp-n-crab leg buffet, and I did not hesitate. After I ate so much shrimp and crab that my stomach hurt, I would quietly sneak in and throw more shrimp and crab onto my plate. I did this over and over again until I had cleaned out the buffet and was full of shrimp-n-crab.
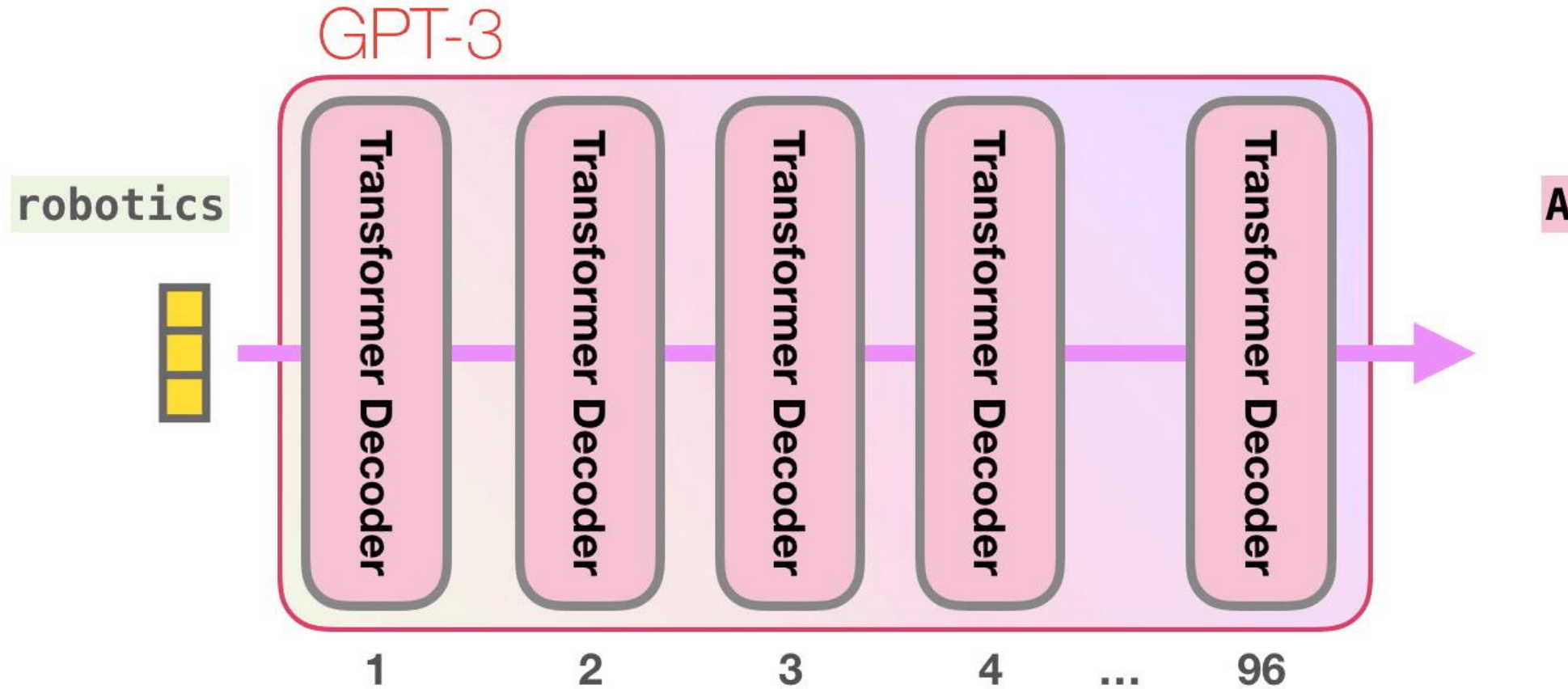
Image from http://jalammar.github.io/how-gpt3-works-visualizations-animations/
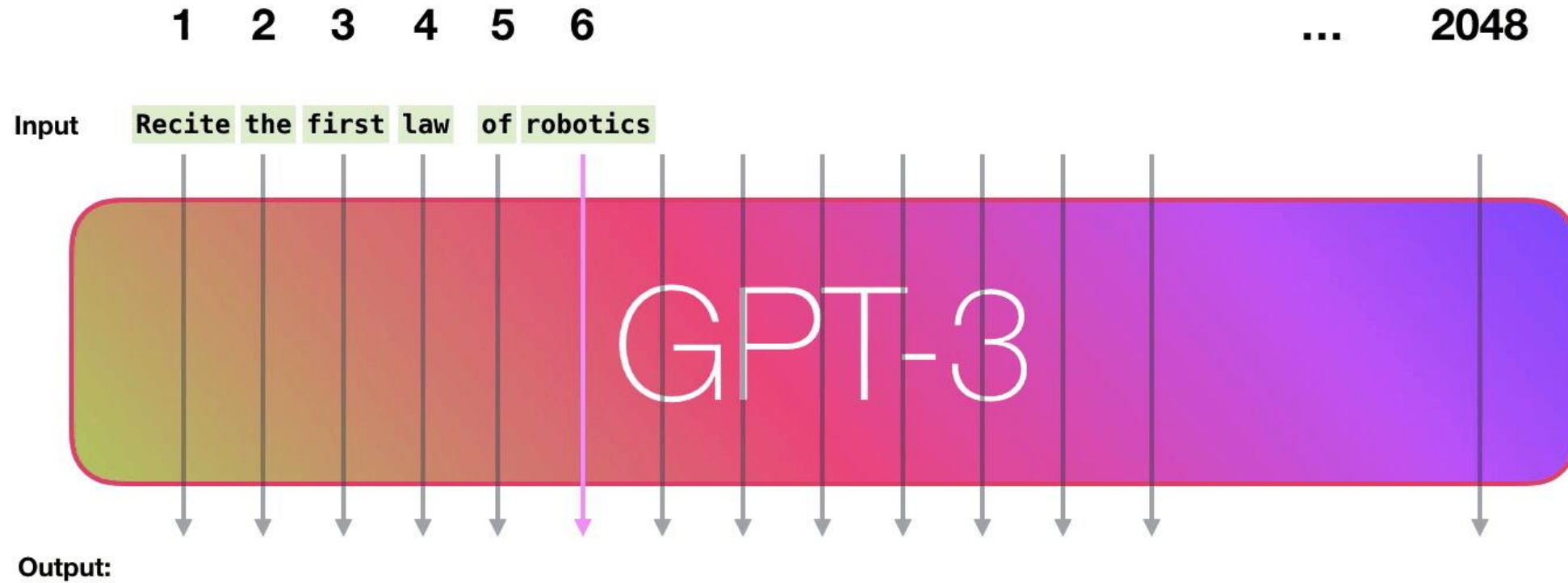
One of the biggest Neural Networks yet

GPT-3 has 175 Billion parameters
(AlexNet has 64 Million)

Solving coding problems using transformers

① **Problem (input)**

D.Backspace

You are given two strings s and t, both consisting of lowercase English letters. You are going to type the string s character by character, from the first character to the last one.

When typing a character, instead of pressing the button corresponding to it, you can press the "Backspace" button. It deletes the last character you have typed among those that aren't deleted yet (or does nothing if there are no characters in the current string). For example, if s is "abcbd" and you press Backspace instead of typing the first and the fourth characters, you will get the string "bd" (the first press of Backspace deletes no character, and the second press deletes the character 'c'). Another example, if s is "abcaa" and you press Backspace instead of the last two letters, then the resulting text is "a".

Your task is to determine whether you can obtain the string t, if you type the string s and press "Backspace" instead of typing several (maybe zero) characters of s.

Input

The first line contains a single integer q ($1 \le q \le 10^5$) — the number of test cases.

The first line of each test case contains the string s ($1 \le |s| \le 10^5$). Each character of s is a lowercase English letter.

The second line of each test case contains the string t ($1 \le |t| \le 10^5$). Each character of t is a lowercase English letter.

It is guaranteed that the total number of characters in the strings over all test cases does not exceed $2 \cdot 10^5$.

Output

For each test case, print "YES" if you can obtain the string t by typing the string s and replacing some characters with presses of "Backspace" button, or "NO" if you cannot.

You may print each letter in any case (YES, yes, Yes will all be recognized as positive answer, NO, no and nO will all be recognized as negative answer).

```
Input

4
ababa
ba
ababa
bb
aaa
aaaa
aababa
ababa
```

```
Output

YES
NO
NO
YES
```

Note

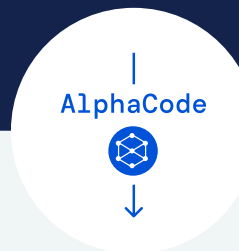Consider the example test from the statement.

In order to obtain "ba" from "ababa", you may press Backspace instead of typing the first and the fourth characters.

There's no way to obtain "bb" while typing "ababa".

There's no way to obtain "aaaa" while typing "aaa".

In order to obtain "ababa" while typing "aababa", you have to press Backspace instead of typing the first character, then type all the remaining characters.

AlphaCode

② **Solution (output)**

① AlphaCode is presented with a problem, in this case to figure out if it's possible to convert one phrase to another by pressing backspace instead of typing some letters.

Solving coding problems using transformers

AlphaCode

② **Solution (output)**

```python
t=int(input())
for i in range(t):
    s=input()
    t=input()
    a=[]
    b=[]
    for j in s:
        a.append(j)
    for j in t:
        b.append(j)
    a.reverse()
    b.reverse()
    c=[]
    while len(b)!=0 and len(a)!=0:
        if a[0]==b[0]:
            c.append(b.pop(0))
            a.pop(0)
        elif a[0]!=b[0] and len(a)!=1:
            a.pop(0)
            a.pop(0)
        elif a[0]!=b[0] and len(a)==1:
            a.pop(0)
    if len(b)==0:
        print("YES")
    else:
        print("NO")
```

First AlphaCode reads the two phrases.

If the letters at the end of both phrases don't match, the last letter must be deleted. If they do match we can move onto the second last letter and repeat.

Backspace deletes two letters. The letter you press backspace instead of, and the letter before it.

If we've matched every letter, it's possible and we output that.

② AlphaCode reads the whole problem statement and produces code, analogous to how a human would approach the problem by reading it, coding a solution, and submitting.

An aside:
There might be more patterns about us
than we may want to admit...

Cambridge Analytica: how 50m Facebook records were hijacked

1
Approx. 320,000 US voters ('seeders') were paid $2-5 to take a detailed personality/political test that required them to log in with their Facebook account

2
The app also collected data such as likes and personal information from the test-taker's Facebook account ...

3
The personality quiz results were paired with their Facebook data – such as likes – to seek out psychological patterns

4
Algorithms combined the data with other sources such as voter records to create a superior set of records (initially 2m people in 11 key states*), with hundreds of data points per person

Friends' data

User's data
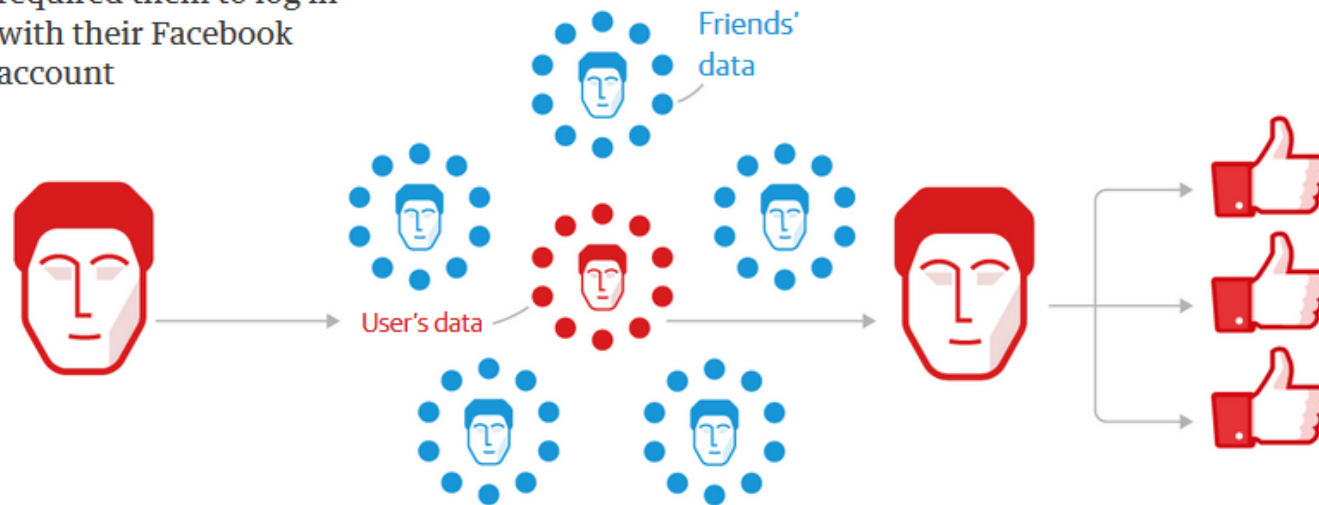
... as well their friends' data, amounting to over 50m people's raw Facebook data

These individuals could then be targeted with highly personalised advertising based on their personality data

[Graphics from https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm]

- **Scientific foundations: the method**

  Two well-known articles by Kosinski et al.

Private traits and attributes are predictable from digital records of human behavior

Michal Kosinski[a,1], David Stillwell[a], and Thore Graepel[b]

[a]Free School Lane, The Psychometrics Centre, University of Cambridge, Cambridge CB2 3RQ United Kingdom; and [b]Microsoft Research, Cambridge CB1 2FB, United Kingdom

Edited by Kenneth Wachter, University of California, Berkeley, CA, and approved February 12, 2013 (received for review October 29, 2012)

Computer-based personality judgments are more accurate than those made by humans

Wu Youyou[a,1,2], Michal Kosinski[b,1], and David Stillwell[a]

[a]Department of Psychology, University of Cambridge, Cambridge CB2 3EB, United Kingdom; and [b]Department of Computer Science, Stanford University, Stanford, CA 94305

Edited by David Funder, University of California, Riverside, CA, and accepted by the Editorial Board December 2, 2014 (received for review September 28, 2014)

- **The "Big Five" personality traits are predictable from Facebook likes**
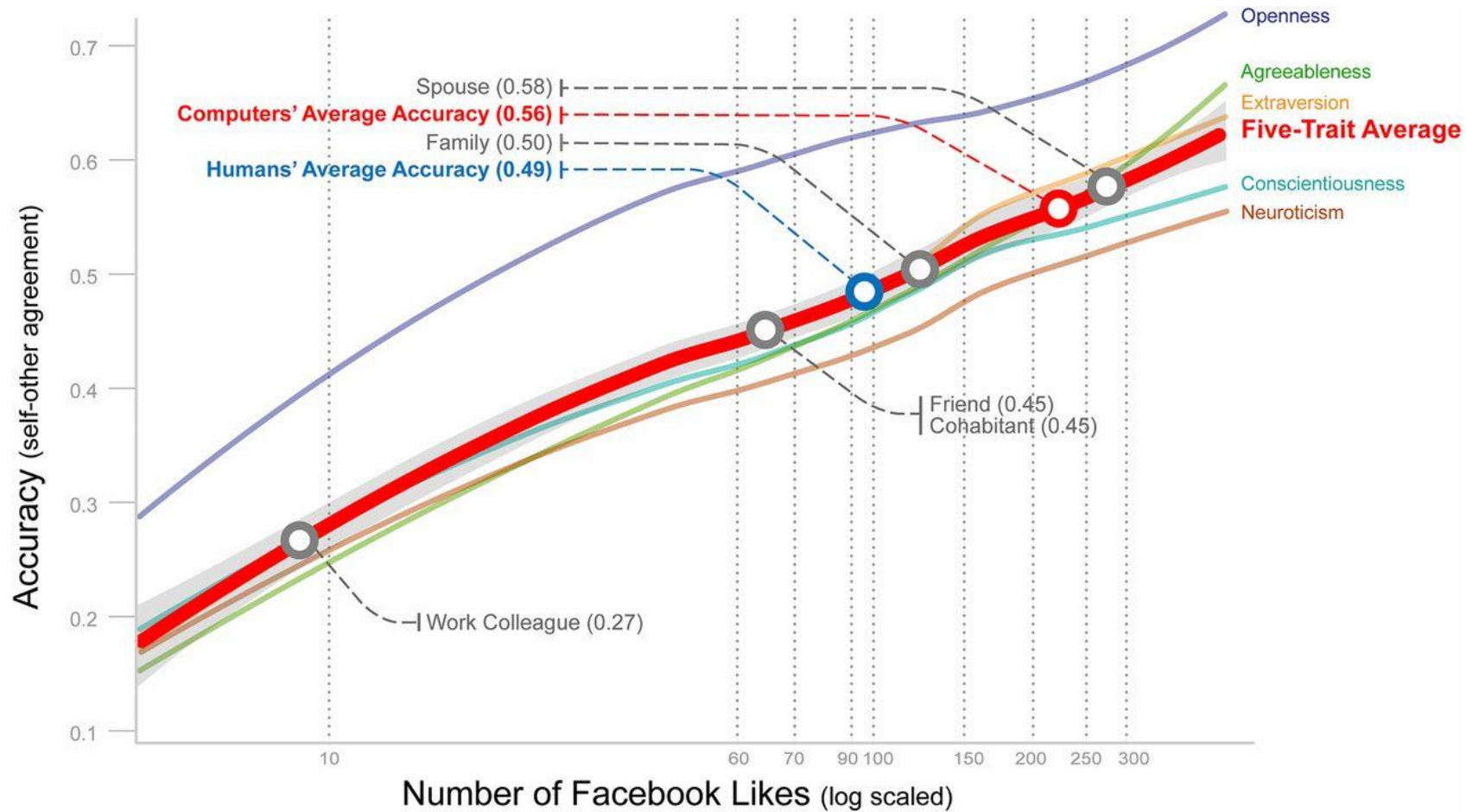
Wu Youyou et al. PNAS 2015;112:4:1036-1040

OK, about _this_ course ...

1) **Fundamentals**

2) **Deep Supervised Learning**

3) **Deep Convolutional Neural Networks**

4) **Deep Recurrent Networks**

5) **Deep Reinforcement Learning**

Labs with Google Colab

Make sure you have a look!

Better yet: follow the tutorial at https://colab.research.google.com/notebooks/intro.ipynb

PyCharm (optional)

However, if you really want to see how things work...

- ### Set it up on your computer

  Python 3.7+
  Numpy 1.17+
  TensorFlow 2.+

  ... and, above all ...

  Install PyCharm (Community Edition)

  https://www.jetbrains.com/pycharm/

  It's free

  Learn using the *debugger* :
  it will change the way you learn

  CAUTION: no assistance whatsoever will be provided for this ...
  "Legions of students succeeded in doing this on their own: you can make it too!"

1) Choose and propose a final project

The topic could be any of your choice
The techniques adopted must be (strongly) related with the course
Groupwork is allowed, with a maximum of <u>two</u> *(no exceptions)*

2) Develop and submit your projects

Each final project must be submitted as a Google Colab notebook plus dataset *(if required)*

Submission must be made at least <u>two</u> days before the exams

3) Be strong on theory as well

Alone, not even a *phenomenal* final project, will give you a final grade *(sorry)*

Appropriate knowledge about theory will be fundamental

4) Final Exam

Oral interview, about both theory and final project
Relative weights: theory 60%, final project 40%

Some References Books for the Course

- **Deep Learning**

  Ian Goodfellow, Yoshua Bengio and Aaron Courville

  MIT Press, 2017

  See also
  http://www.deeplearningbook.org/

- **Hands-On Machine Learning
  with Scikit-Learn, Keras, and TensorFlow:
  Concepts, Tools, and Techniques
  to Build Intelligent Systems, 2nd Edition**

  Aurélien Géron

  O'Reilly, 2020