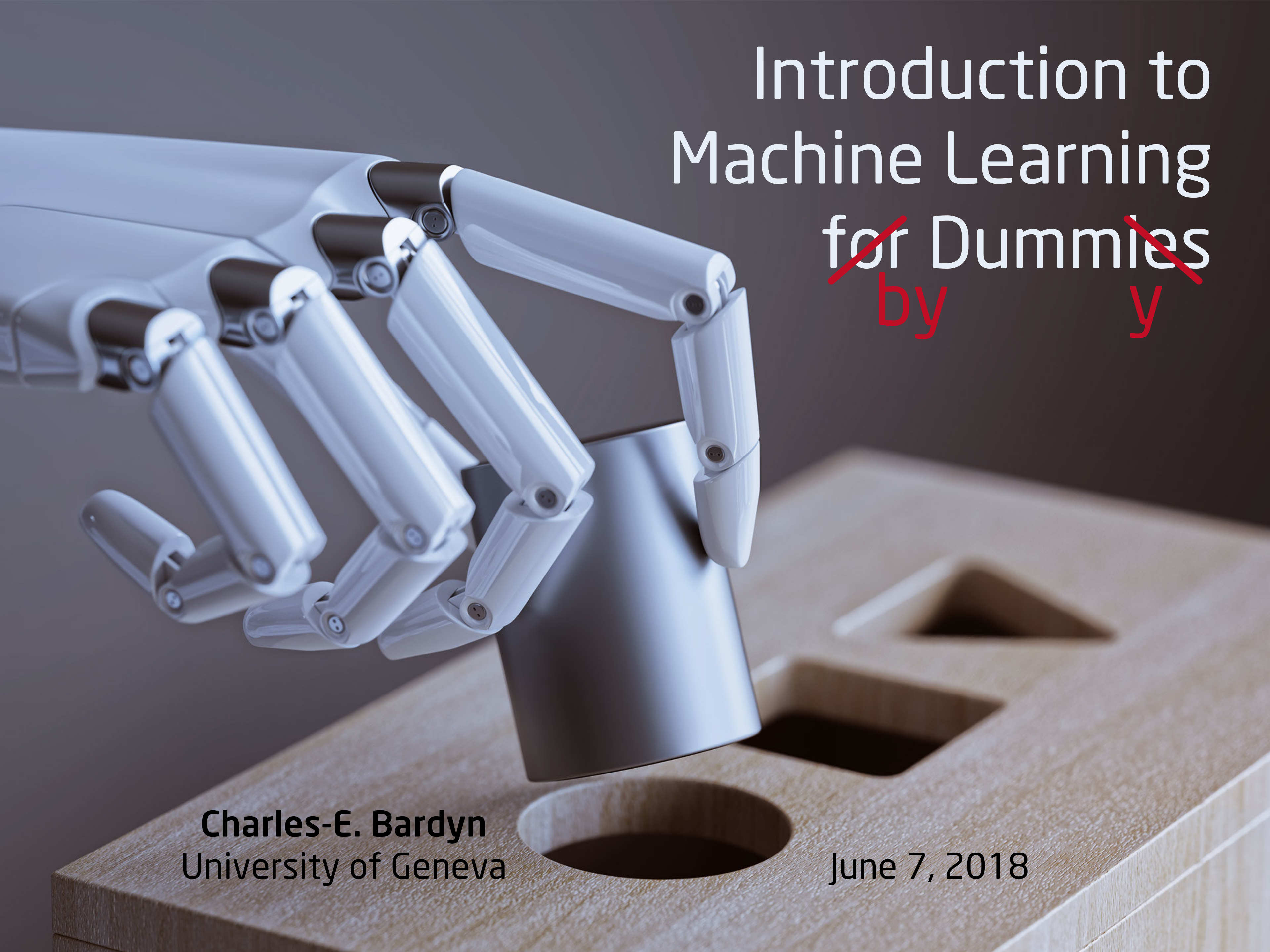
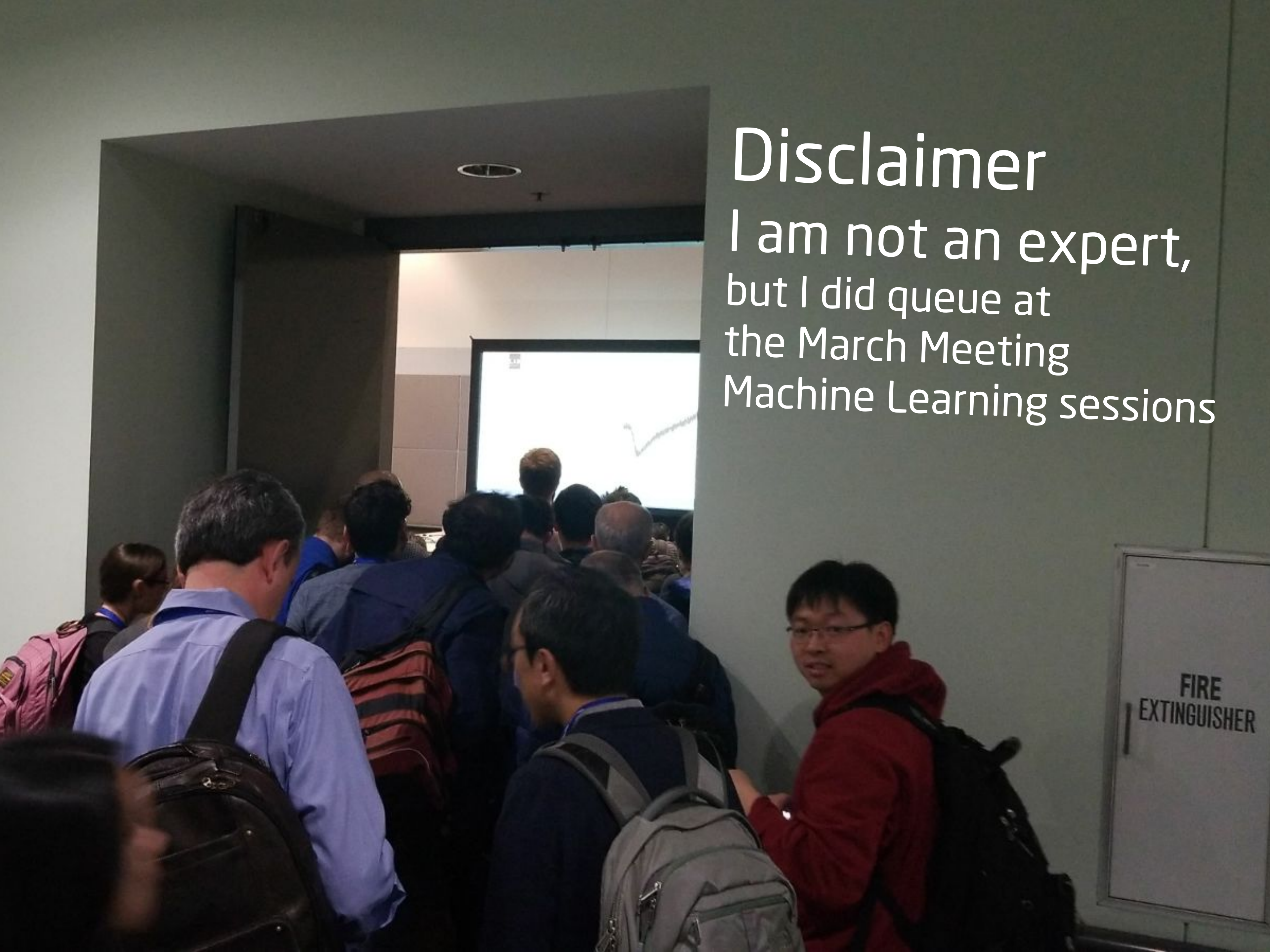


Introduction to Machine Learning ~~for~~ by Dumm~~ies~~y

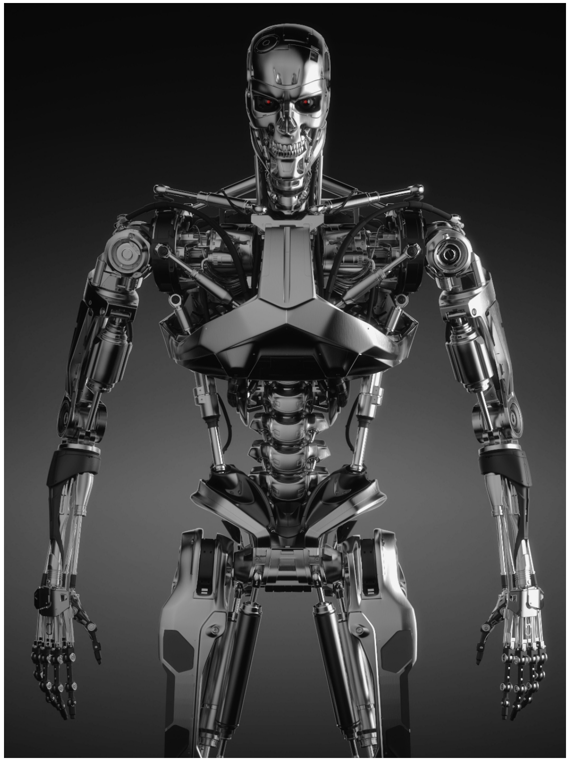
Charles-E. Bardyn
University of Geneva

June 7, 2018


Disclaimer
I am not an expert,
but I did queue at
the March Meeting
Machine Learning sessions
FIRE
EXTINGUISHER

What is machine learning?

Not killer robots

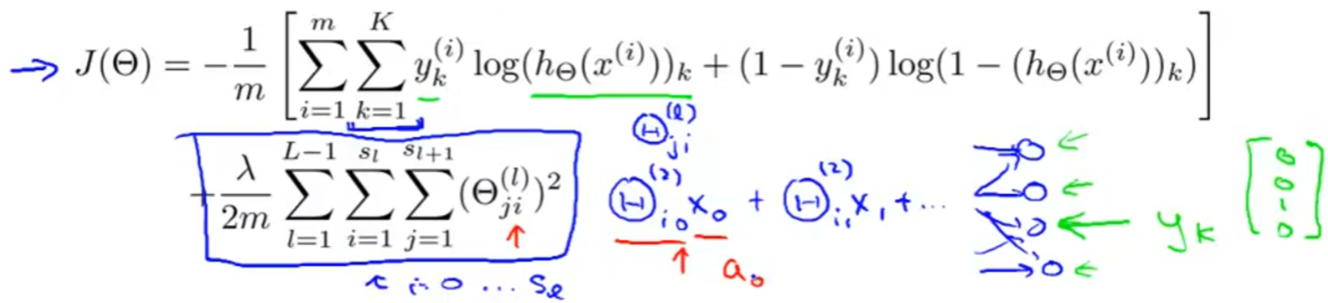


Not friendly robots

Not magic

What is machine learning?

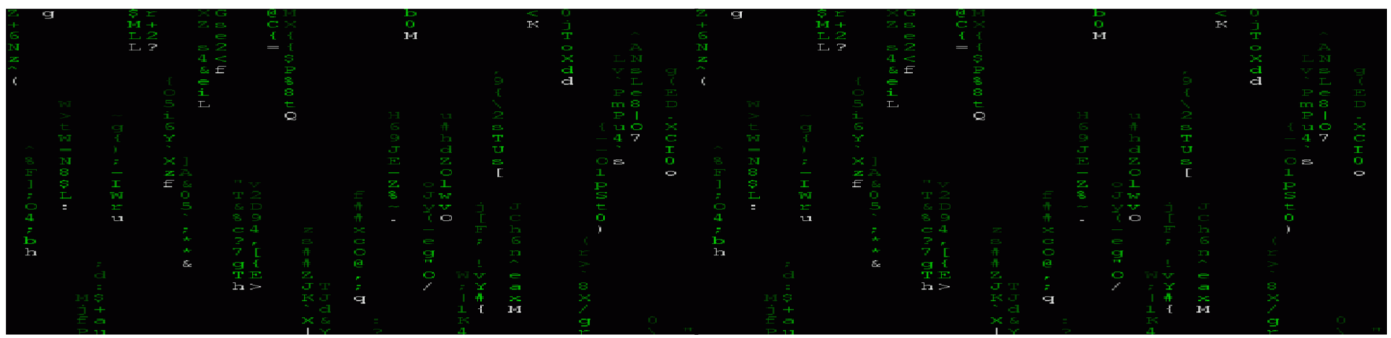
A bunch of
math



Andrew Ng, Stanford

+

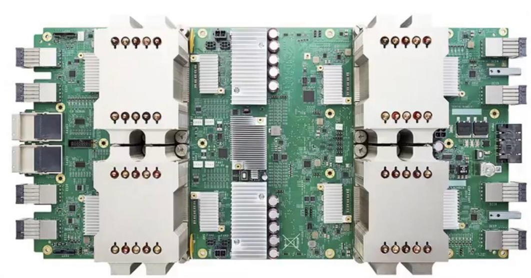
A lot of
data



+

A lot of
**computational
power**

Nvidia GPUs



Google TPUs

What is machine learning?

WIKIPEDIA

Machine learning

Machine learning is a field of computer science that often uses statistical techniques to give computers the ability to "learn" (i.e., progressively improve performance on a specific task) with data, without being explicitly programmed.[1]

Setup

Dataset X

Model $p(X|w)$

Probability of observing X

Tunable model parameters

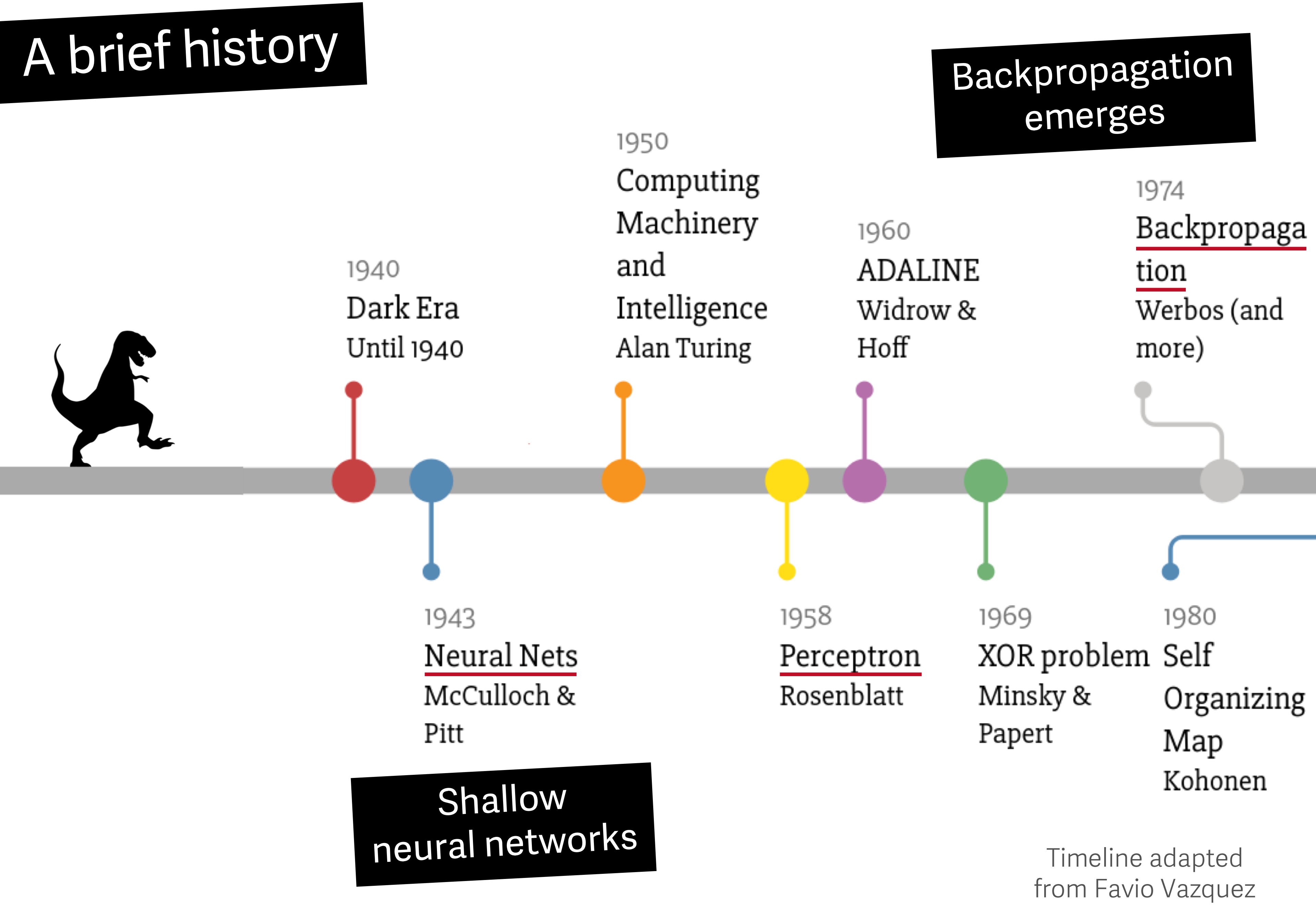
Goal

Optimize $p(X|w)$ not to fit the data, but to **make predictions**

Problem origins

Comes from **statistics, computational neuroscience**, and **physics**

Many connections to statistical physics
(Monte Carlo, simulated annealing, variational methods, etc.)

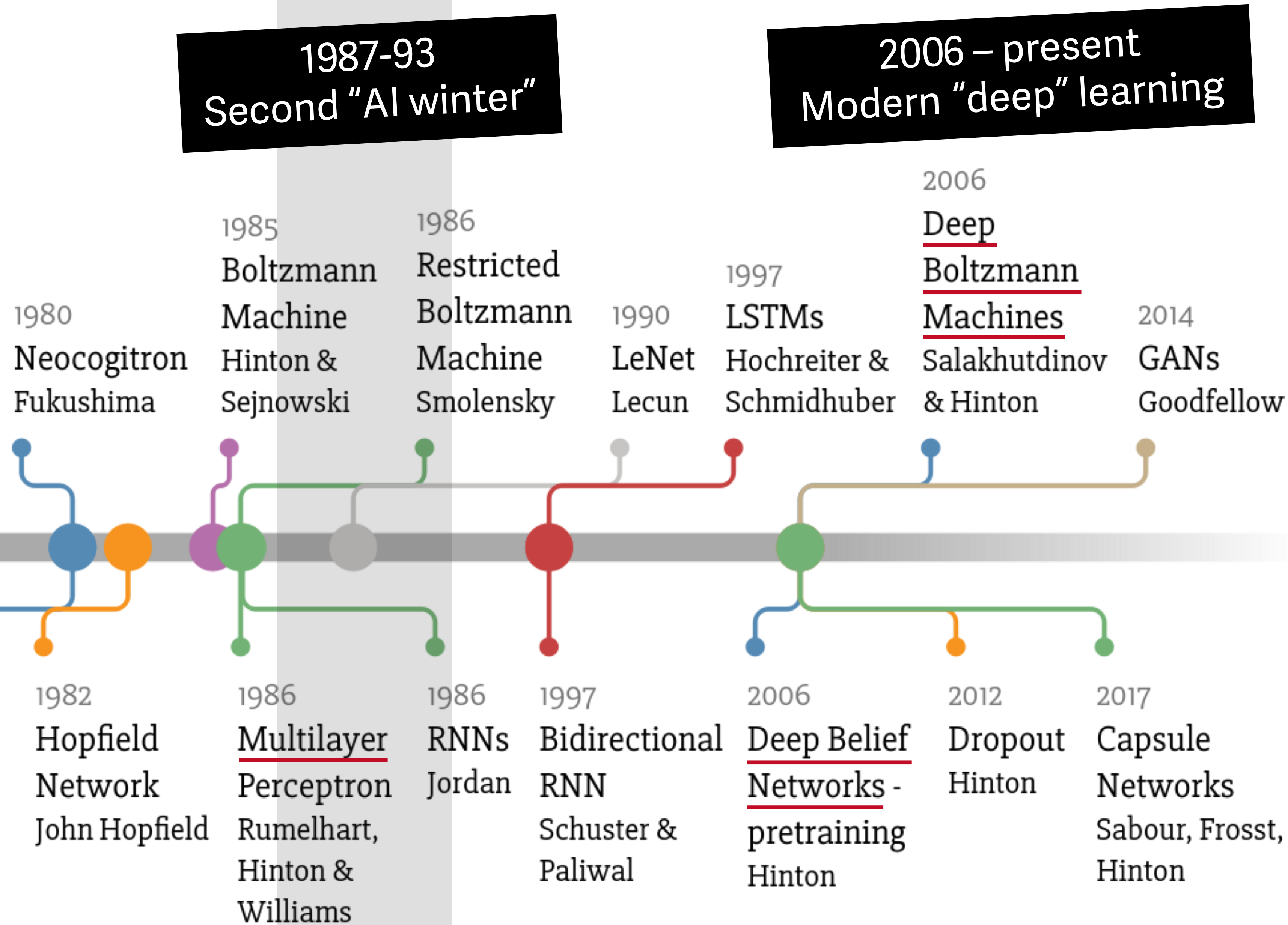

A brief history
Backpropagation emerges
1940
Dark Era
Until 1940
1943
Neural Nets
McCulloch & Pitt
1950
Computing Machinery and Intelligence
Alan Turing
1958
Perceptron
Rosenblatt
1960
ADALINE
Widrow & Hoff
1969
XOR problem
Minsky & Papert
1974
Backpropagation
Werbos (and more)
1980
Self Organizing Map
Kohonen
Shallow neural networks
Timeline adapted from Favio Vazquez

1974-80
First "AI winter"

Wikipedia

AI winter

In the history of artificial intelligence, an **AI winter** is a period of reduced funding and interest in artificial intelligence research.[1]


1987-93
Second "AI winter"
2006 – present
Modern "deep" learning
1980
Neocogitron
Fukushima
1985
Boltzmann Machine
Hinton & Sejnowski
1986
Restricted Boltzmann Machine
Smolensky
1990
LeNet
Lecun
1997
LSTMs
Hochreiter & Schmidhuber
2006
Deep Boltzmann Machines
Salakhutdinov & Hinton
2014
GANs
Goodfellow
1982
Hopfield Network
John Hopfield
1986
Multilayer Perceptron
Rumelhart, Hinton & Williams
1986
RNNs
Jordan
1997
Bidirectional RNN
Schuster & Paliwal
2006
Deep Belief Networks - pretraining
Hinton
2012
Dropout
Hinton
2017
Capsule Networks
Sabour, Frosst, Hinton

1987-93
Second "AI winter"

2006 – present
Modern "deep" learning

1980
Neocogitron
Fukushima

1985
Boltzmann
Machine
Hinton &
Sejnowski

Restricted
Boltzmann
Machine
Smolensky

1990
LeNet
Lecun

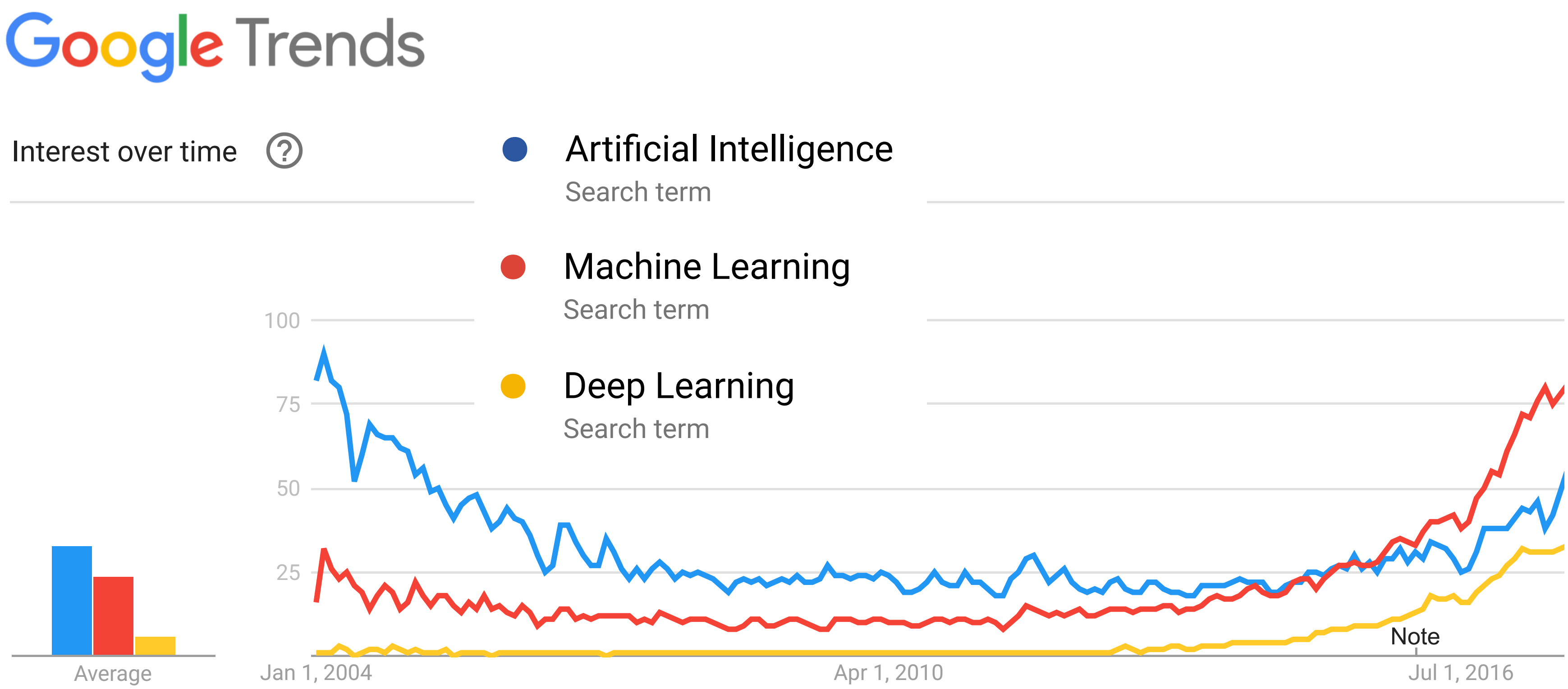
LSTMs
Hochreiter &
Schmidhuber

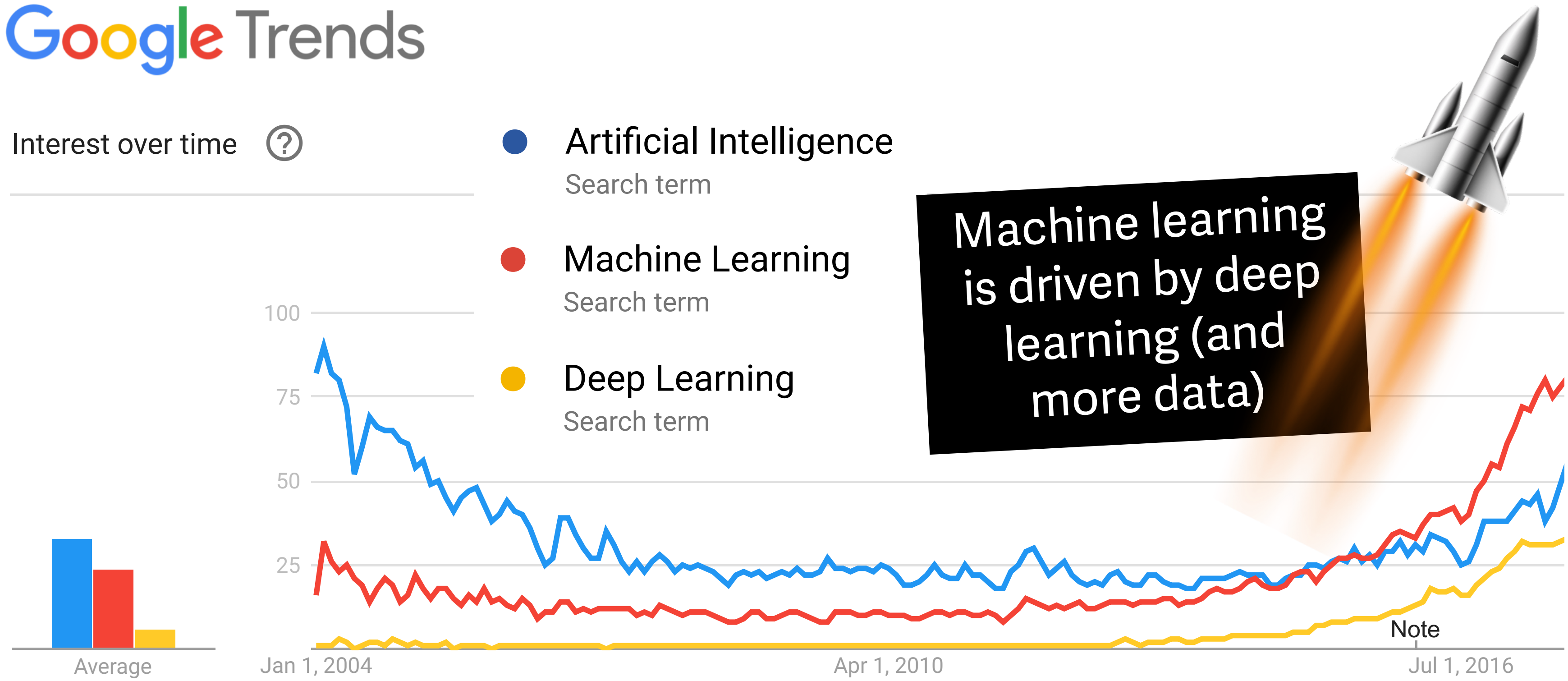
2006
Deep
Boltzmann
Machines
Salakhutdin
& Hinton

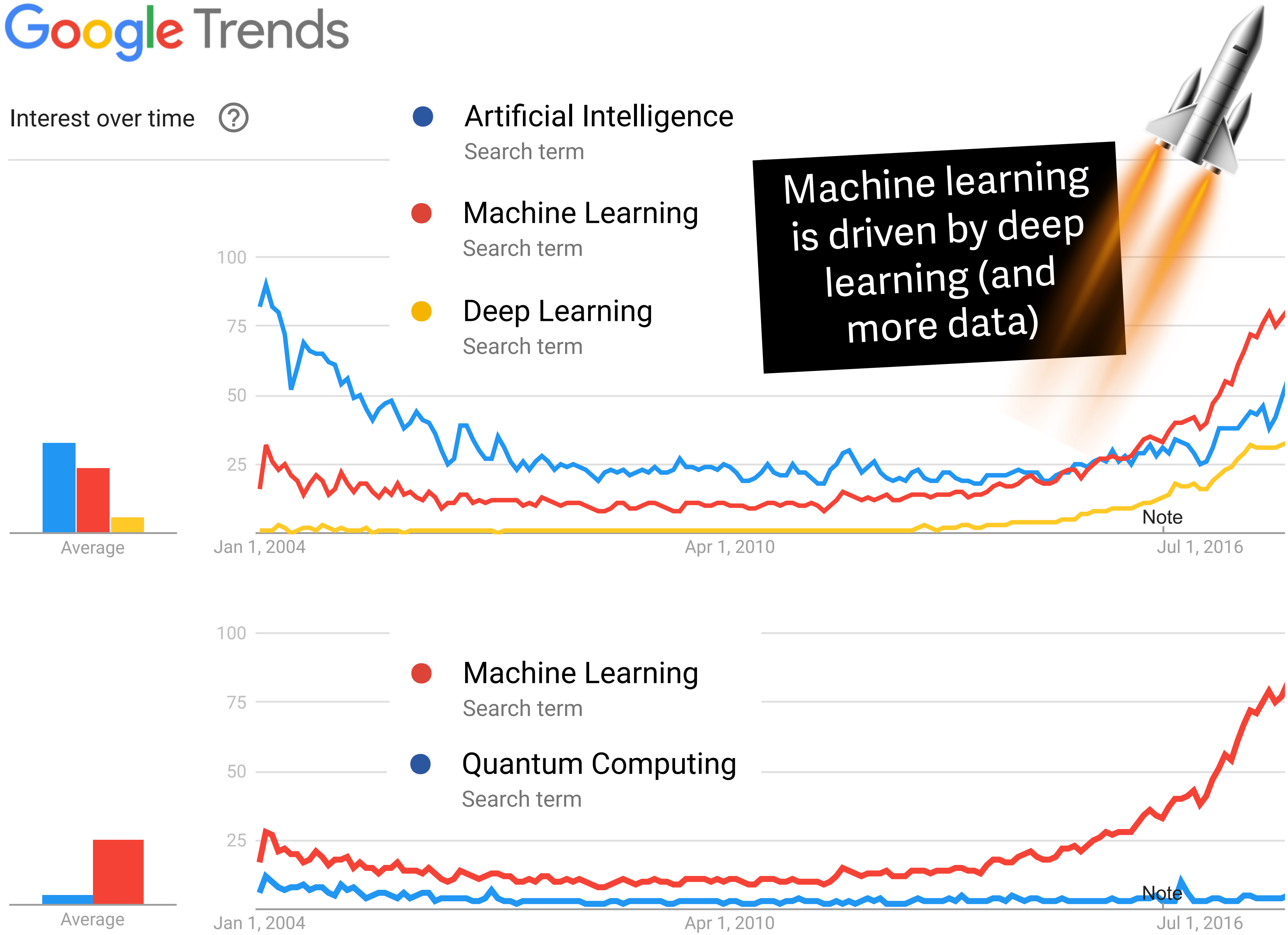
GANs
Goodfellow

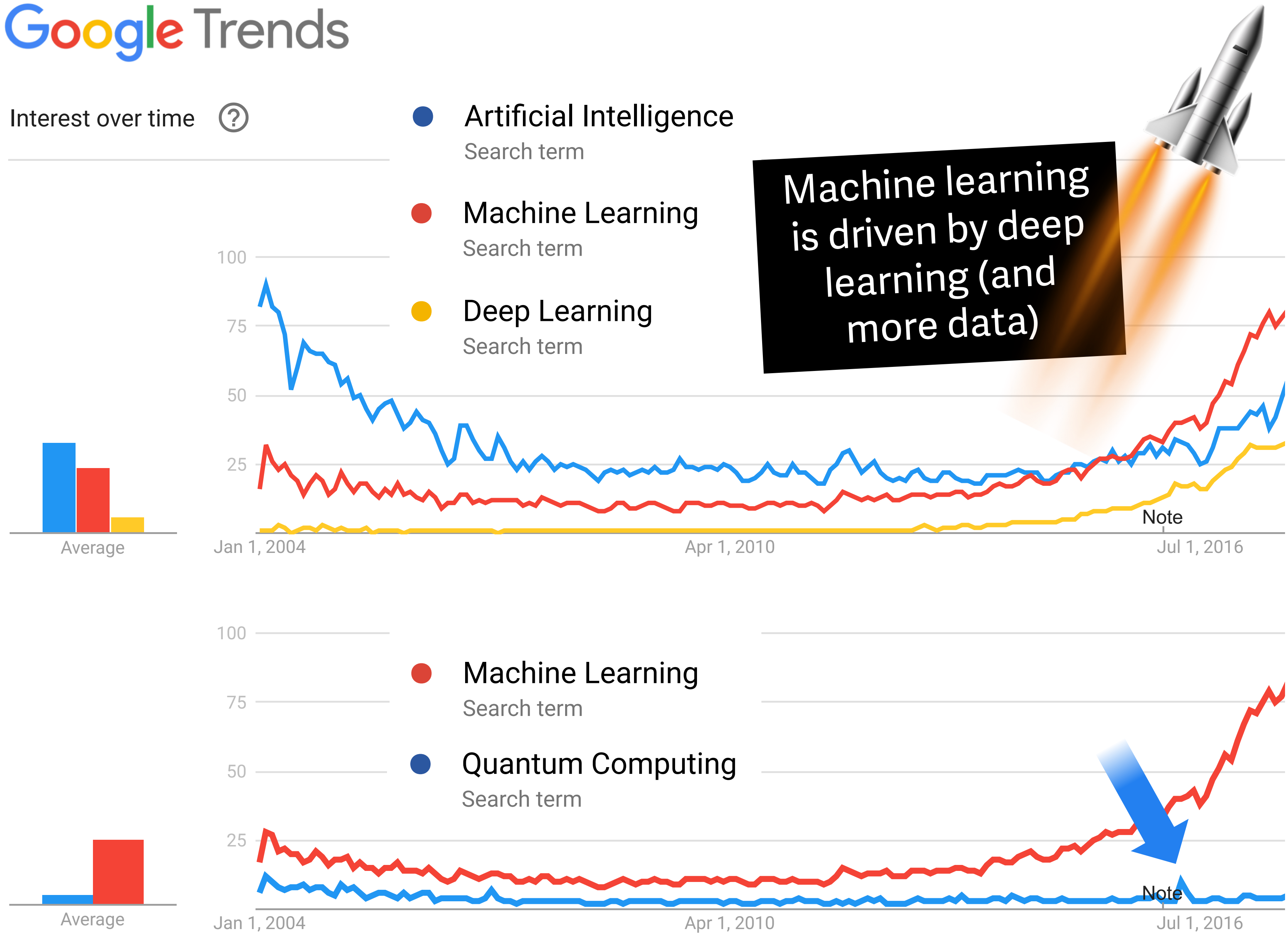
"I think people need to understand that deep learning is making a lot of things, behind-the-scenes, much better."

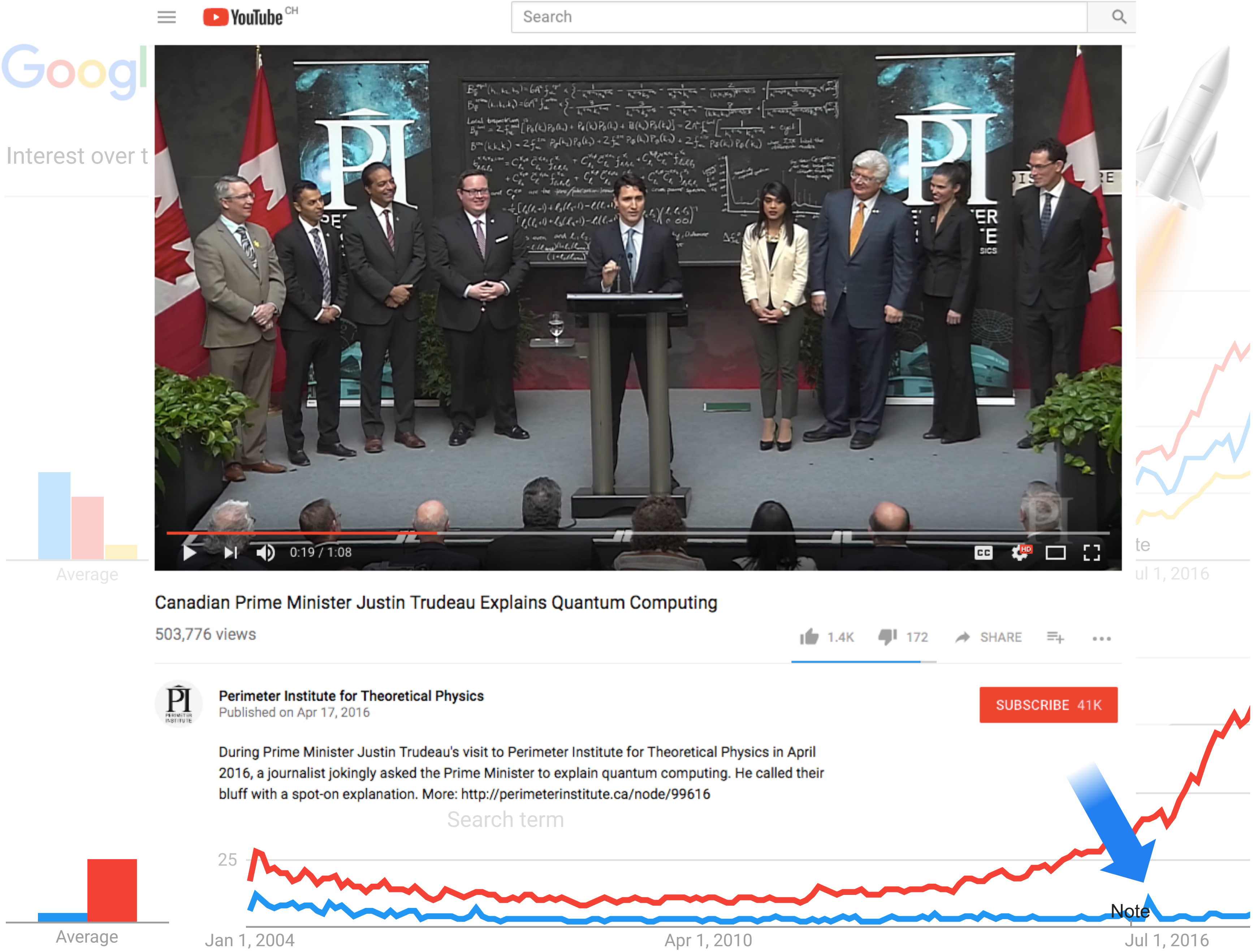
Geoffrey Hinton


Google Trends
Interest over time
Artificial Intelligence
Search term
Machine Learning
Search term
Deep Learning
Search term
100
75
50
25
Average
Jan 1, 2004
Apr 1, 2010
Jul 1, 2016
Note


Google Trends
Interest over time
Artificial Intelligence
Search term
Machine Learning
Search term
Deep Learning
Search term
Machine learning is driven by deep learning (and more data)
100
75
50
25
Average
Jan 1, 2004
Apr 1, 2010
Jul 1, 2016
Note


Google Trends
Interest over time
Artificial Intelligence
Search term
Machine Learning
Search term
Deep Learning
Search term
Machine learning is driven by deep learning (and more data)
100
75
50
25
Average
Jan 1, 2004
Apr 1, 2010
Jul 1, 2016
Note
100
75
50
25
Machine Learning
Search term
Quantum Computing
Search term
Average
Jan 1, 2004
Apr 1, 2010
Jul 1, 2016
Note


Google Trends
Interest over time
Artificial Intelligence
Search term
Machine Learning
Search term
Deep Learning
Search term
Machine learning is driven by deep learning (and more data)
100
75
50
25
Average
Jan 1, 2004
Apr 1, 2010
Jul 1, 2016
Note
100
75
50
25
Machine Learning
Search term
Quantum Computing
Search term
Average
Jan 1, 2004
Apr 1, 2010
Jul 1, 2016
Note


YouTube CH
Search
Interest over t
Average
Canadian Prime Minister Justin Trudeau Explains Quantum Computing
503,776 views
1.4K
172
SHARE
Perimeter Institute for Theoretical Physics
Published on Apr 17, 2016
SUBSCRIBE 41K
During Prime Minister Justin Trudeau's visit to Perimeter Institute for Theoretical Physics in April 2016, a journalist jokingly asked the Prime Minister to explain quantum computing. He called their bluff with a spot-on explanation. More: http://perimeterinstitute.ca/node/99616
0:19 / 1:08
Search term
25
Note
Average
Jan 1, 2004
Apr 1, 2010
Jul 1, 2016

Entering the zoo of ML algorithms

Different types of learning tasks

Data / input from:

Data / input from	Learning task
Human	**Supervised** learning
Human, Machine	**Augmented supervised** learning
Human, Machine	**Semi-supervised** learning
Human, Machine	**Reinforcement** learning
Machine	**Unsupervised** learning

Inspired by Lex Fridman, MIT

Entering the zoo of ML algorithms

Different types of learning tasks

Data / input from:		
Human	**Supervised** learning	**Current** successes
Human, Machine	**Augmented supervised** learning	
Human, Machine	**Semi-supervised** learning	**Current / near-term** future successes
Human, Machine	**Reinforcement** learning	
Machine	**Unsupervised** learning	**Longer-term future** successes

Inspired by Lex Fridman, MIT

The zoo of ML algorithms

Classification

Logistic regression,
support vector machines (SVMs),
k-nearest neighbors,
decision trees,
random forests,
etc.

Supervised learning

Regression

Linear regression,
polynomial regression,
lasso regression,
ridge regression,
etc.

Clustering

k-means,
hierarchical clustering,
etc.

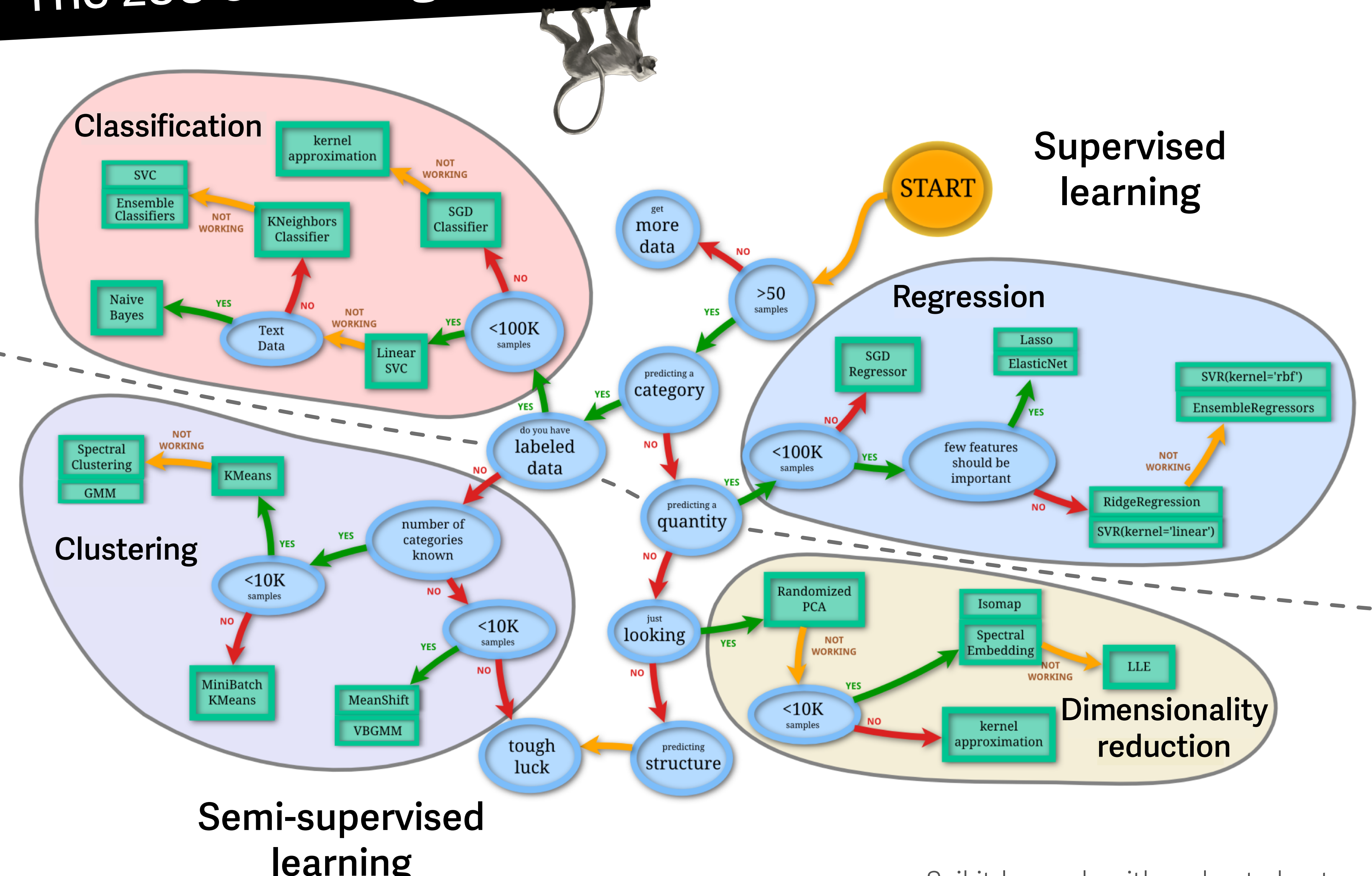
Principal components analysis (PCA),
singular value decomposition (SVD),
etc.

Dimensionality reduction

Semi-supervised learning

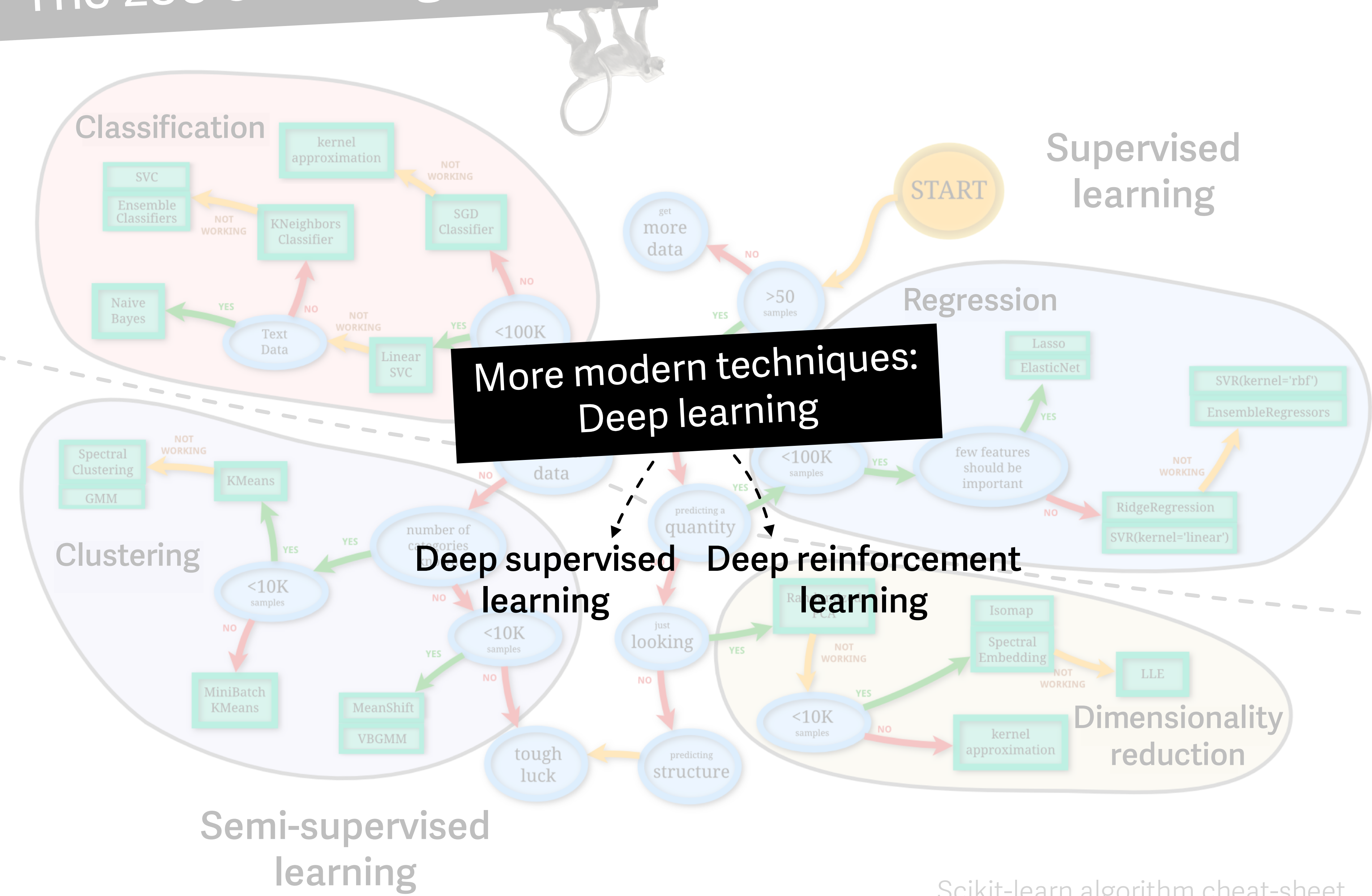
Scikit-learn algorithm cheat-sheet

The zoo of ML algorithms



Scikit-learn algorithm cheat-sheet

The zoo of ML algorithms



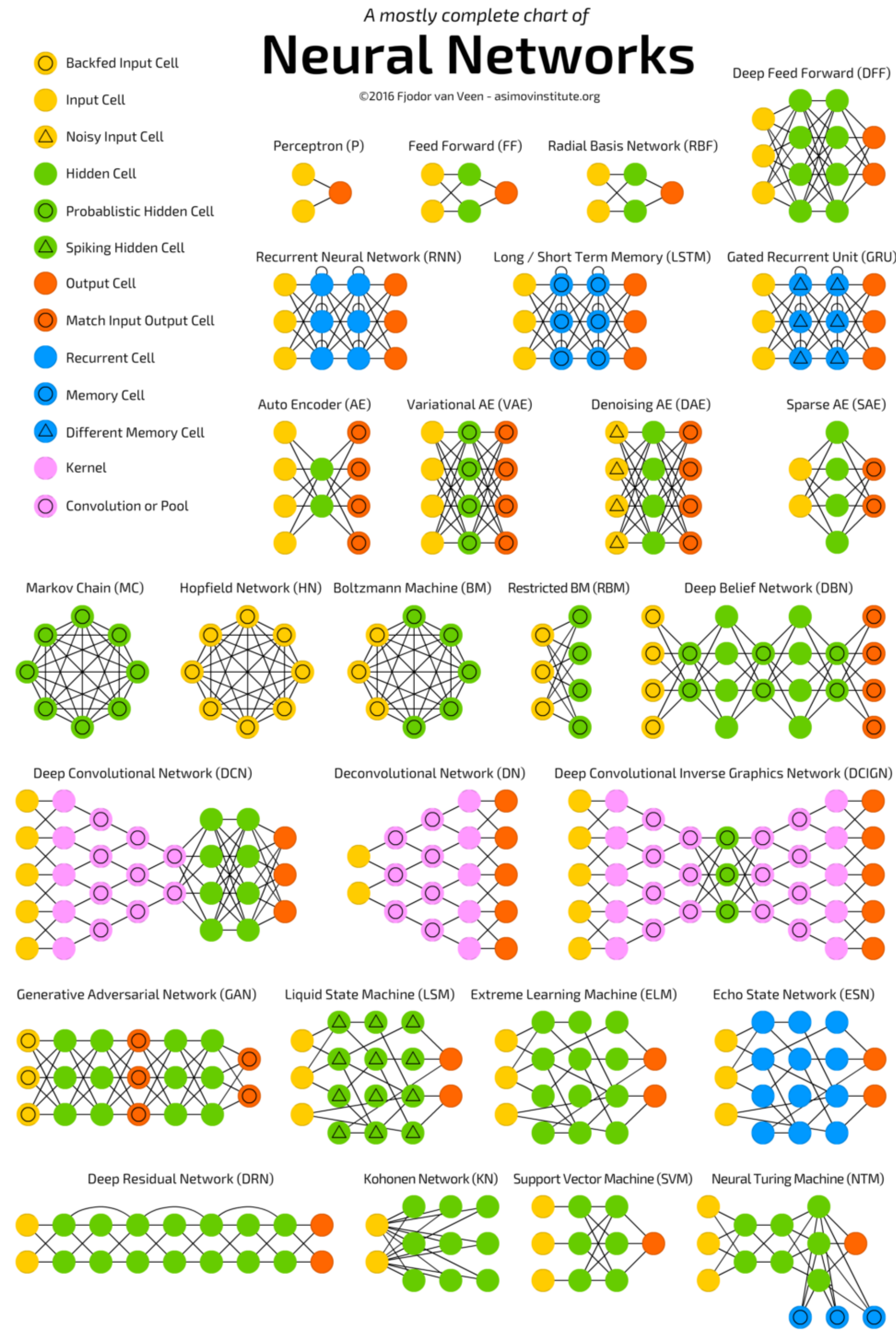
Scikit-learn algorithm cheat-sheet

The deep learning jungle

Deep learning
≈
Multilayer (deep) neural networks

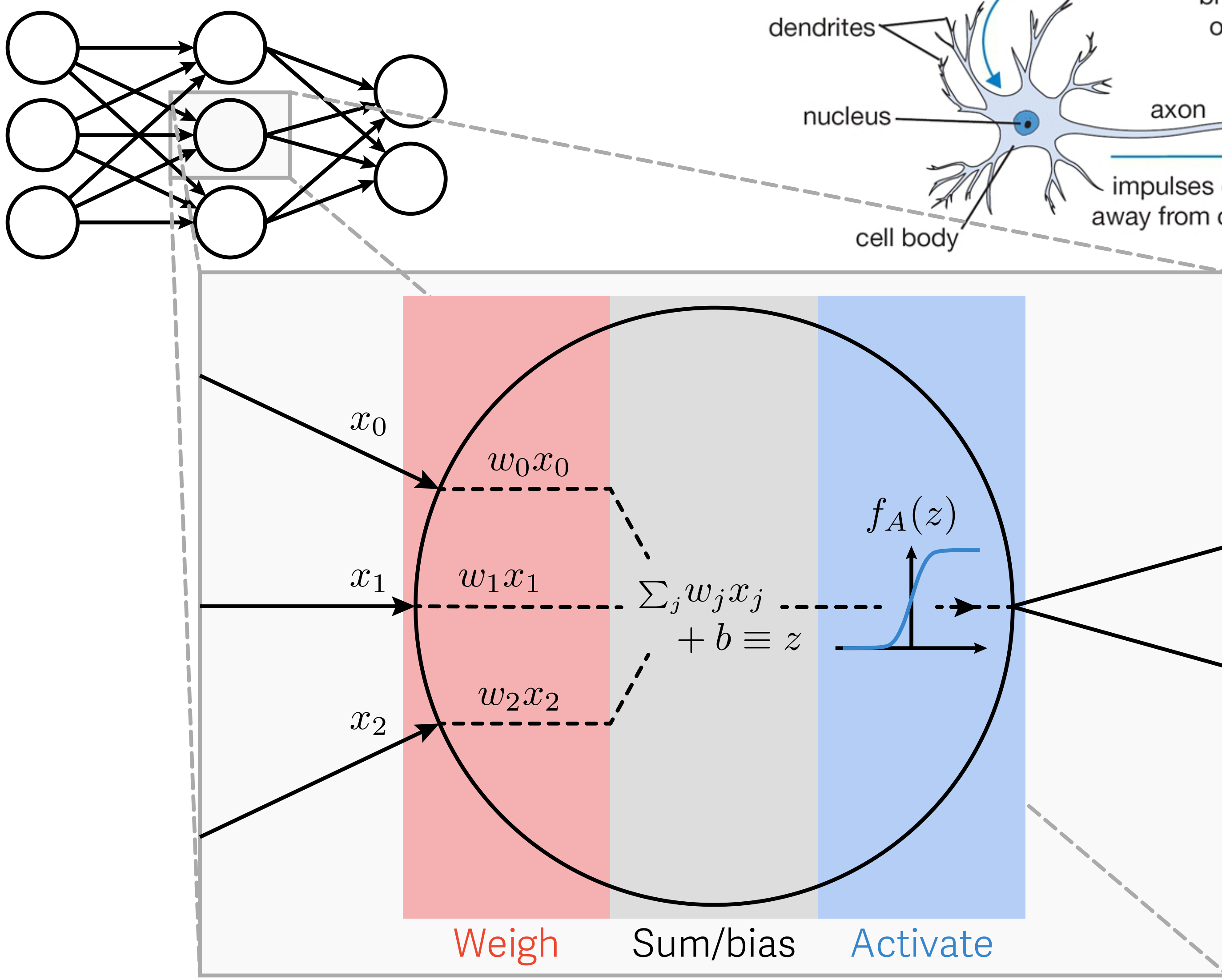
Many types of neural networks

Convolutional neural networks (CNN),
Recurrent neural networks (RNN),
Long short-term memory (LSTM),
Restricted Boltzmann machines (RBM),
etc.

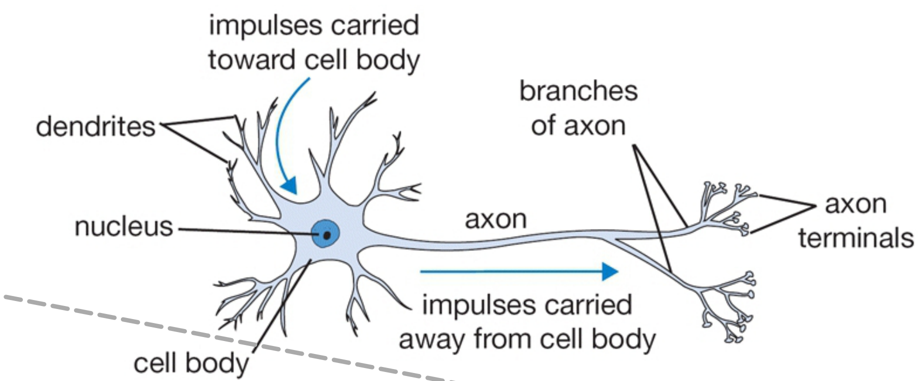


Deep learning

Building blocks: artificial neurons



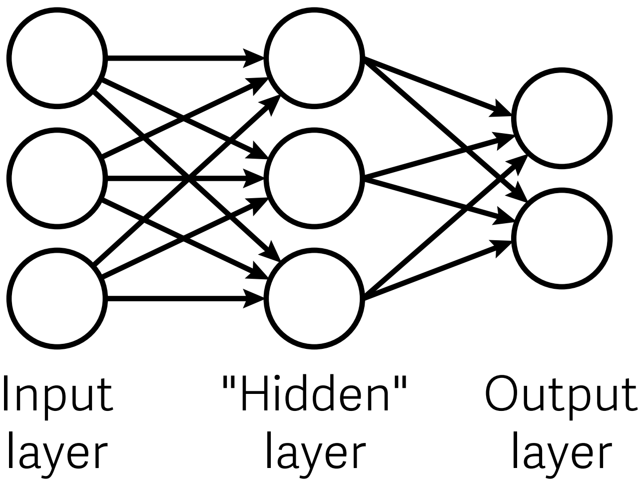
Vaguely inspired by the brain



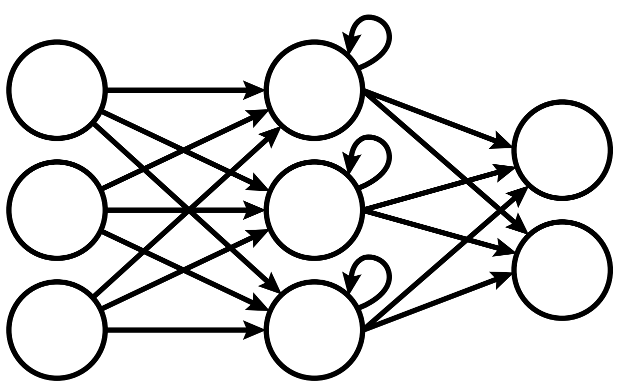
Deep learning

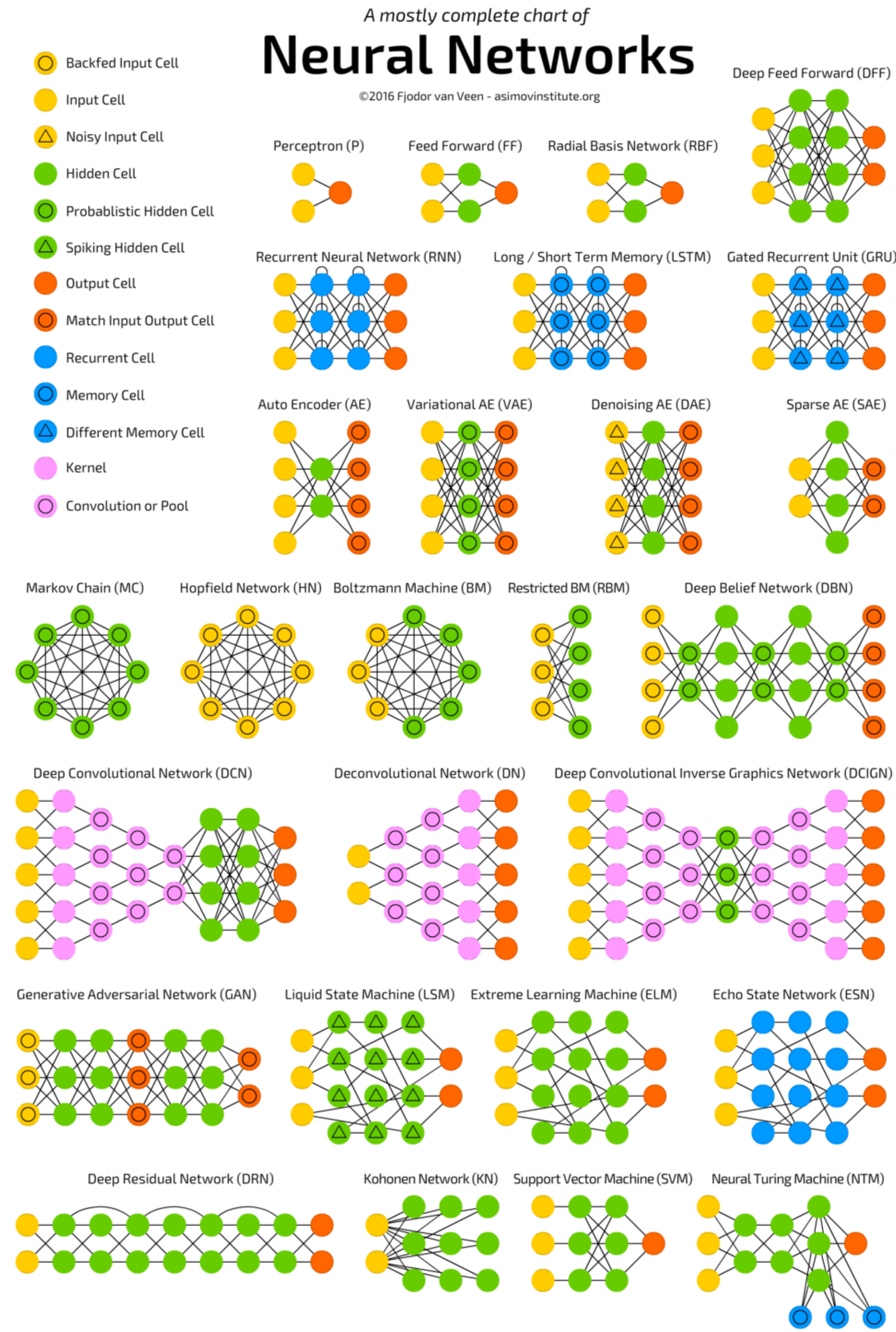
Combine neurons into layers

Feed-forward neural network



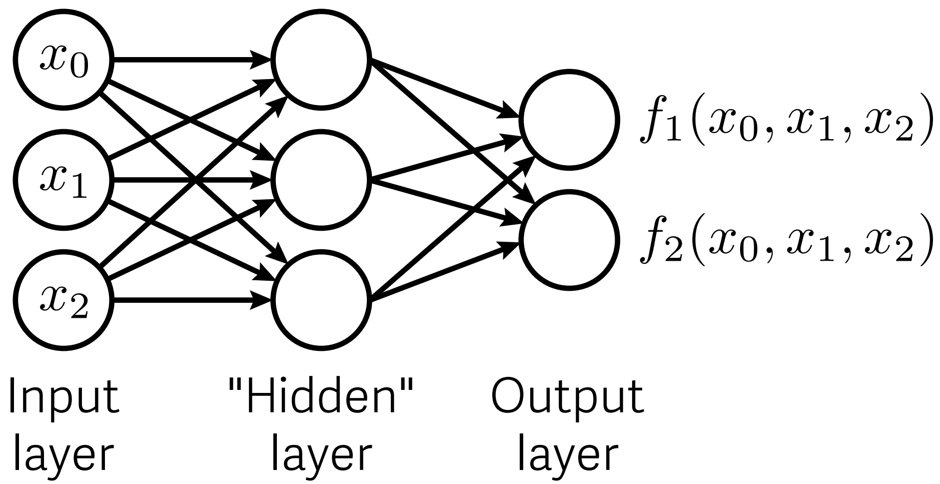
Recurrent neural network





Deep learning

Universal approximation theorem

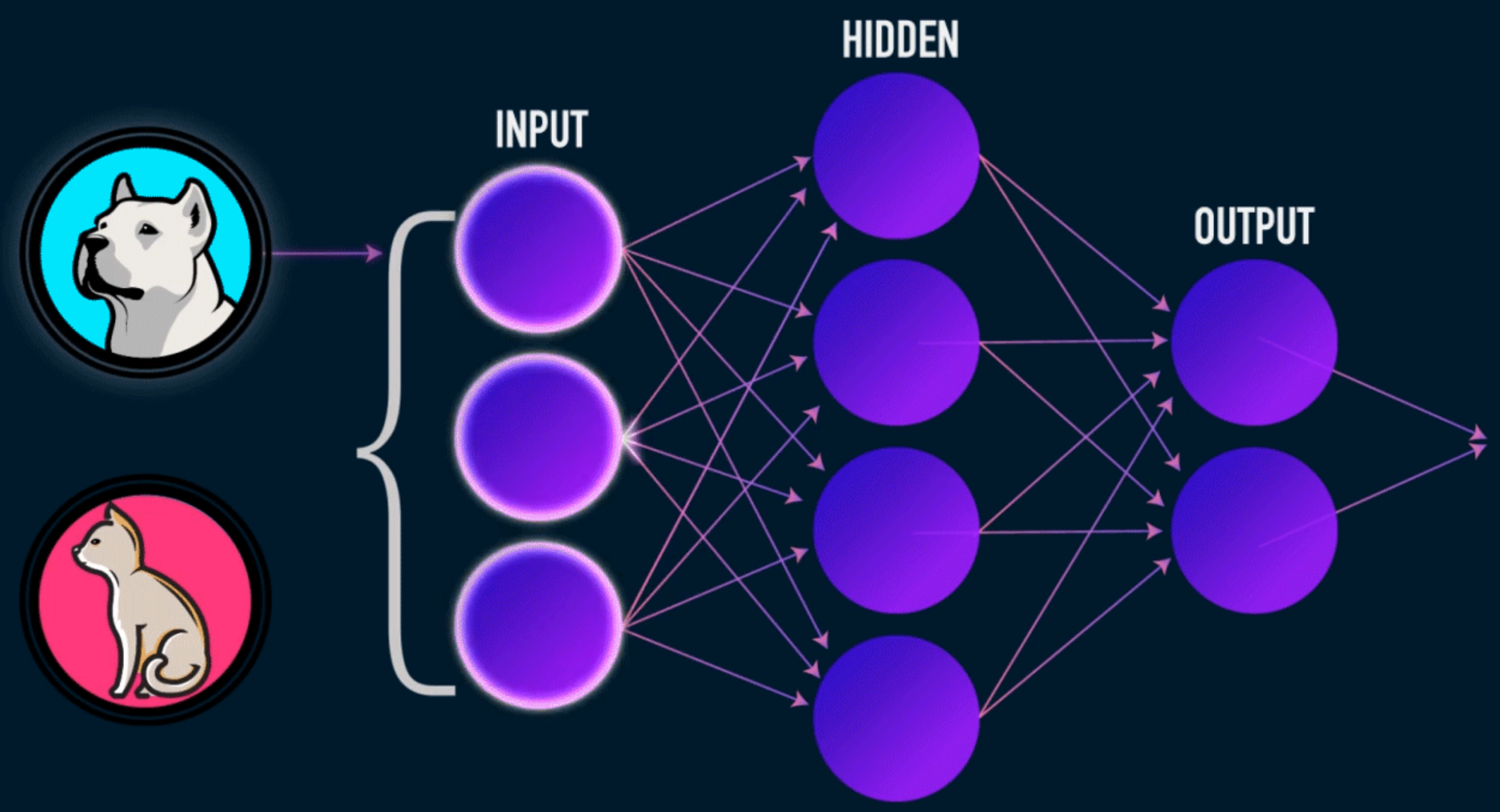


Neural networks can approximate any function (given enough neurons)

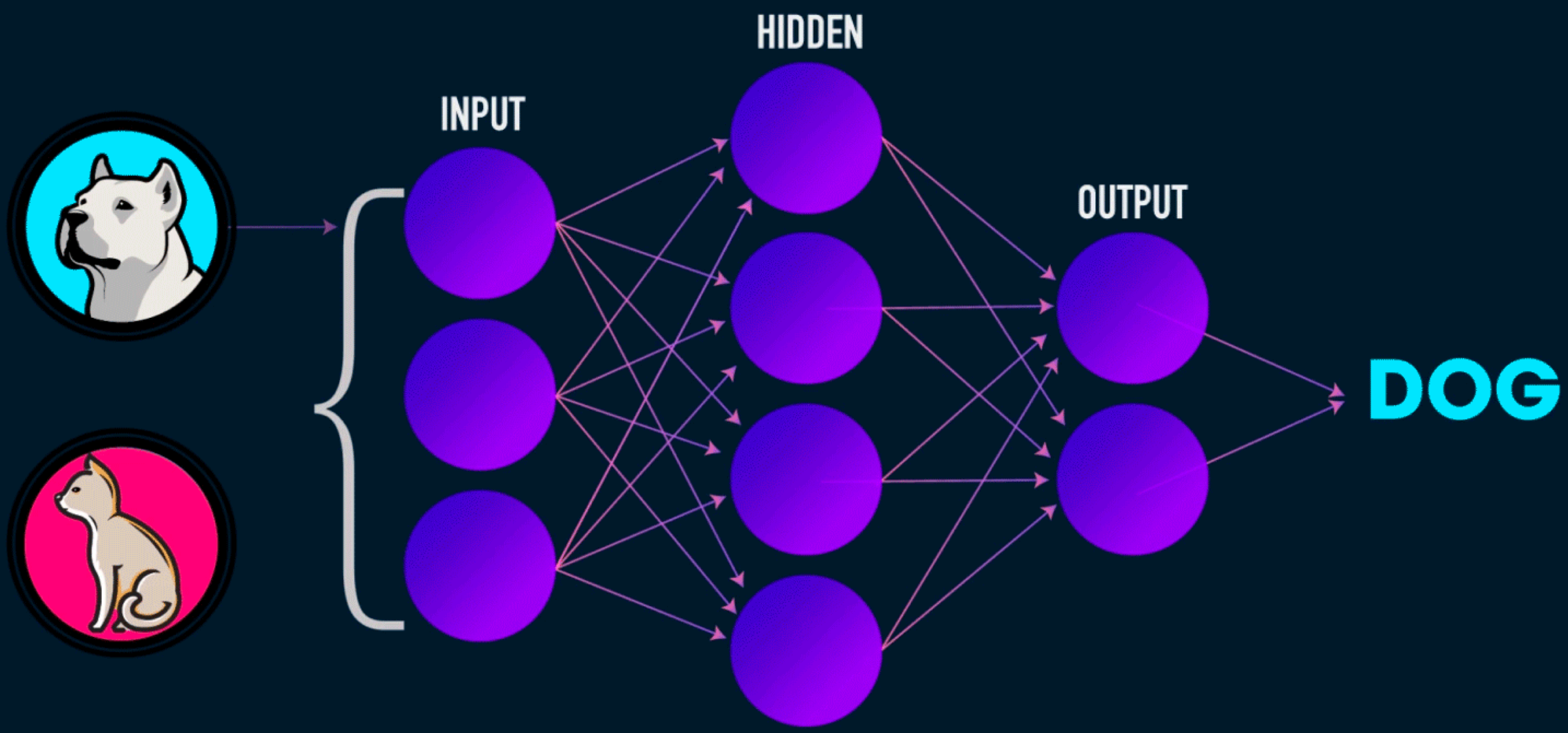
Cybenko (1989)

For any function f(x), there exists a neural network that closely approximate f(x) for any input x

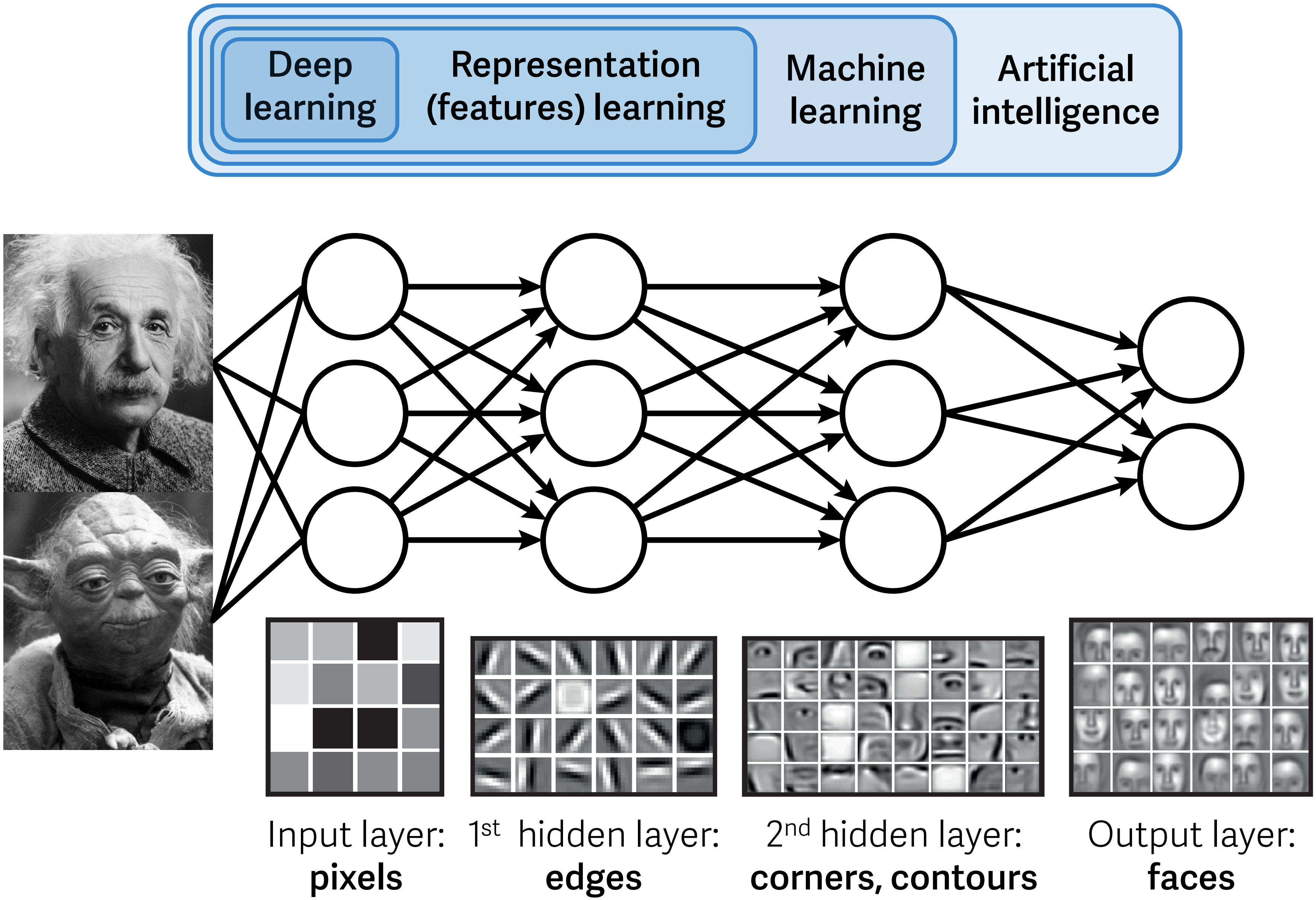
One hidden layer is enough!



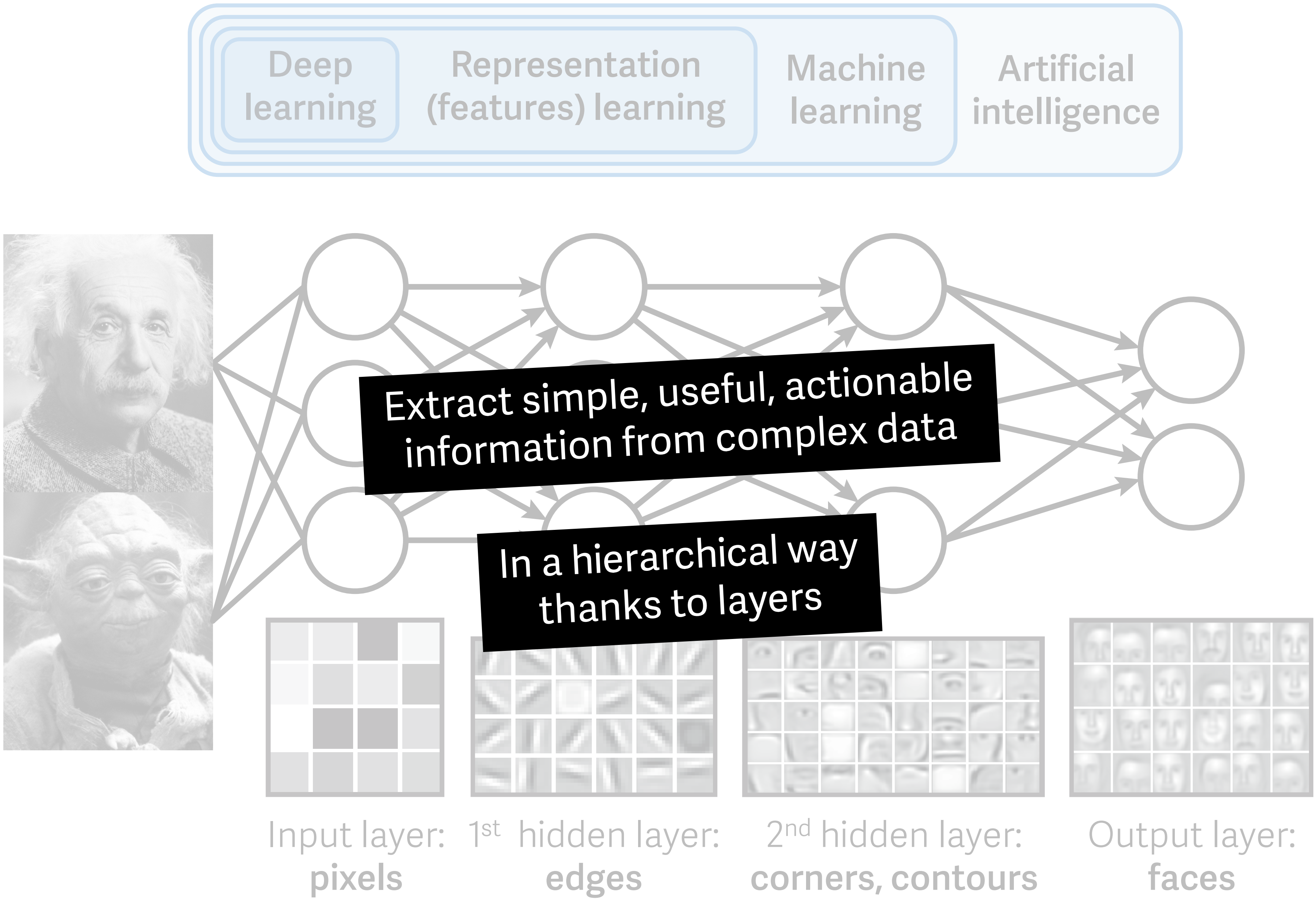
Source: James Loy, Georgia Tech



Source: James Loy, Georgia Tech

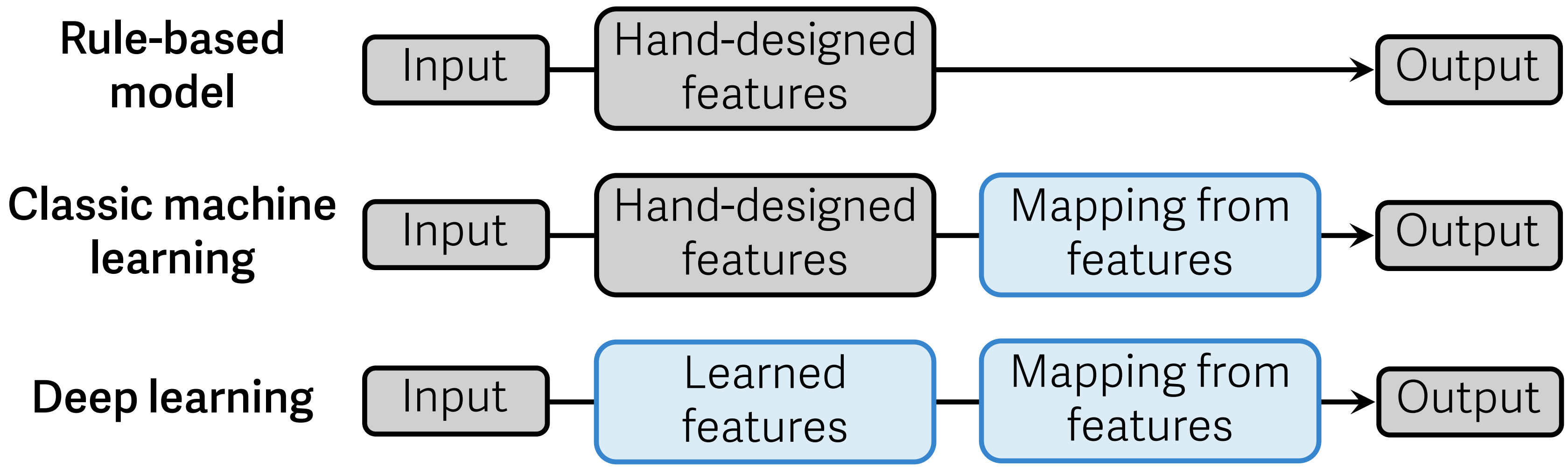


Hierarchy pictures from Jones, Nature 505, 148 (2014)

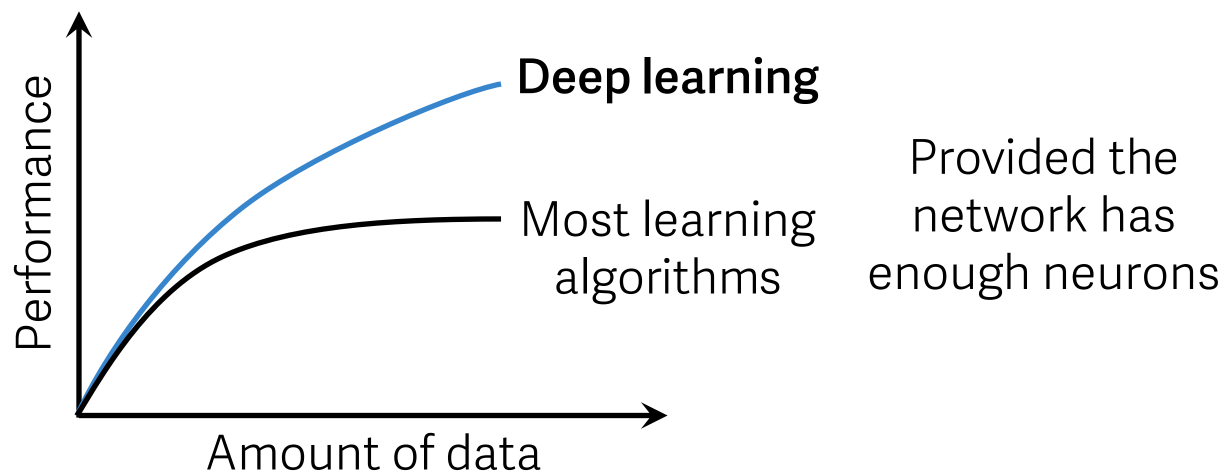


Hierarchy pictures from Jones, Nature 505, 148 (2014)

Deep learning

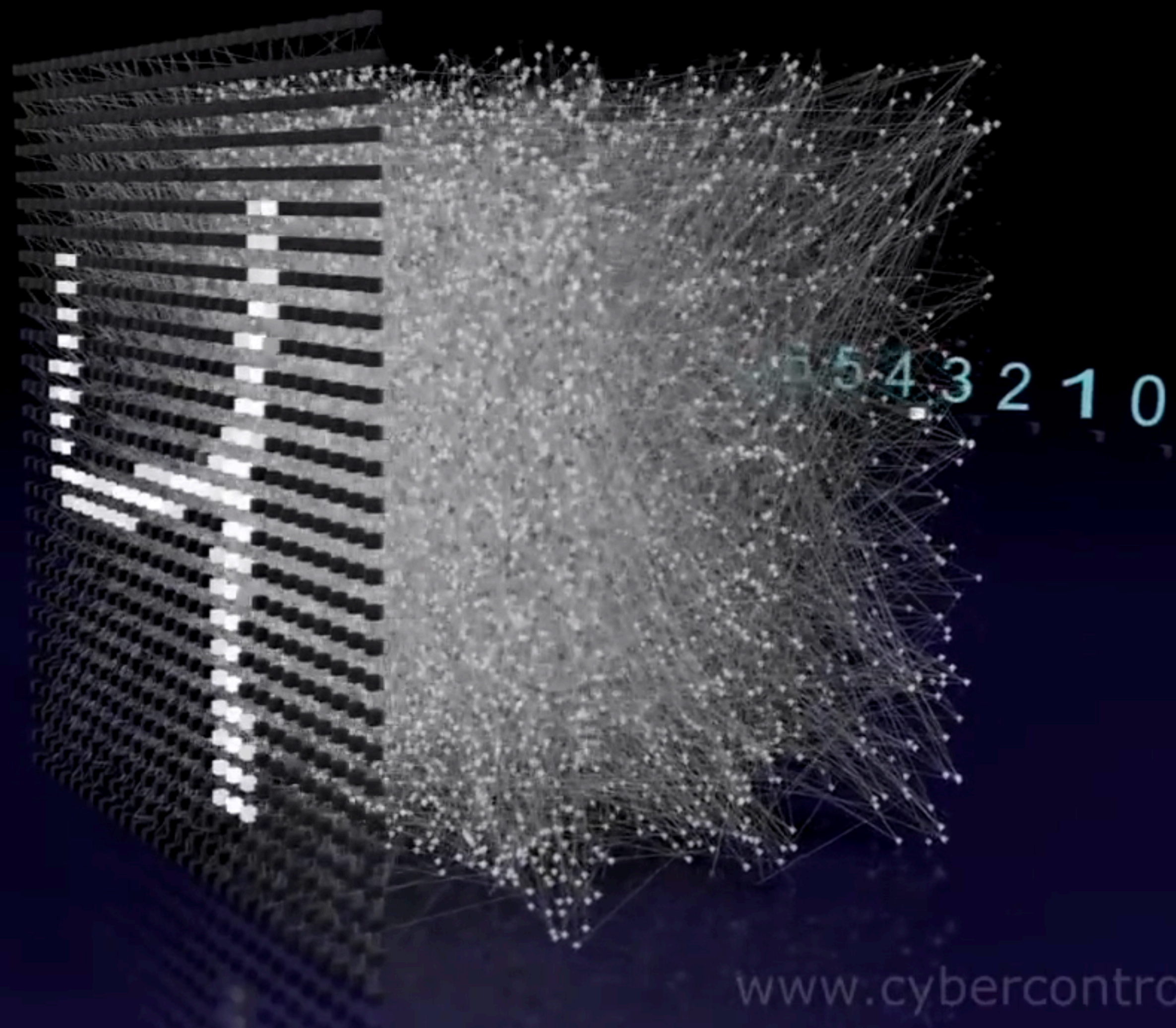


Scalable machine learning



Inspired by Lex Fridman, MIT

Type: ML Perceptron
Data Set: MNIST
Hidden Layers: 3
Hidden Neurons: 10000
Synapses: 24864180
Synapses shown: 2%
Learning: BP




"You take the red pill,
you stay in
Wonderland"
Applications
Technical details

Machine learning

The devil is in the

Generic setup

- **Dataset** X
- **Model** $f_M(w)$
- **Cost function** $f_C(X; f_M(w))$

Model parameters

Need something **to quantify the model performance!**

Goal

Find a model that best **predicts new (unseen) data**

Generic learning procedure

1. **Divide the dataset into training and test sets,** X_{train} and X_{test}
2. Train the model, i.e., **minimize the cost function on X_{train} alone**
3. **Evaluate the** generalization (prediction) **performance on X_{test}**

"Cross-validation"

Machine learning is hard

Example: polynomial regression

Dataset $X = (x_i, y_i)$ sampled from:

$$y_i = f(x_i) + \epsilon_i$$

Unknown function

Noise (e.g., Gaussian)

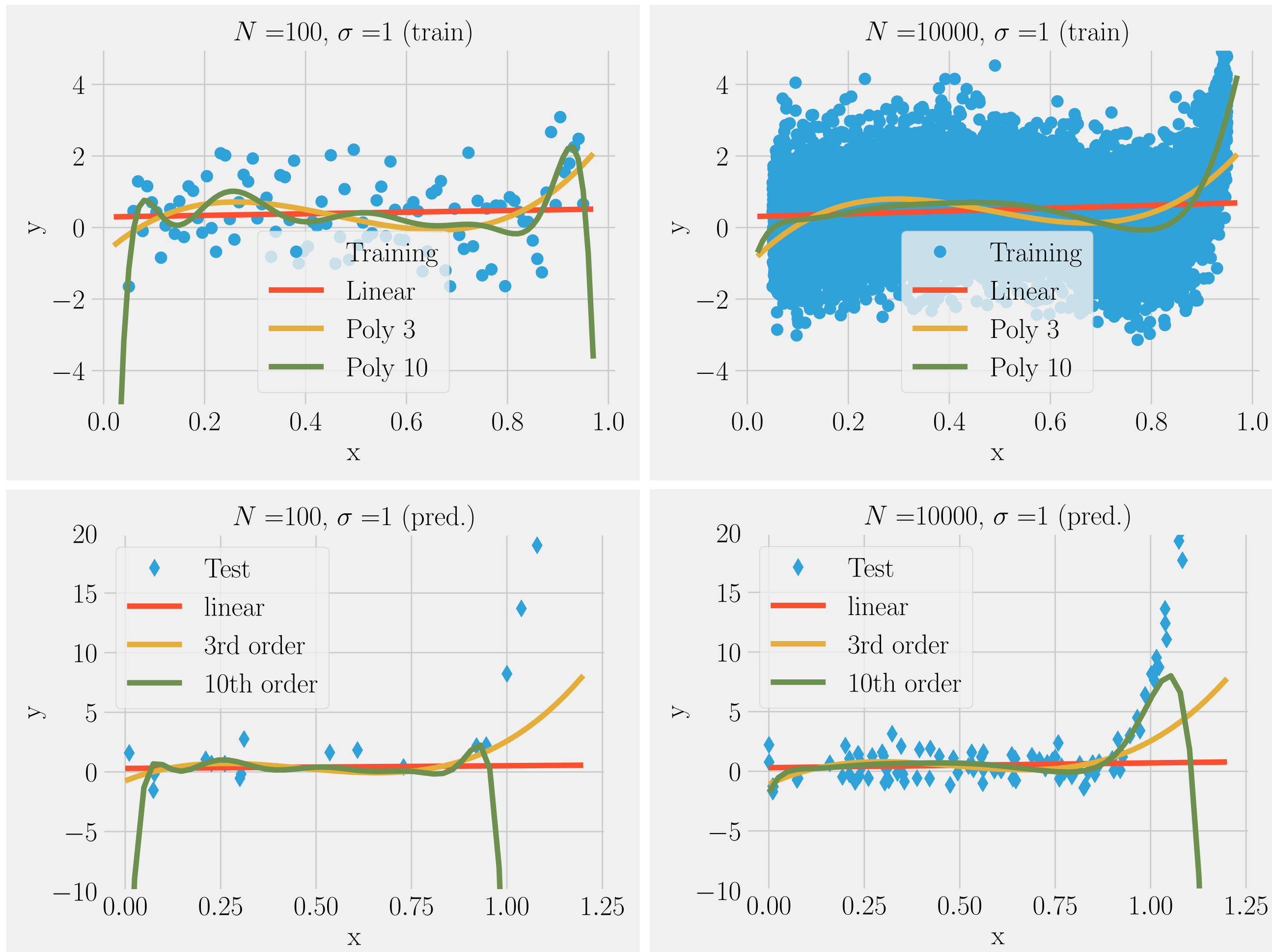
$$\langle \epsilon_i \rangle = 0$$

$$\langle \epsilon_i \epsilon_j \rangle = \sigma^2 \delta_{ij}$$

Cost function:

$$f_C(X; f_M(w)) = \sum_i (y_i - f_M(x_i; w))^2$$

Mean squared error (MSE)



Mehta et al., arXiv:1803.08823 (2018)

Machine learning is hard

Example: polynomial regression

Dataset $X = (x_i, y_i)$
sampled from:

$$y_i = f(x_i) + \epsilon_i$$

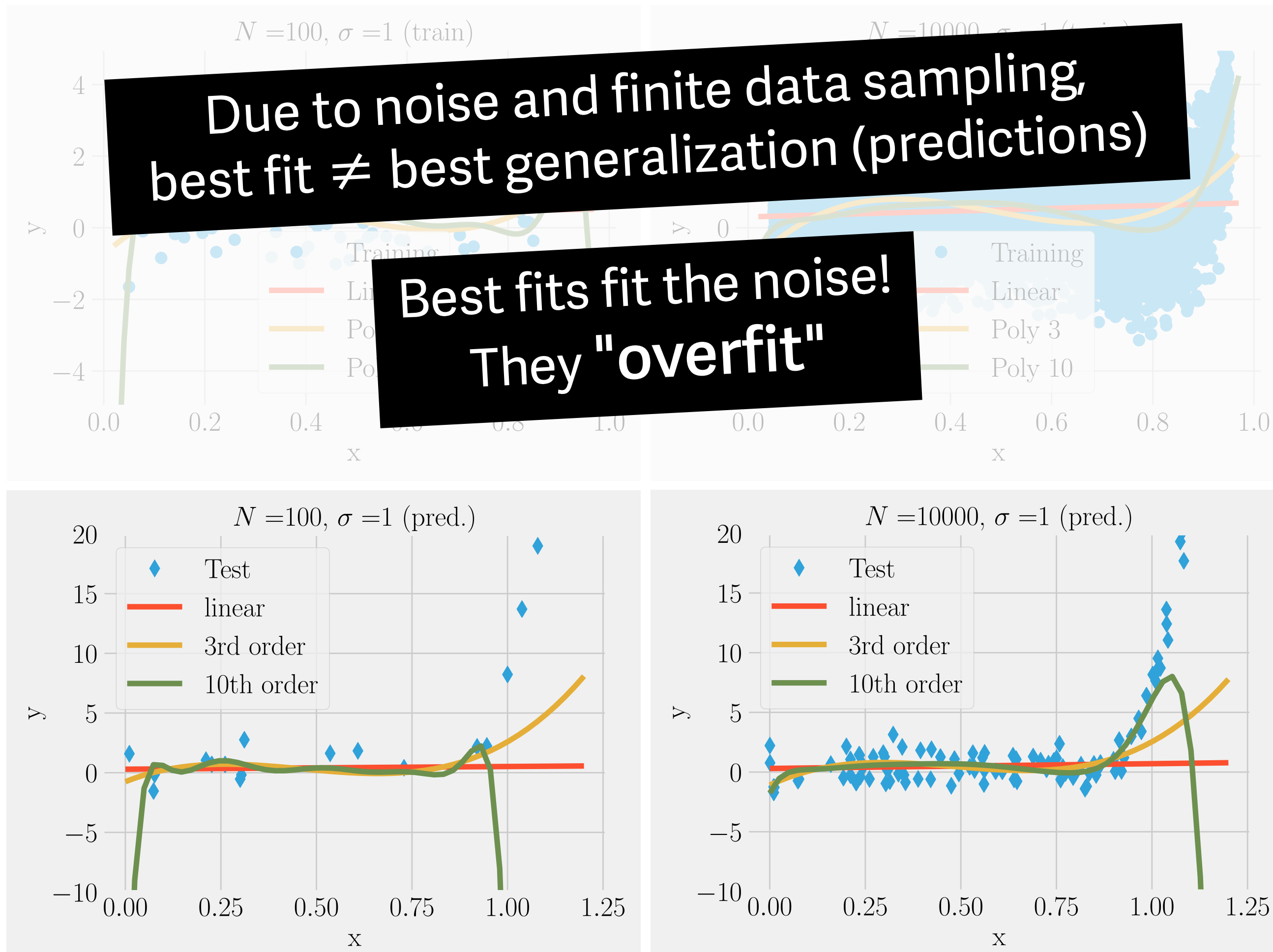
Unknown function — **Noise** (e.g., Gaussian)

$$\langle \epsilon_i \rangle = 0$$
$$\langle \epsilon_i \epsilon_j \rangle = \sigma^2 \delta_{ij}$$

Cost function:

$$f_C(X; f_M(w)) = \sum_i (y_i - f_M(x_i; w))^2$$

Mean squared error (MSE)



Due to noise and finite data sampling, best fit ≠ best generalization (predictions)

Best fits fit the noise! They **"overfit"**

Mehta et al., arXiv:1803.08823 (2018)

Machine learning is hard

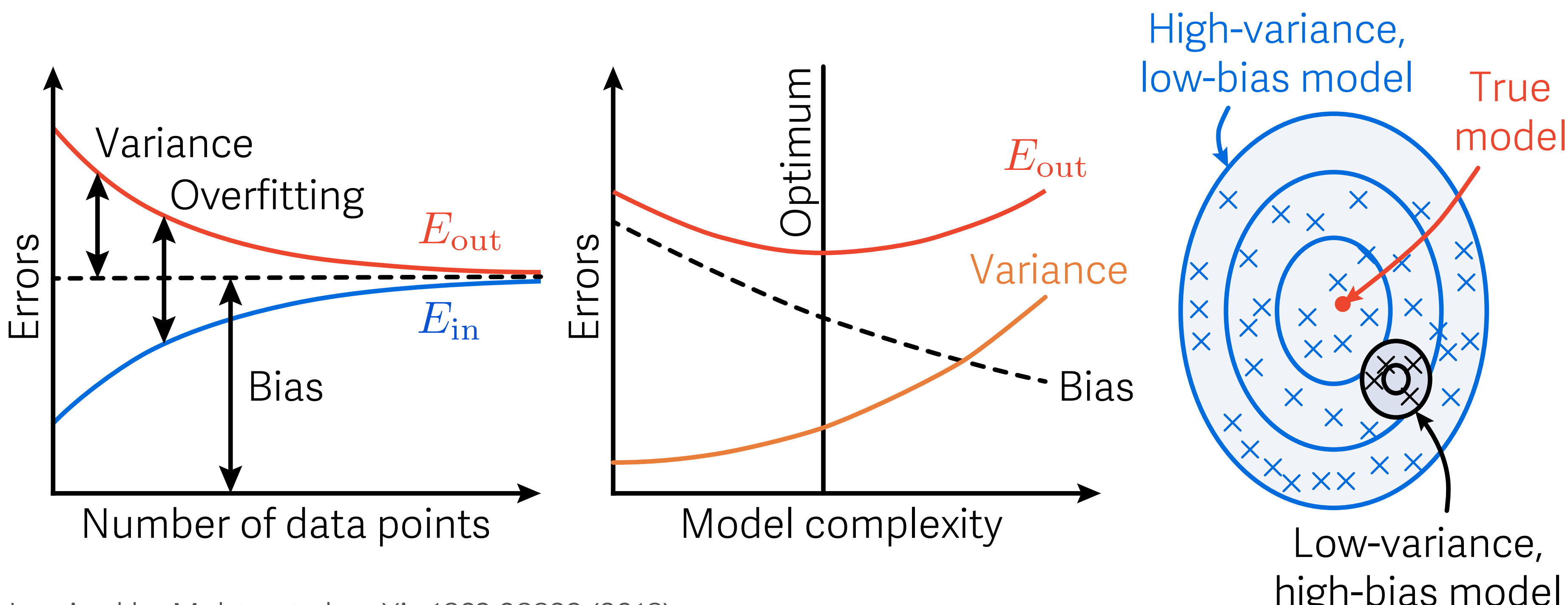
Generic difficulties in 3 plots

Training dataset → Model with optimized parameters from training

In-sample error: $E_{\text{in}} = f_C(X_{\text{train}}; f_M(w_{\text{opt}}))$

Out-of-sample error: $E_{\text{out}} = f_C(X_{\text{test}}; f_M(w_{\text{opt}}))$ ← **Quantifies the generalizing (predicting) performance** of the trained model

Test dataset

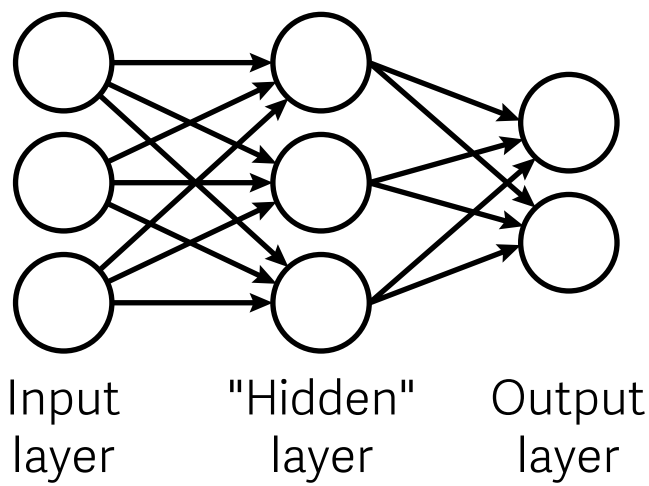


Inspired by Mehta et al., arXiv:1803.08823 (2018)

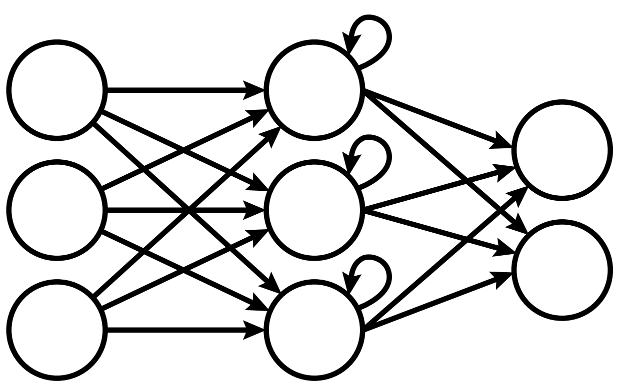
Deep learning

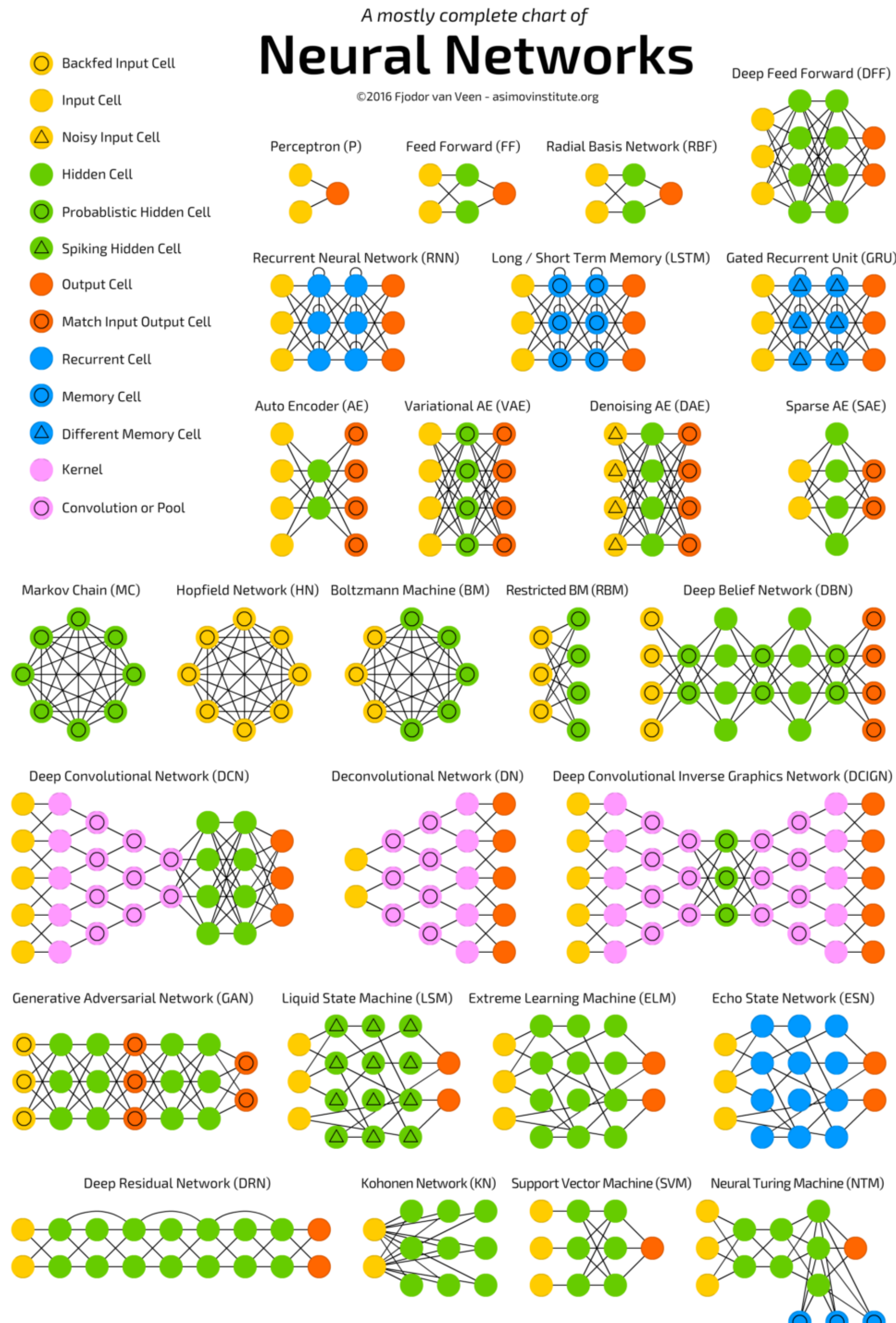
Combine neurons into layers

Feed-forward neural networks



Recurrent neural networks

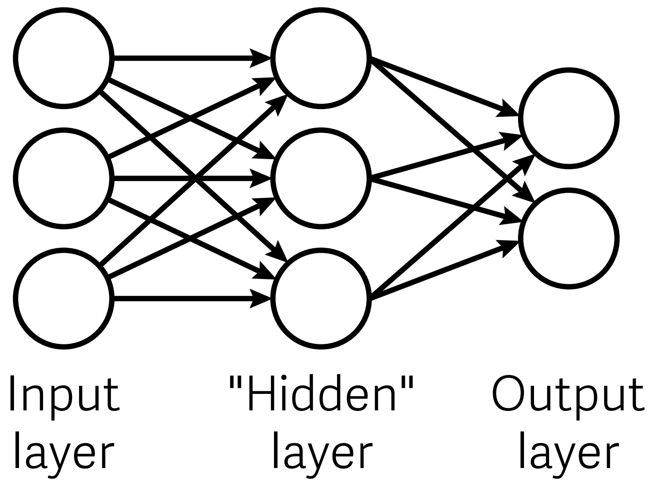




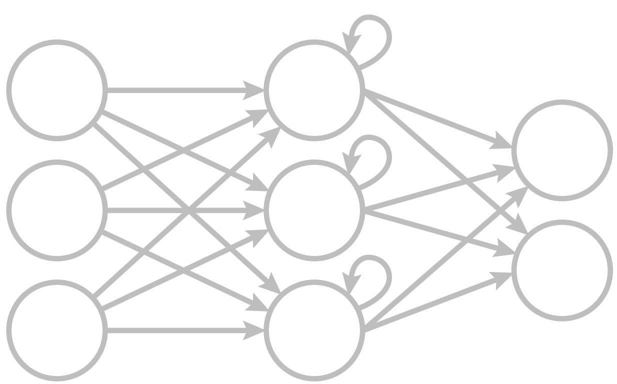
Deep learning

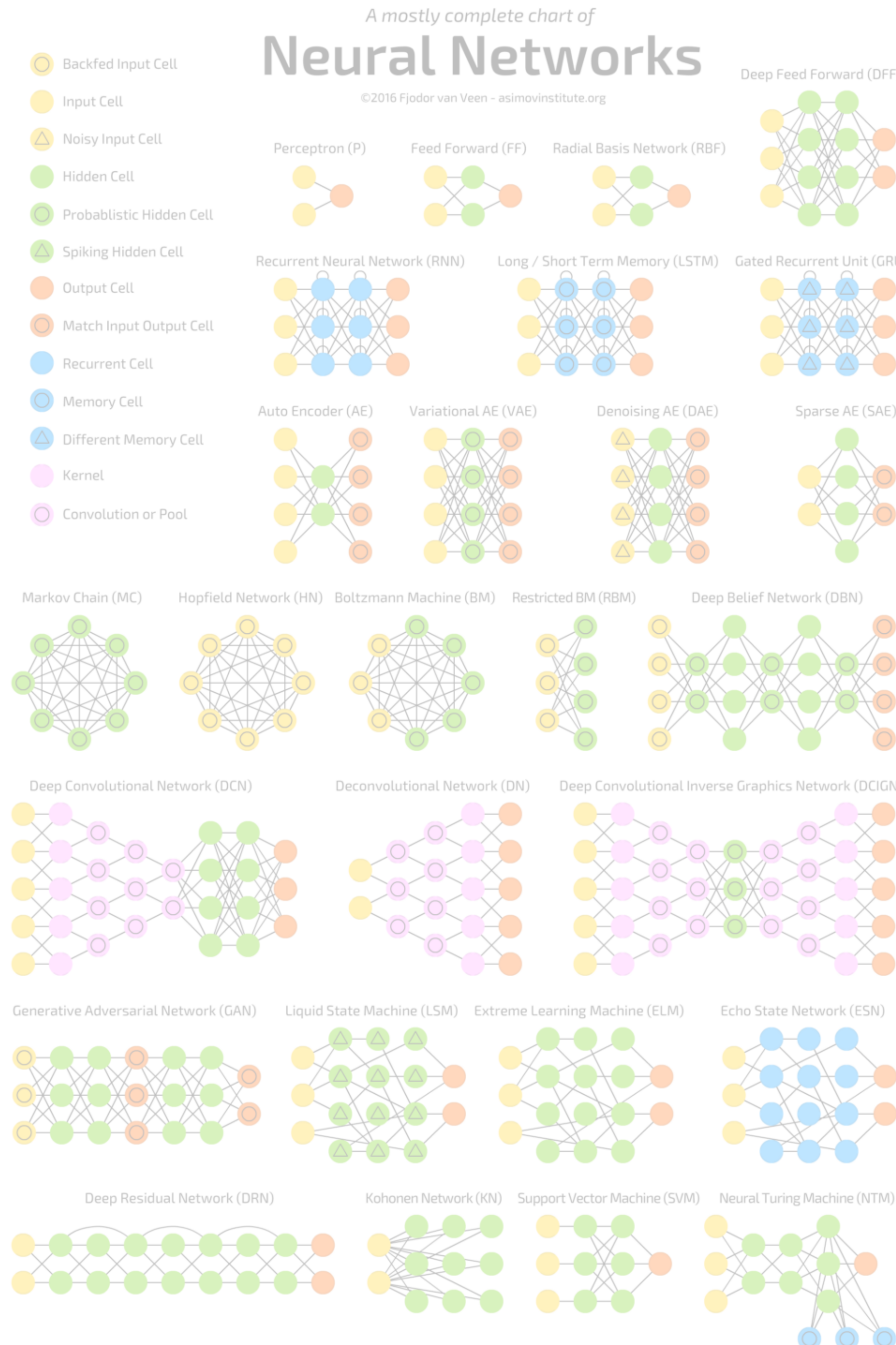
Combine neurons into layers

Feed-forward neural networks



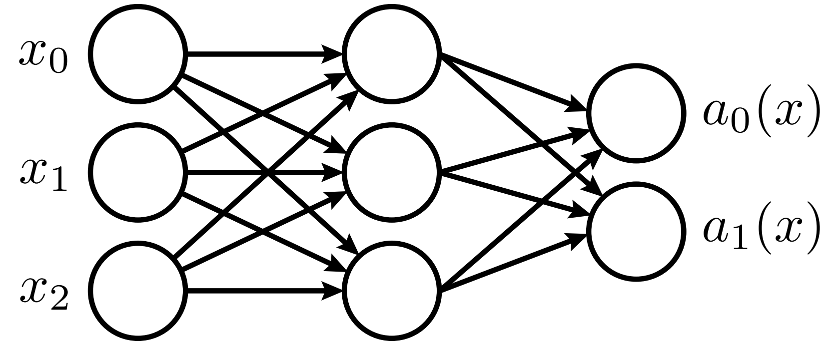
Recurrent neural networks





Feed-forward neural nets

How do they learn?



"Learning" or "training" = minimizing the chosen cost function

E.g., **mean squared error** (MSE)

$$f_C(w) = \frac{1}{N_{\text{train}}} \sum_x ||y(x) - a(x)||^2$$

Model parameters (weights and biases)

Expected output

Actual output (neuron activations)

Number of training data points (vectors) x

Learning algorithm

Batch gradient descent

$$\Delta w = -\eta \nabla f_C$$

Learning rate

Gradient estimated from the **whole training data (batch)**

Stochastic gradient descent
(typically better)

Gradient estimated from **one data point**

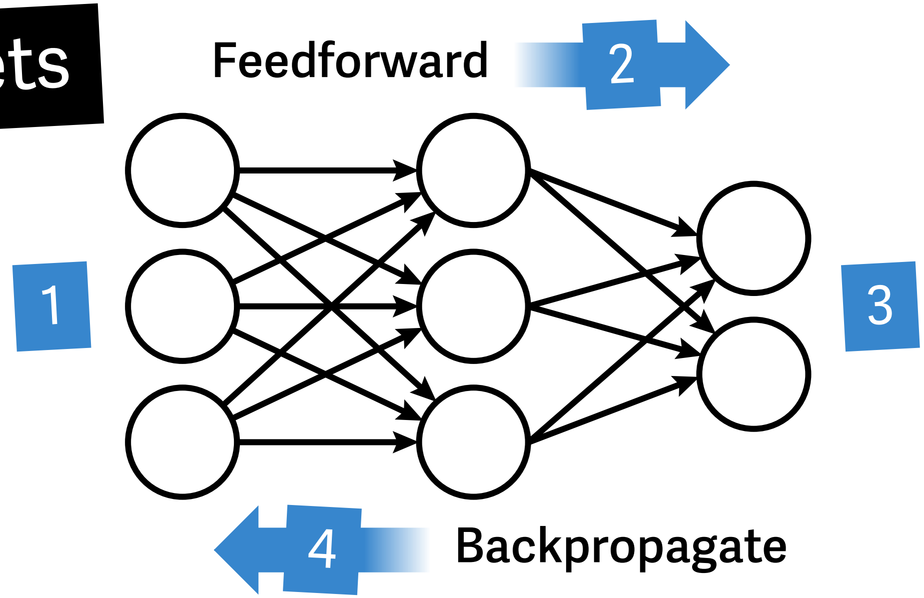
Mini-batch gradient descent
(typically even better)

Gradient estimated from **subsets of data points (mini-batches)**

Feed-forward neural nets

Backpropagation

Or how to **compute the gradient of the cost function efficiently**



Main steps

1 Compute the input activations: $a^{(1)} = f_A(x)$

2 **Feedforward:** Compute $z^{(l)} = w^{(l)} a^{(l-1)} + b^{(l)}$ and $a^{(l)} = f_A(z^{(l)})$ for successive layers $l = 2, 3, \ldots, L$

3 Compute the output error: $\delta^{(L)} = \nabla_a f_C \odot f'_A(z^{(L)})$

Comes from the usual chain rule

4 **Backpropagate the error:** Compute $\delta^{(l)} = [(w^{(l+1)})^T \delta^{(l+1)}] \odot f'_A(z^{(l)})$ for successive layers $l = L-1, L-2, \ldots, 2$

Output

$$\frac{\partial f_C}{\partial w_{jk}^{(l)}} = a_k^{(l-1)} \delta_j^{(l)} \qquad \frac{\partial f_C}{\partial b_j^{(l)}} = \delta_j^{(l)}$$

Gradient computed from only two passes (forward and backward)

Machine learning is hard

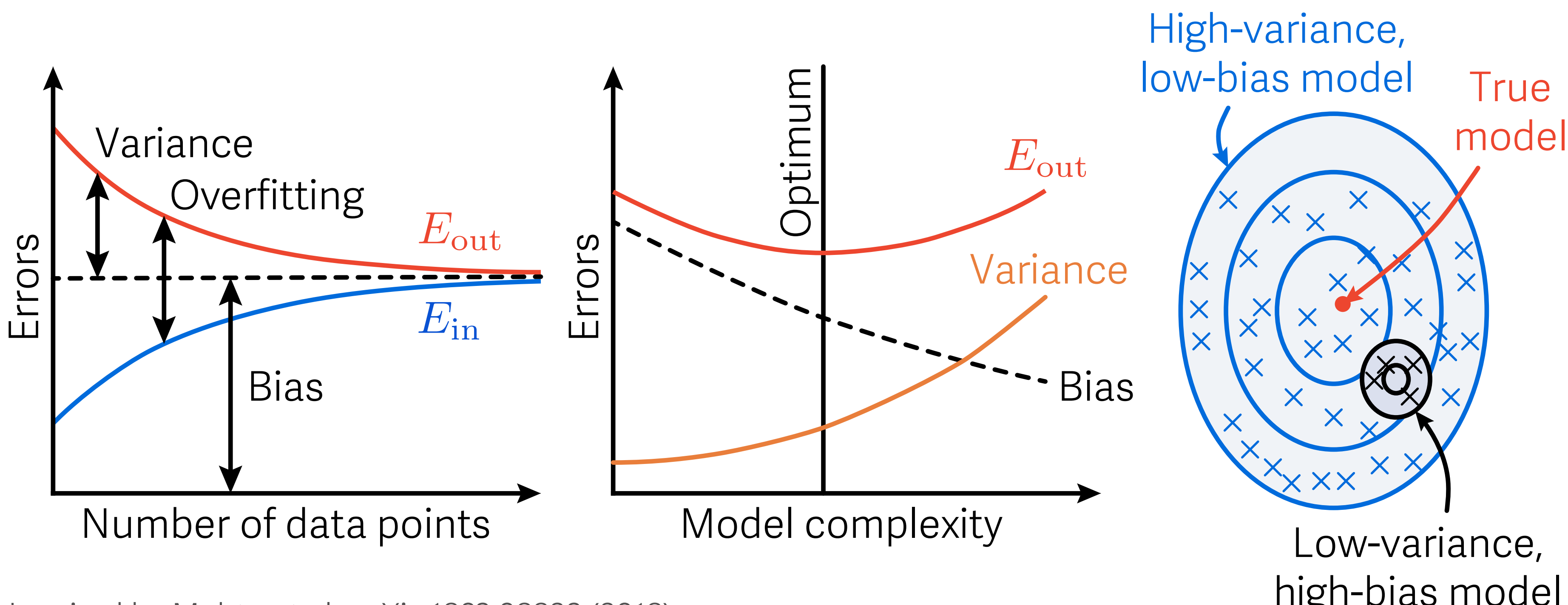
Generic difficulties in 3 plots

Training dataset → X_{train}; Model with optimized parameters from training → $f_M(w_{\text{opt}})$

In-sample error: $E_{\text{in}} = f_C(X_{\text{train}}; f_M(w_{\text{opt}}))$

Out-of-sample error: $E_{\text{out}} = f_C(X_{\text{test}}; f_M(w_{\text{opt}}))$ ← **Quantifies the generalizing (predicting) performance** of the trained model

Test dataset → X_{test}



Inspired by Mehta et al., arXiv:1803.08823 (2018)

Machine learning is hard

Generic difficulties in 3 plots

Training dataset

Model with optimized parameters from training

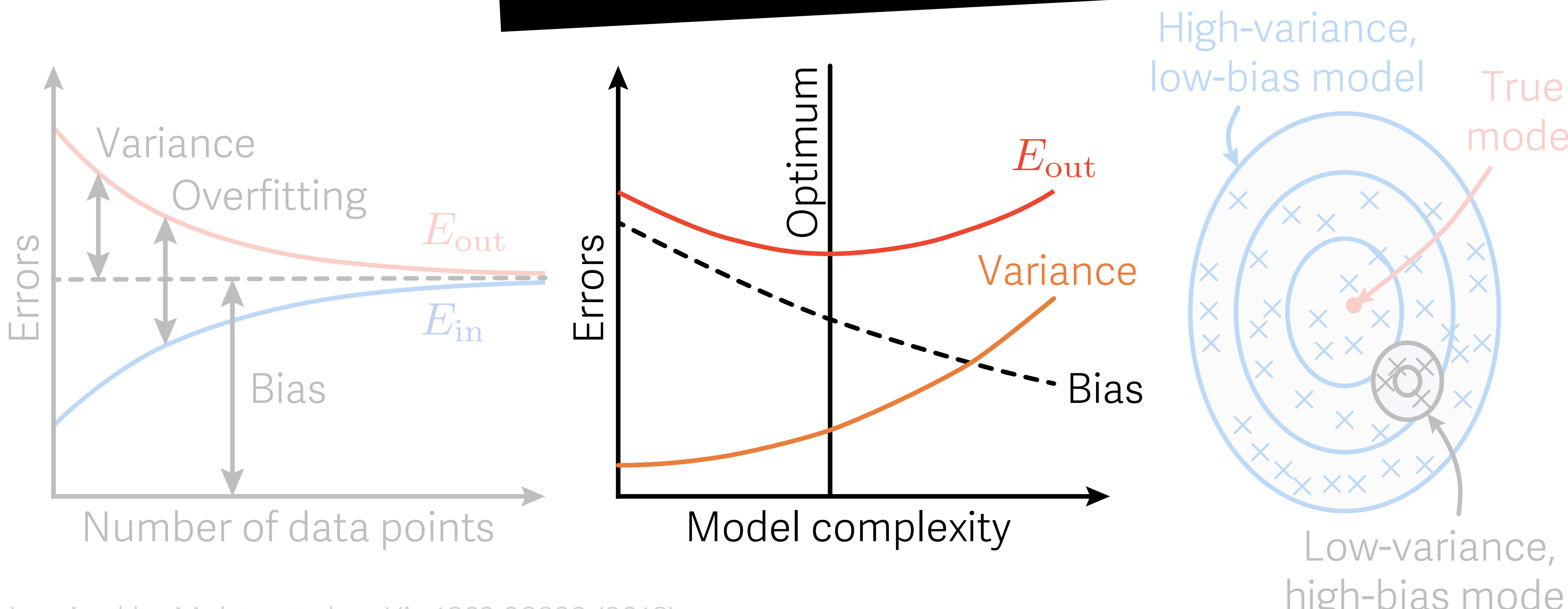
In-sample error: $E_{\text{in}} = f_C(X_{\text{train}}; f_M(w_{\text{opt}}))$

Out-of-sample error: E

...ntifies the generalizing (...dicting) performance
...f the trained model

T

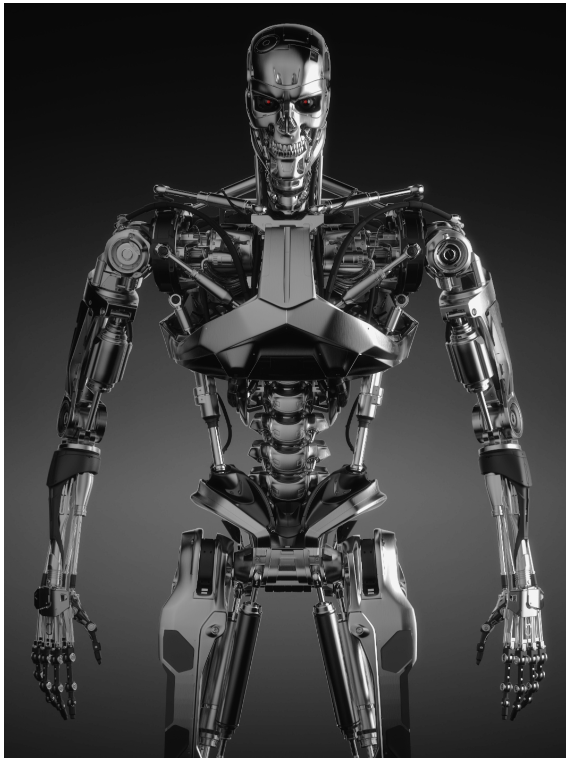
Modern deep learning models are very complex. They should massively overfit!



Inspired by Mehta et al., arXiv:1803.08823 (2018)

What is machine learning?

Not killer robots



Not friendly robots

Some

~~Not~~ magic

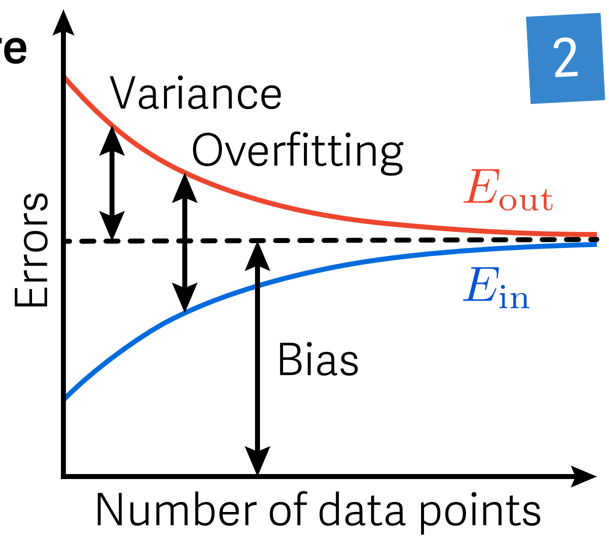
Deep learning

How to reduce overfitting effects

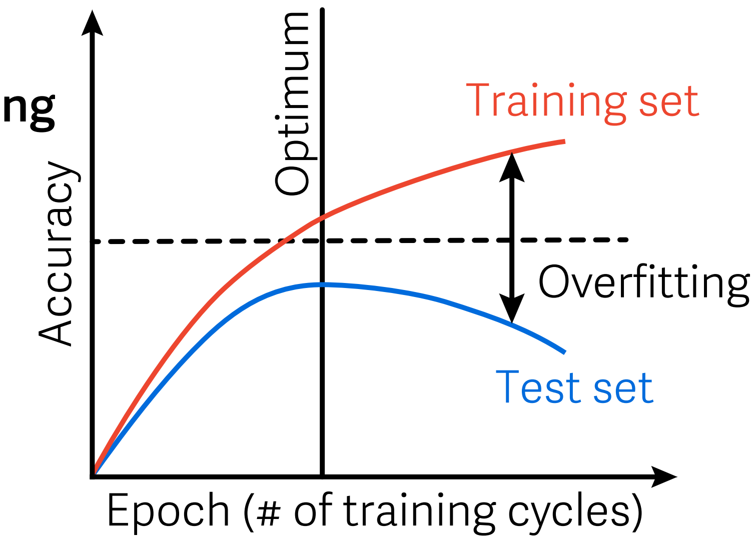
Recall: Overfitting is the fitting of random noise due to too large model complexity and/or too small amount of data

Several options

1 Get **more data!**



2 **Early stopping**



3 **Regularization** (L1, L2, etc.)

Add a term to the cost function to **penalize large weights**

$$f_C \to f_C + \frac{\lambda}{N_{\text{train}}} \sum_w |w|^n$$

4 **Dropout**

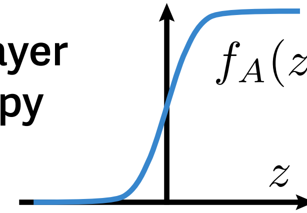
Remove a (random) subset of neurons before each gradient computation

Effectively reduces the number of model parameters (ability to fit the noise)

Deep learning

Something of an art

Sigmoid output layer and **cross-entropy cost function**



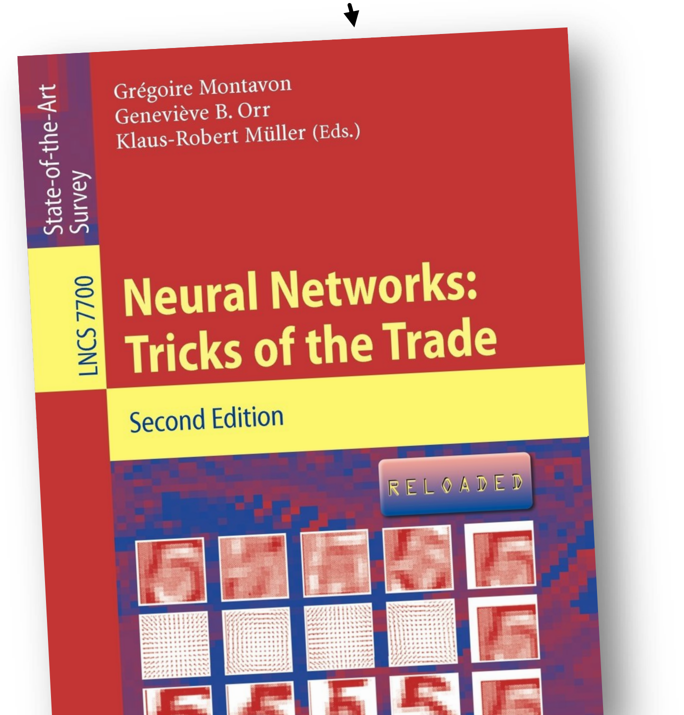
$$f_A(z) = \frac{1}{1+e^{-z}}$$

$$f_C = -\frac{1}{N_{\text{train}}} \sum_x [y \ln a + (1-y) \ln(1-a)]$$

Many other tricks of the trade

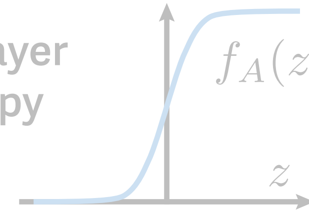
- Choice of **activation and cost functions**
  - To avoid learning slowdowns
- Choice of **hyperparameters** (learning rate, mini-batch size, etc.)
  - Grid search, Bayesian optimization, etc.
- **Parameters initialization** (weights and biases)
- **Improved gradient descents**
  - Hessian methods (computationally expensive), **momentum-based gradient descent** (better), etc.
- And on and on...

One good reference



Deep learning
Something of an art

Sigmoid output layer and **cross-entropy cost function**



$$f_A(z) = \frac{1}{1+e^{-z}}$$

$$f_C = -\frac{1}{N_{\text{train}}} \sum_x [y \ln a + (1-y) \ln(1-a)]$$

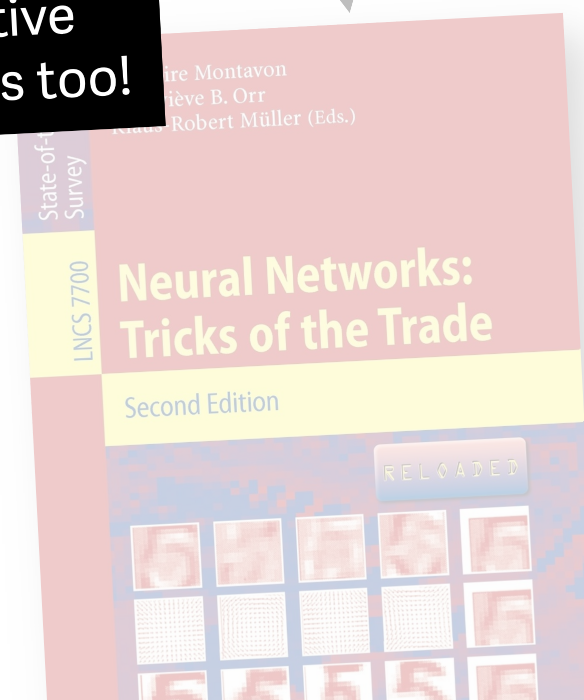
Many other tricks of the trade

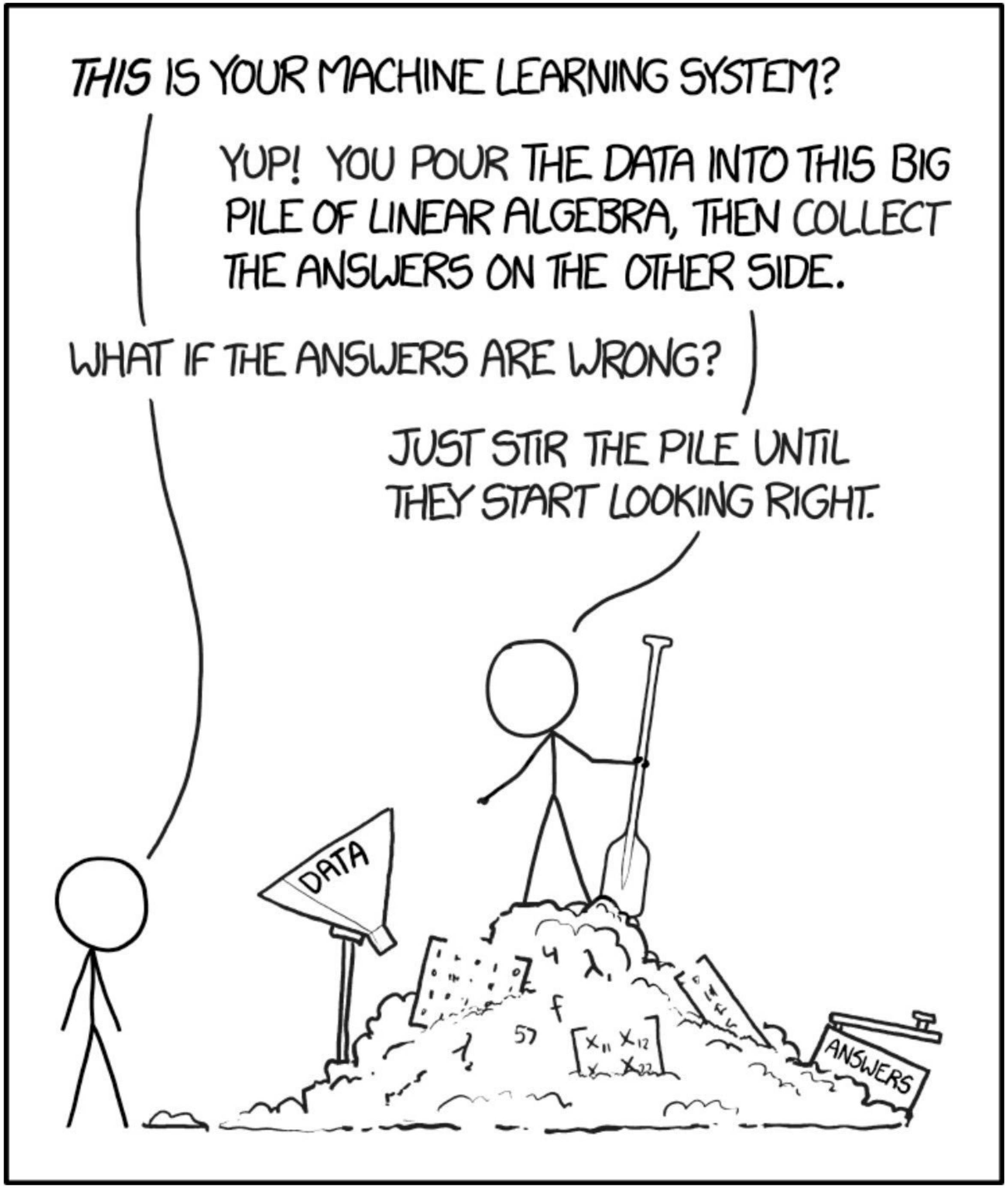
- Choice of **activation and**
  To avoid learning slo
- Choice of **hyperparameters** (learning rate, mini-b
  Grid search, Bay
- **Parameters initialization** (weights and biases)
- **Improved gradient descents**
  Hessian methods (computationally expensive), **momentum-based gradient descent** (better), etc.
- And on and on...

Machine learning is partly empirical ...like physics!

What matters is the predictive power of models... like physics too!

One good reference




THIS IS YOUR MACHINE LEARNING SYSTEM?
YUP! YOU POUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.
WHAT IF THE ANSWERS ARE WRONG?
JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT.
DATA
ANSWERS

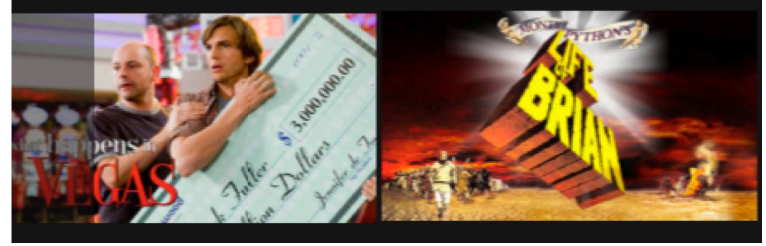
What can machine learning do?

Google

Will quantum computers

will quantum computers **break blockchain**
will quantum computers **threaten modern cryptography**
will quantum computers **kill bitcoin**
will quantum computers **replace**
will quantum computers **break bitcoin**
will quantum computers **break encryption**
will quantum computers **ever work**
will quantum computers **work**
will quantum computers **cure cancer**

NETFLIX

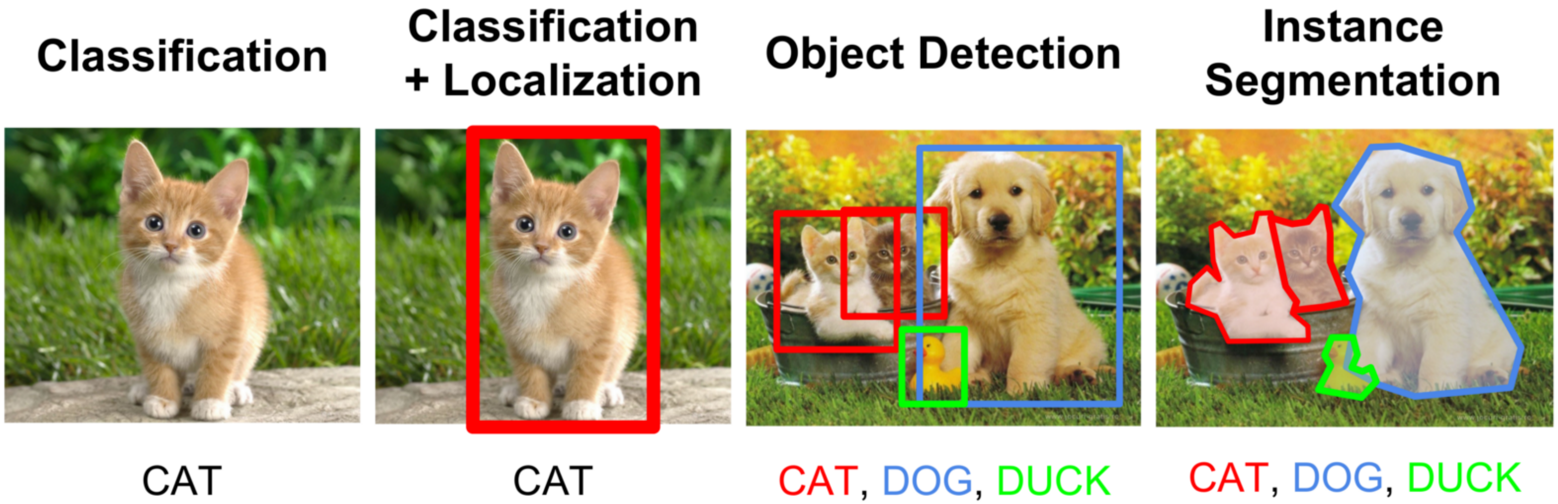


amazon

We Have Recommendations for You

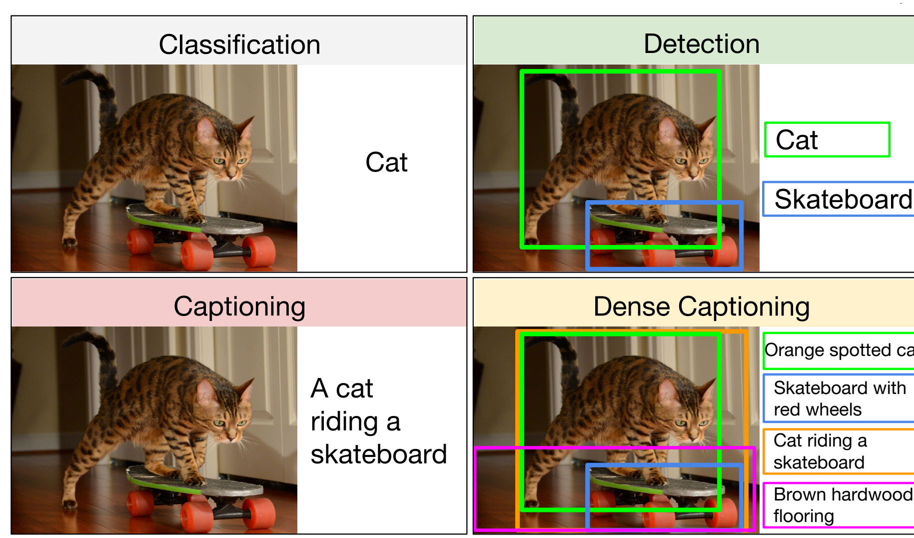
Sign in to see personalized recommendations

Computer vision



With super- or near-human performances

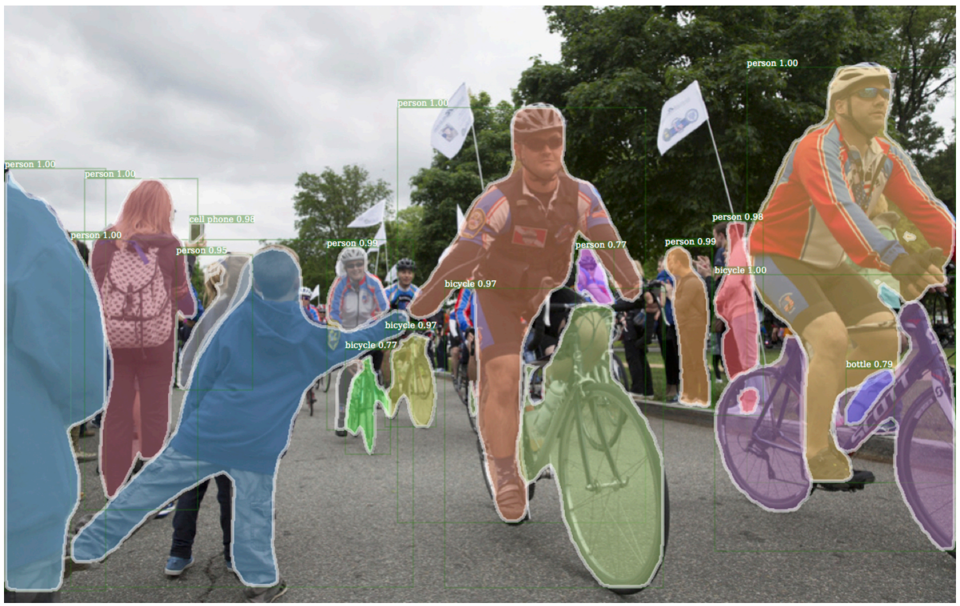
Handwriting recognition,
face recognition (Facebook, etc.),
human pose estimation,
motion recognition (Xbox Kinect, etc.),
human action recognition,
etc.



Johnson et al., arXiv:1511.07571 (2015)

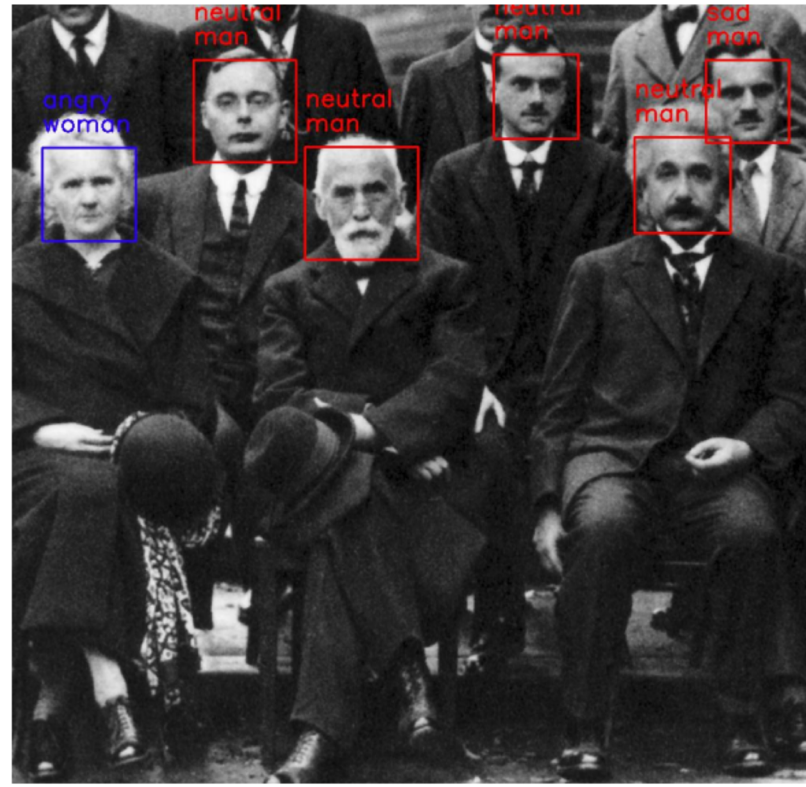
Computer vision

Object / people detection



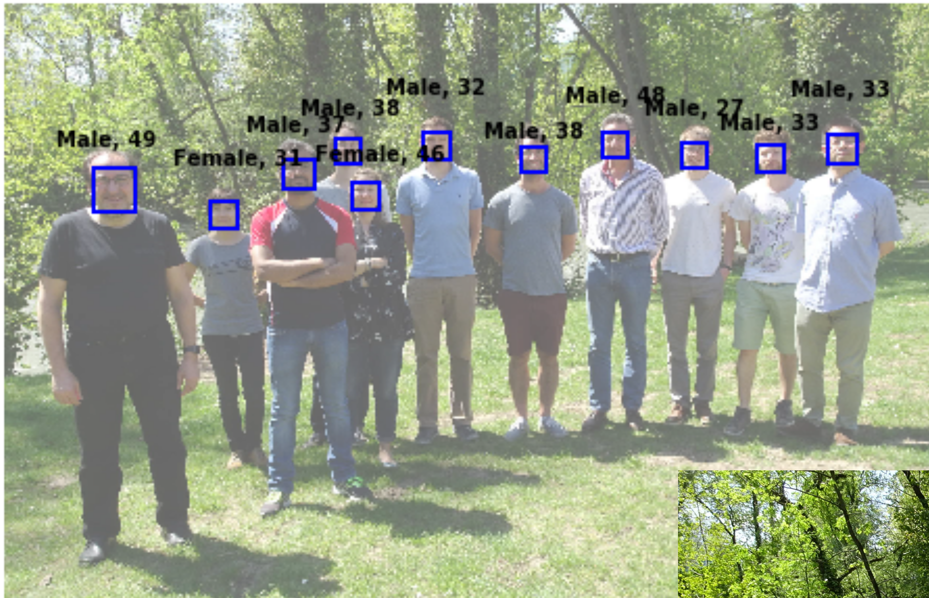
Detectron, Facebook AI Research (FAIR) (2018)

Emotion / gender classification



Arriaga et al., arXiv:1710.07557 (2017)

Computer vision



Gender / age classification

Image captioning

A group of people standing in front of a tree posing for the camera

Made with Microsoft Computer Vision and Face APIs

Natural language processing

Google Assistant

Apple Siri

Microsoft Cortana

Amazon Alexa

Google's robot assistant now makes eerily lifelike phone calls for you

Google Duplex contacts hair salon and restaurant in demo, adding 'er' and 'mmm-hmm' so listeners think it's human

Olivia Solon *in San Francisco*

Tue 8 May 2018 21.13 BST

Google's virtual assistant can now make phone calls on your behalf to schedule appointments, make reservations in restaurants and get holiday hours.

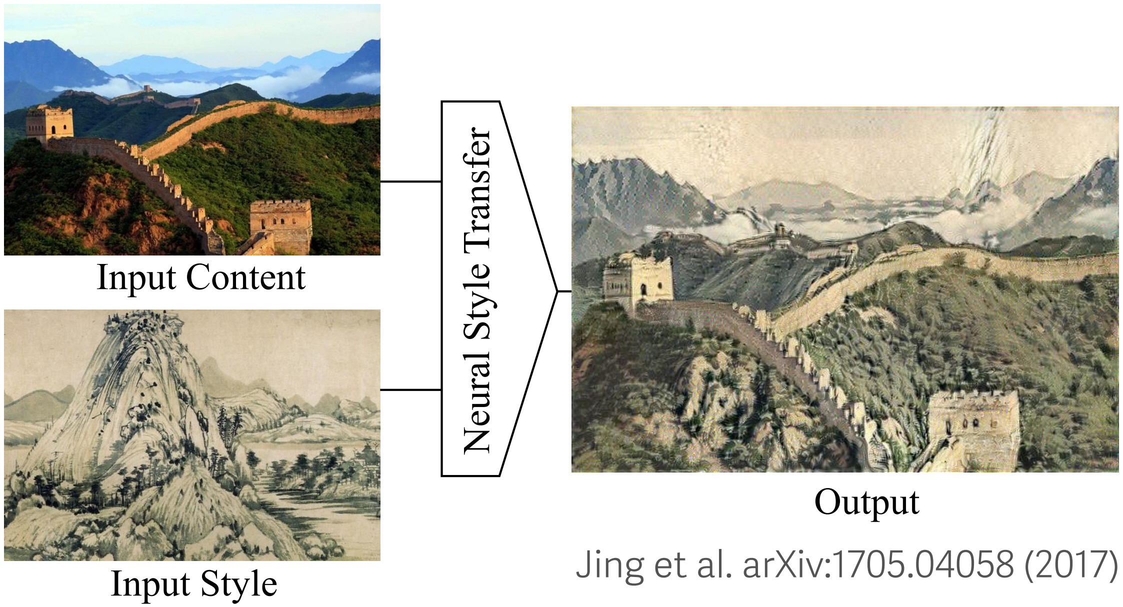
The robotic assistant uses a very natural speech pattern that includes hesitations and affirmations such as "er" and "mmm-hmm" so that it is extremely difficult to distinguish from an actual human phone call.

The unsettling feature, which will be available to the public later this year, is enabled by a

Near-human

Speech recognition, text-to-speech conversion, language translation, etc.

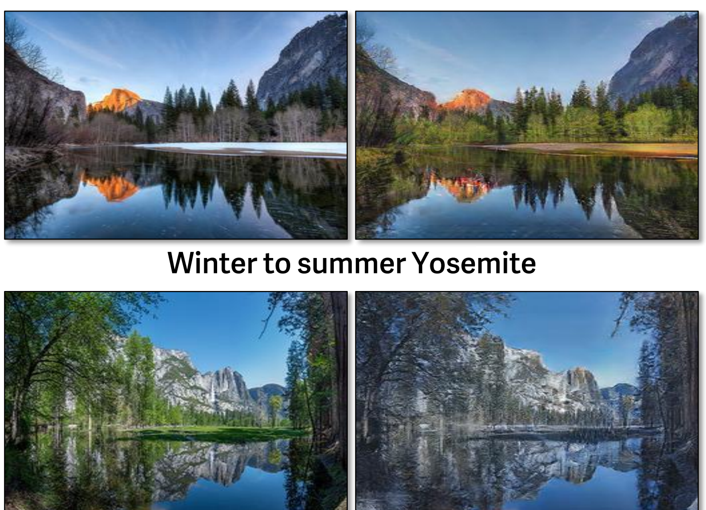
Image generation



Jing et al. arXiv:1705.04058 (2017)

Machine learning can also create / generate from examples

Using generative adversarial networks (GANs), in particular



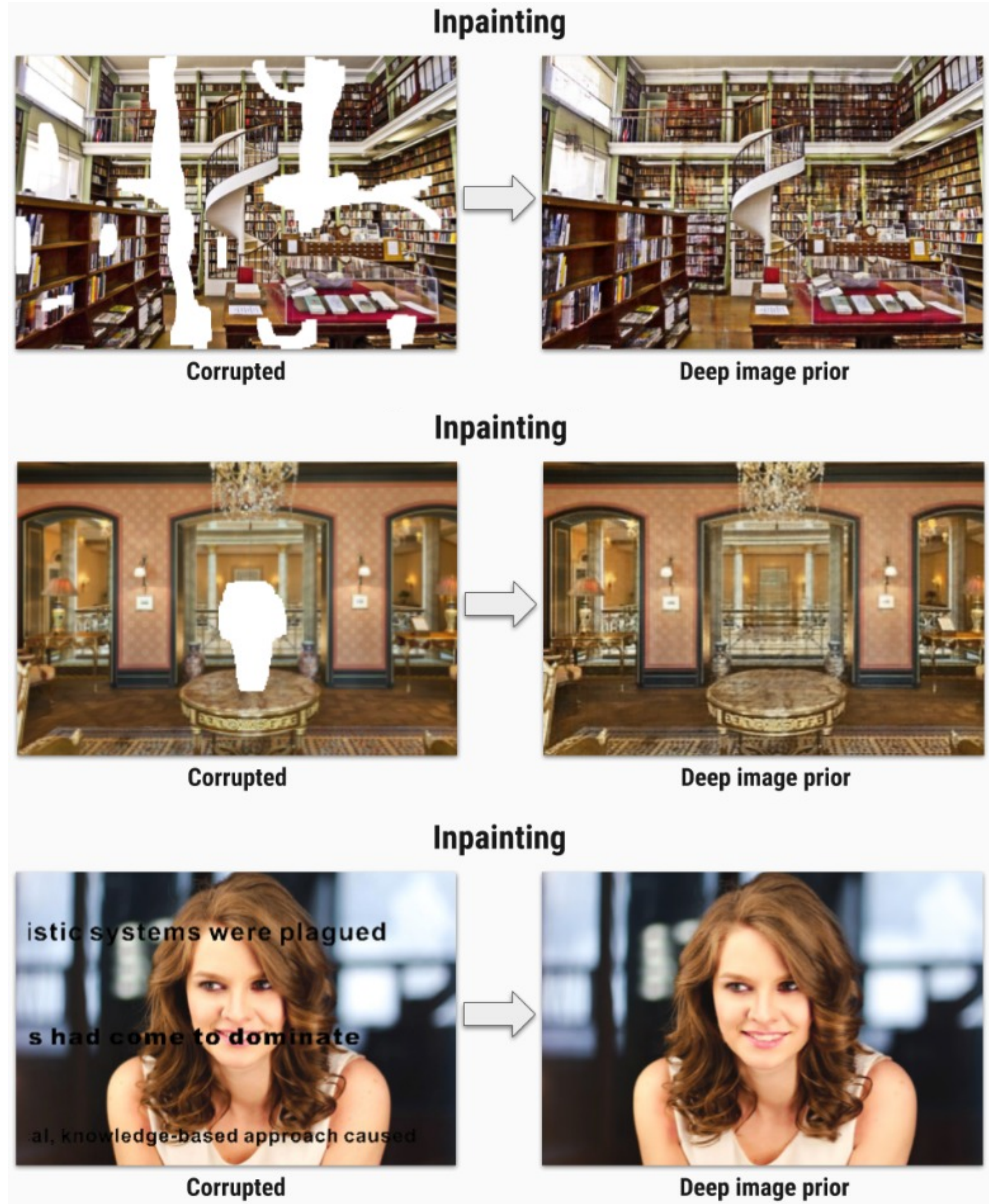
Winter to summer Yosemite

Zhu et al. arXiv:1703.10593 (2017)

Image generation

Can also generate likely missing parts from learned pictures

Using generative adversarial networks (GANs) too



Ulyanov et al., arXiv:1711.10925 (2017)

Music generation

The BachBot Challenge

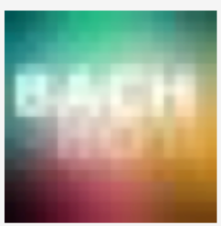
Can you tell the difference between Bach and a computer?

Challenge description

We will present you with some short samples of music which are either extracted from Bach's own work or generated by BachBot. Your task is to listen to both and identify the Bach originals.

To ensure fair comparison, all scores are transposed to C-major or A-minor and set to 120 BPM.

Want to Listen?



BachBot
Sample-1
SOUNDCLOUD
▶ 5K

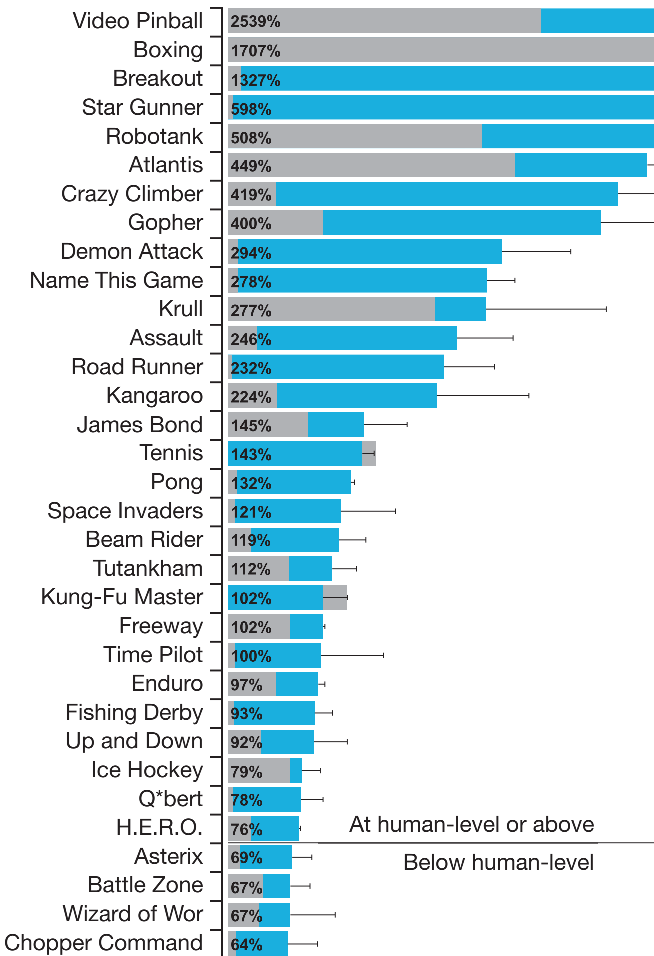
BachBot - Sample-1	▶ 5K
BachBot - Happy Birthday	▶ 3.4K
BachBot - Twinkle Twinkle Little Star	▶ 16K

Liang's thesis (2016), University of Cambridge

Game playing

Powered by deep reinforcement learning

Atari Games

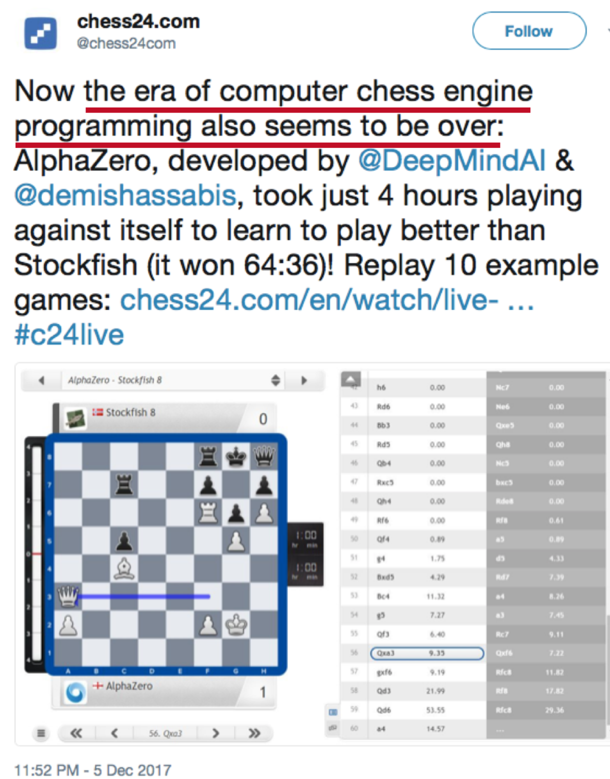


Deepmind,
Nature 518, 529 (2015)

Go

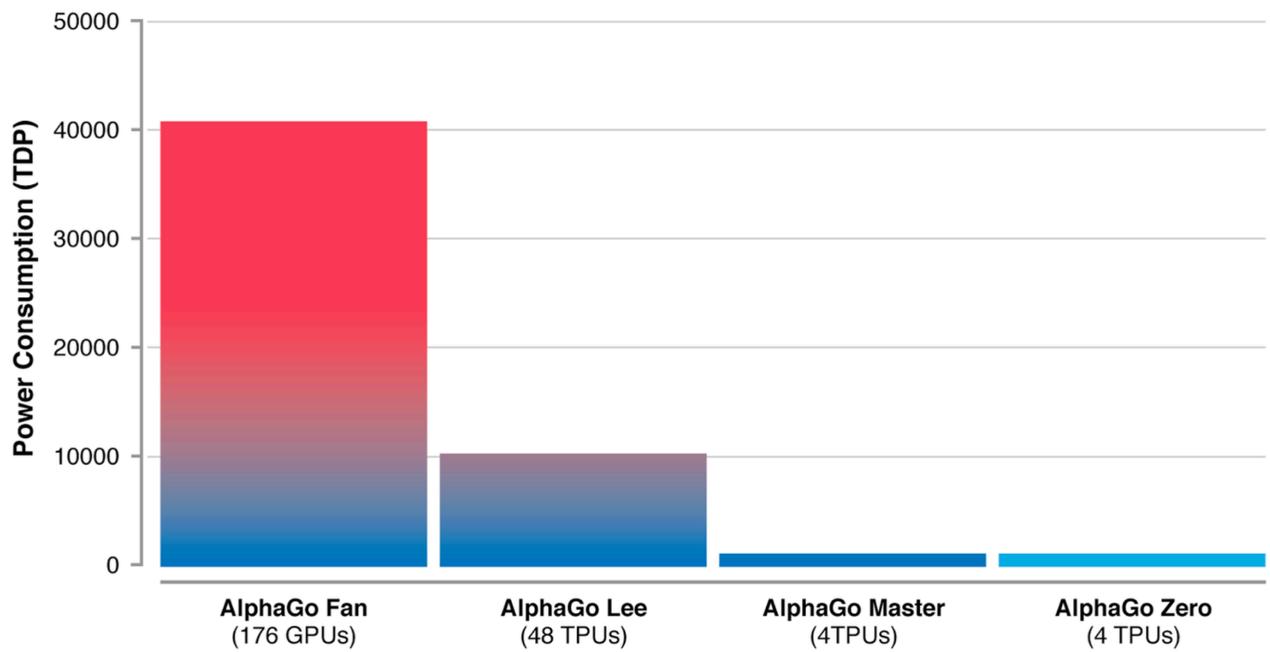
Deepmind's AlphaGo,
Nature 550, 354 (2017)

Chess



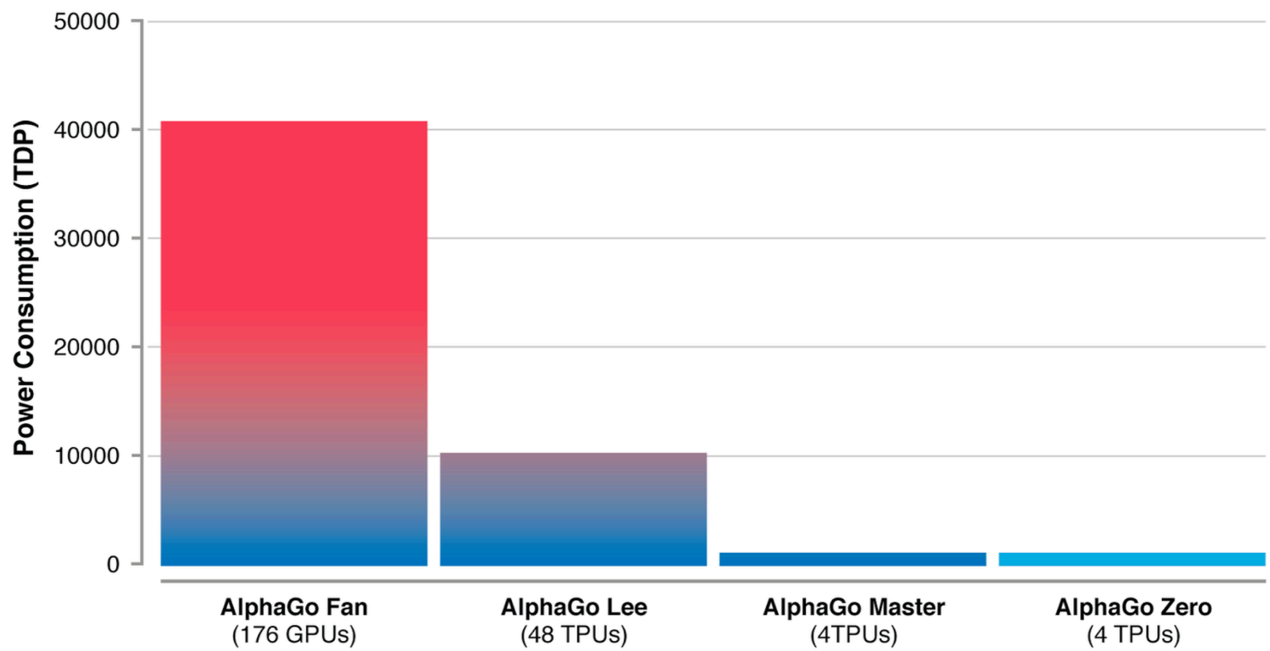
Deepmind's AlphaZero (2017)

Deepmind AlphaGo Zero (2017)



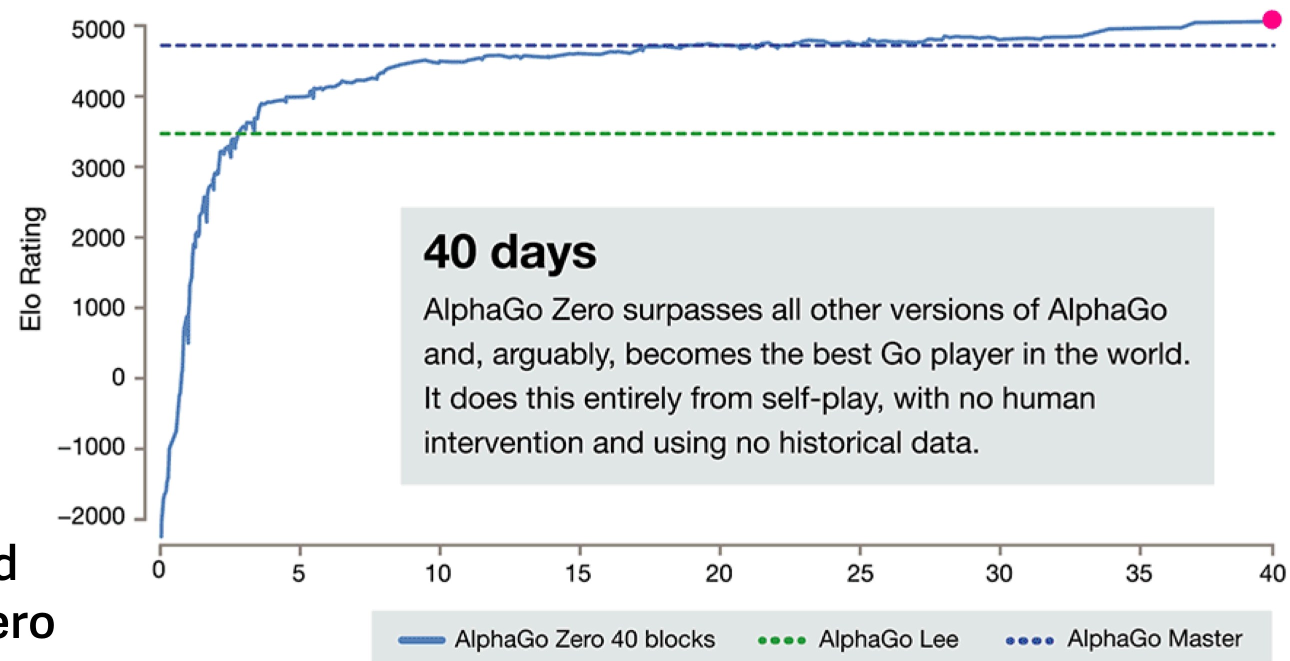
Pictures from Deepmind's blog

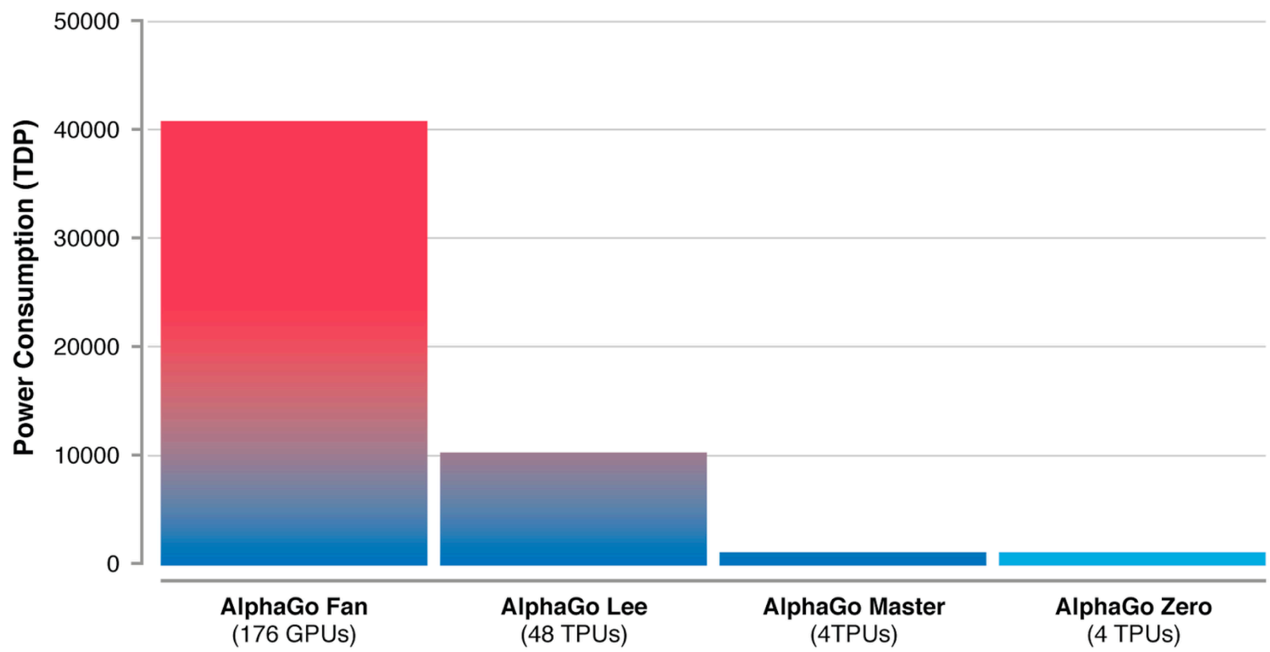
Deepmind AlphaGo Zero (2017)



Pictures from Deepmind's blog

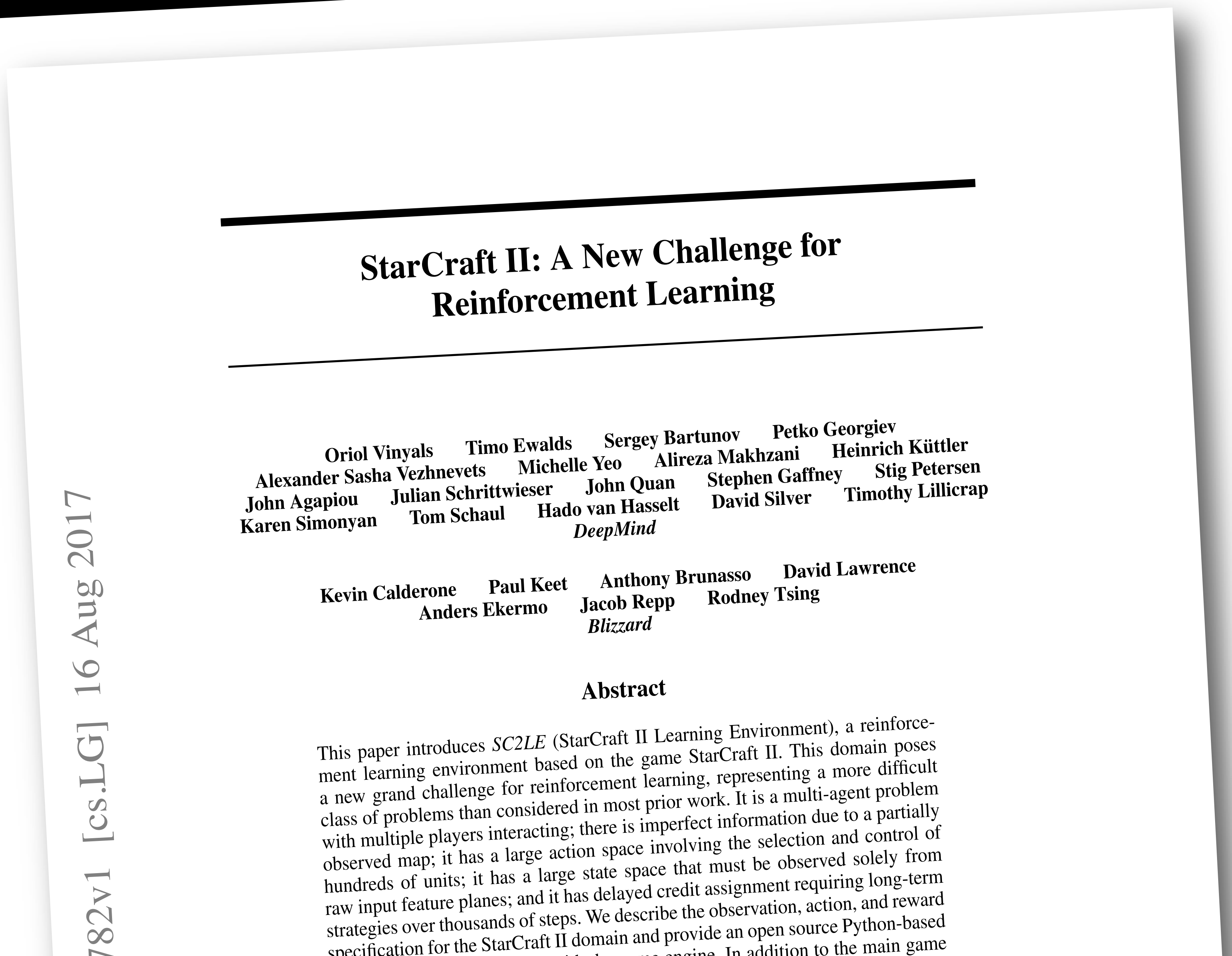
Deepmind AlphaGo Zero (2017)





Pictures from Deepmind's blog

Already moved on to more complex games



782v1 [cs.LG] 16 Aug 2017

StarCraft II: A New Challenge for Reinforcement Learning

Oriol Vinyals Timo Ewalds Sergey Bartunov Petko Georgiev
Alexander Sasha Vezhnevets Michelle Yeo Alireza Makhzani Heinrich Küttler
John Agapiou Julian Schrittwieser John Quan Stephen Gaffney Stig Petersen
Karen Simonyan Tom Schaul Hado van Hasselt David Silver Timothy Lillicrap
DeepMind

Kevin Calderone Paul Keet Anthony Brunasso David Lawrence
Anders Ekermo Jacob Repp Rodney Tsing
Blizzard

Abstract

This paper introduces *SC2LE* (StarCraft II Learning Environment), a reinforcement learning environment based on the game StarCraft II. This domain poses a new grand challenge for reinforcement learning, representing a more difficult class of problems than considered in most prior work. It is a multi-agent problem with multiple players interacting; there is imperfect information due to a partially observed map; it has a large action space involving the selection and control of hundreds of units; it has a large state space that must be observed solely from raw input feature planes; and it has delayed credit assignment requiring long-term strategies over thousands of steps. We describe the observation, action, and reward specification for the StarCraft II domain and provide an open source Python-based

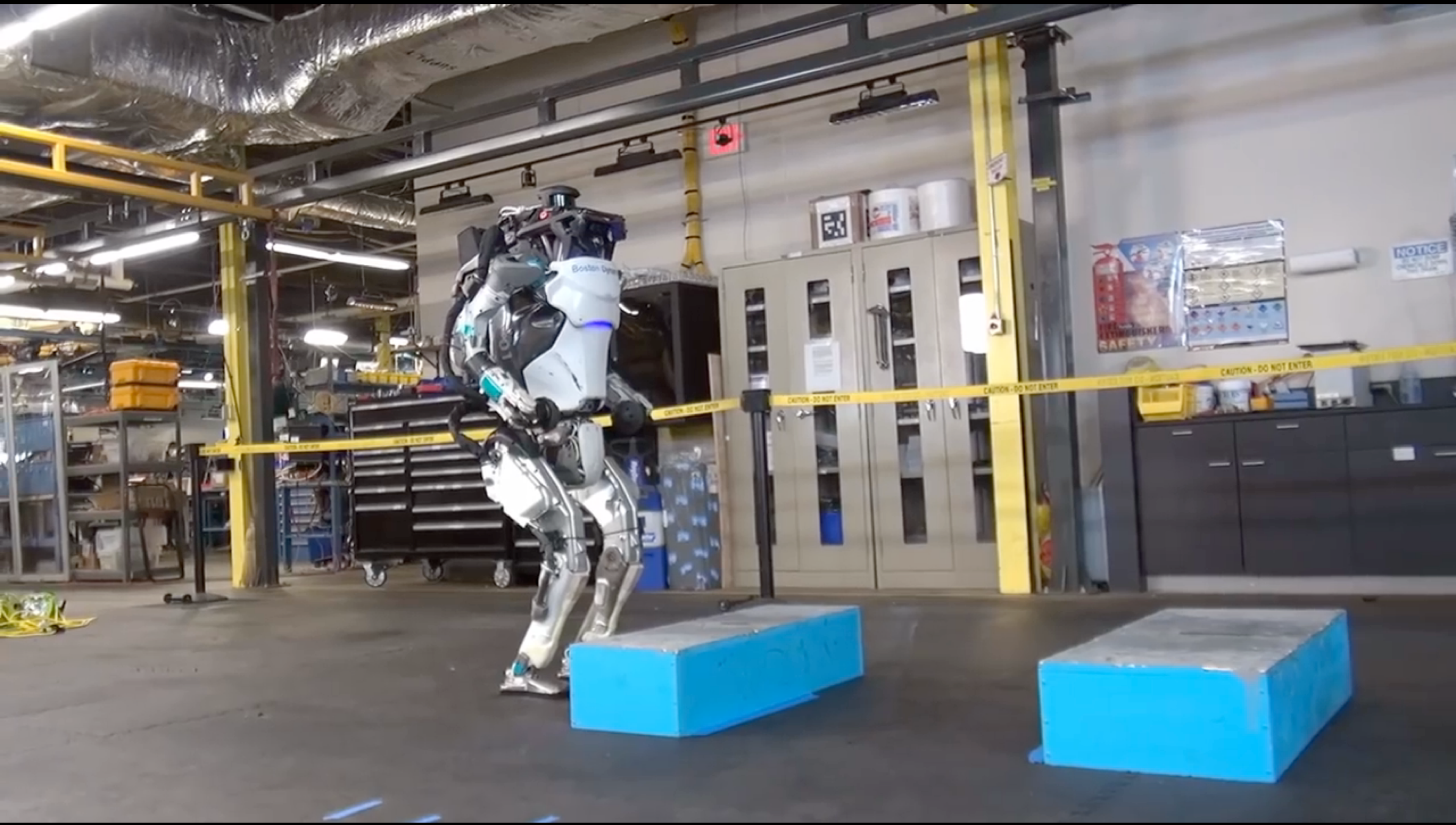
Self-driving cars

Waymo (Google) self-driving cars, February 28, 2018

"Only **Waymo has tested Level 4 vehicles** on passengers who aren't its employees. **No one has yet demonstrated at Level 5**, where the car is so independent that there's no steering wheel or pedals to operate."

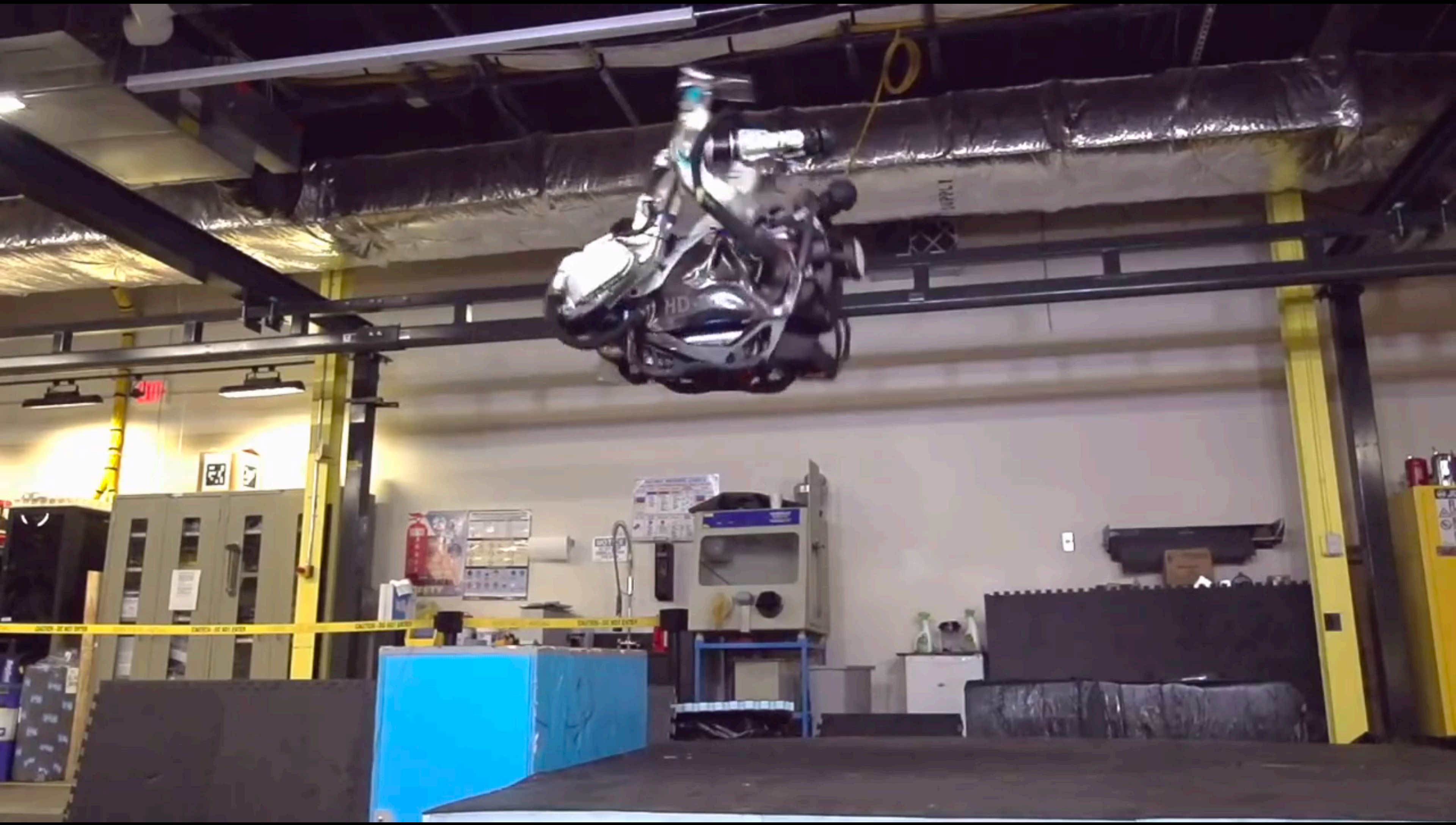
LA Times, May 11, 2018

Robotics

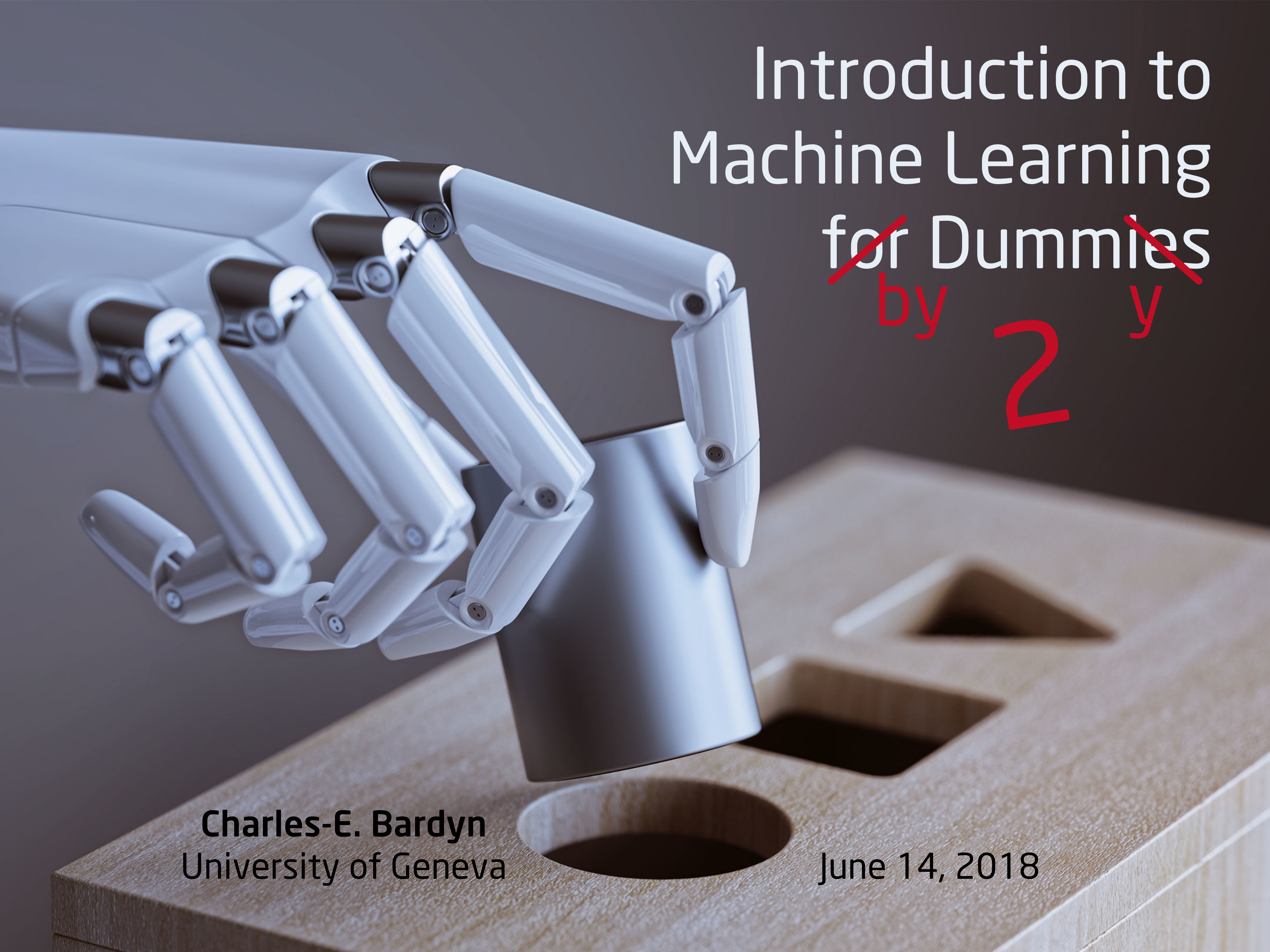


"Atlas" robot from Boston Dynamics (2018)

Robotics



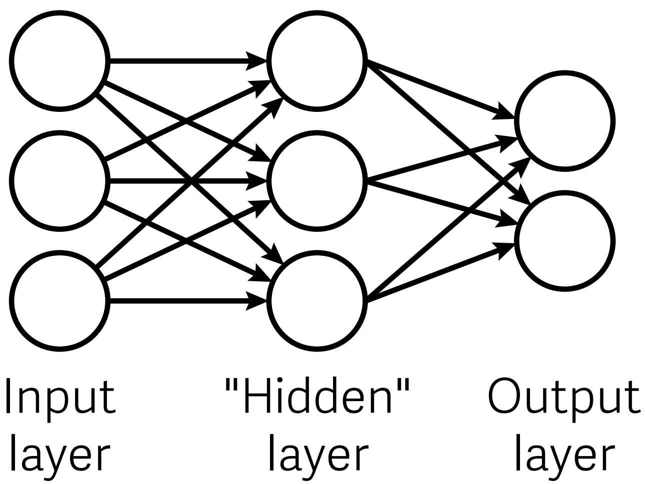
"Atlas" robot from Boston Dynamics (2018)


Introduction to
Machine Learning
~~for~~ Dummi~~e~~s
by y
2
Charles-E. Bardyn
University of Geneva
June 14, 2018

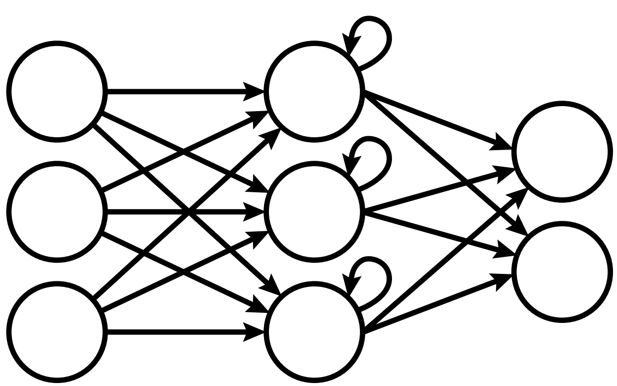
Total recall

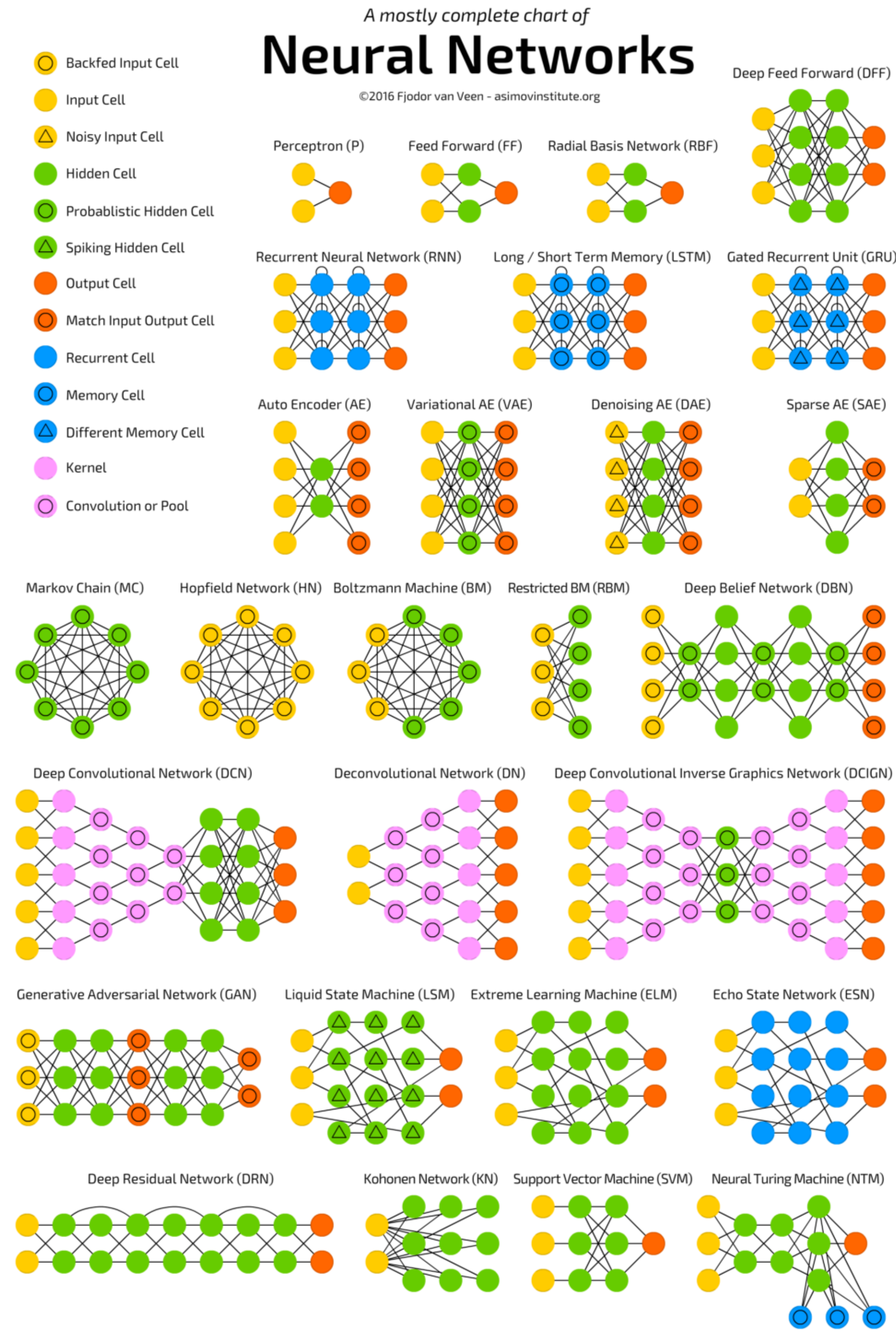
Deep learning: multi-layer networks

Feed-forward neural networks



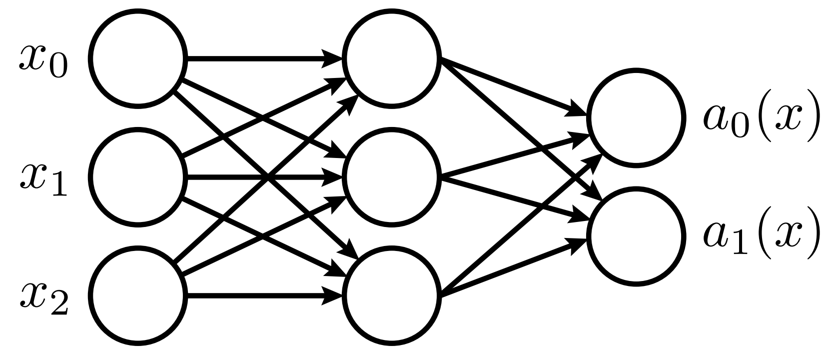
Recurrent neural networks





Total recall

How do deep neural nets learn?



"Learning" or "training" = minimizing the chosen cost function

E.g., **mean squared error** (MSE)

Model parameters (weights and biases)

Expected output

Actual output (neuron activations)

$$f_C(w) = \frac{1}{N_{\text{train}}} \sum_x ||y(x) - a(x)||^2$$

Number of training data points (vectors) x

Learning algorithm

Batch gradient descent

$$\Delta w = -\eta \nabla f_C$$

Learning rate

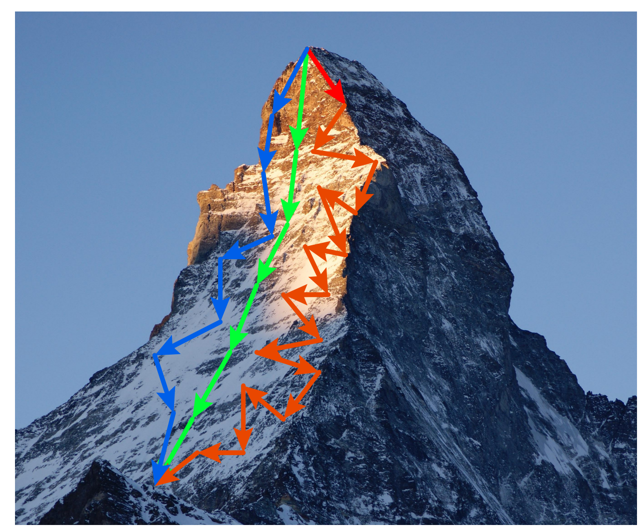
Gradient estimated from the **whole training data (batch)**

Stochastic gradient descent (typically better)

Gradient estimated from **one data point**

Mini-batch gradient descent (typically even better)

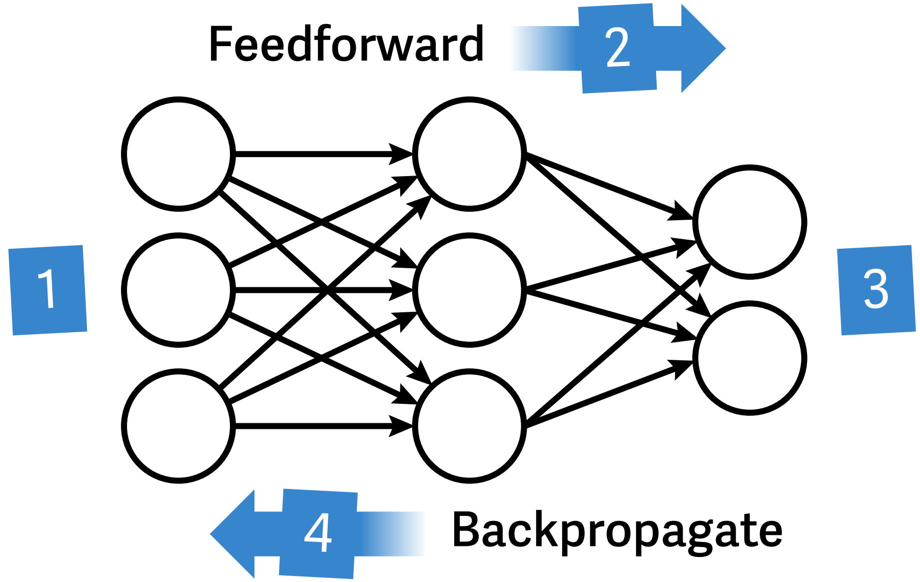
Gradient estimated from **subsets of data points (mini-batches)**



Total recall

Backpropagation

Or how to **compute the gradient of the cost function efficiently**



Main steps

1 Compute the input activations: $a^{(1)} = f_A(x)$

2 **Feedforward:** Compute $z^{(l)} = w^{(l)} a^{(l-1)} + b^{(l)}$ and $a^{(l)} = f_A(z^{(l)})$ for successive layers $l = 2, 3, \ldots, L$

3 Compute the output error: $\delta^{(L)} = \nabla_a f_C \odot f'_A(z^{(L)})$

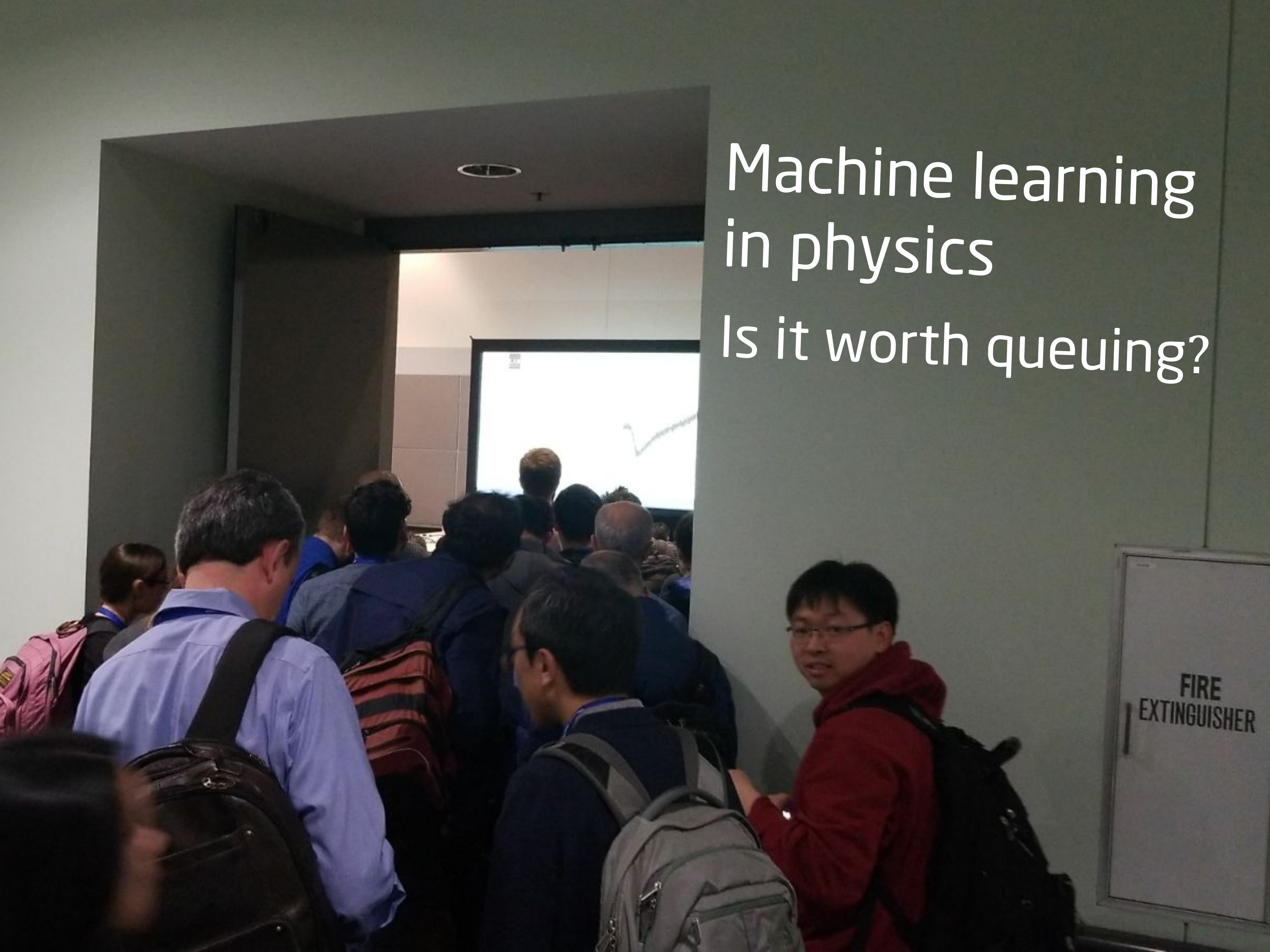
Comes from the usual chain rule

4 **Backpropagate the error:** Compute $\delta^{(l)} = [(w^{(l+1)})^T \delta^{(l+1)}] \odot f'_A(z^{(l)})$ for successive layers $l = L-1, L-2, \ldots, 2$

Output

$$\frac{\partial f_C}{\partial w_{jk}^{(l)}} = a_k^{(l-1)} \delta_j^{(l)} \qquad \frac{\partial f_C}{\partial b_j^{(l)}} = \delta_j^{(l)}$$

Gradient computed from only two passes (forward and backward)


Machine learning in physics
Is it worth queuing?
FIRE EXTINGUISHER

Science applications

Machine learning mostly comes from science!

What goes around comes back around

Already applied in

Biology

In neuroscience, evolution, immunology, genetics, etc.

Libbrecht and Noble (2015)

Medicine

In epidemiology, disease development, etc.

Cleophas and Zwinderman (2015)

Chemistry

In optimization of reactions, search for new molecules, etc.

Cartwright (2007)

Physics

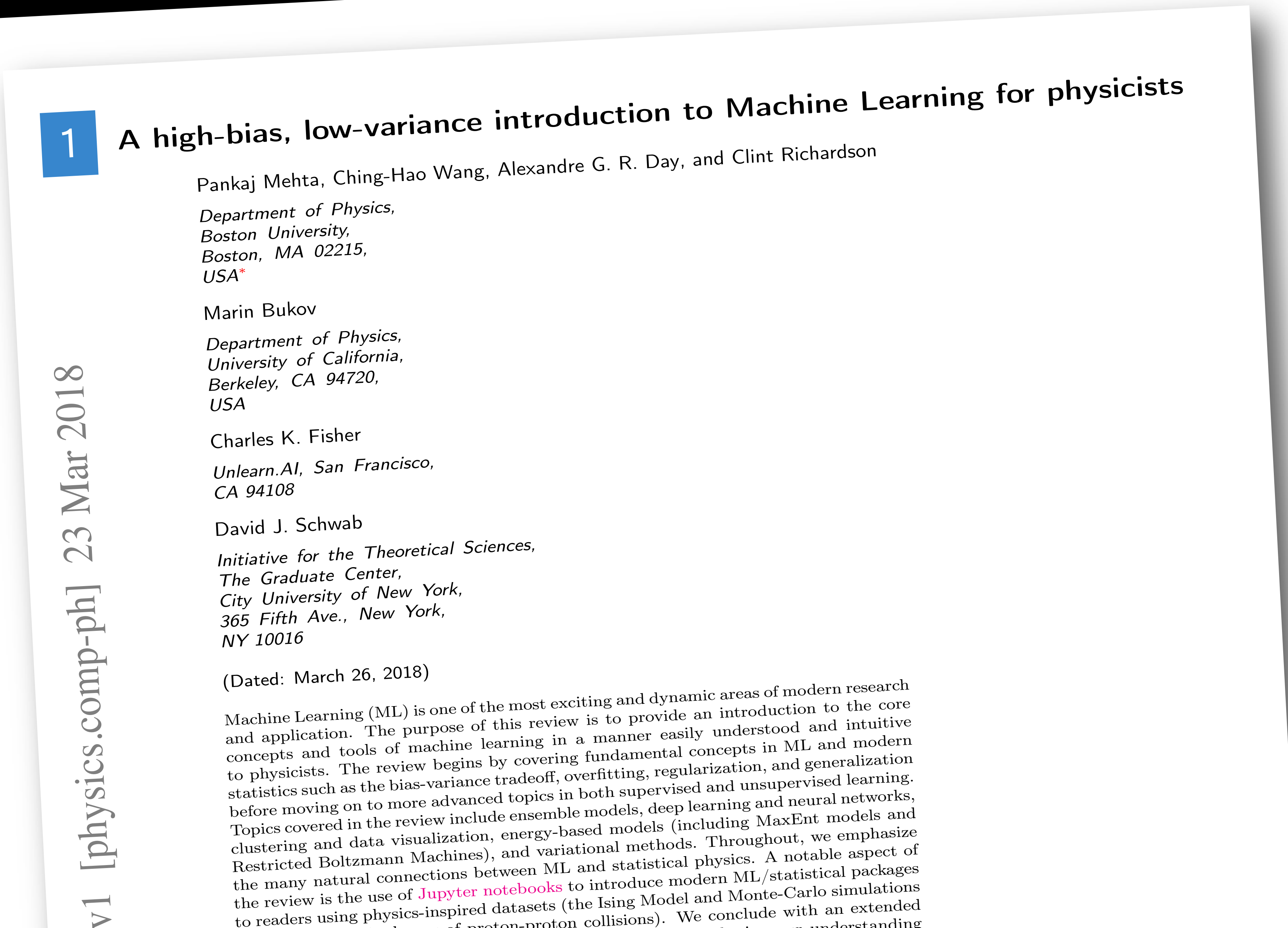
In high-energy physics, astronomy, etc.

Castelvecchi (2015)

And more recently

In condensed matter physics and general quantum physics

Already (at least) two reviews


1

A high-bias, low-variance introduction to Machine Learning for physicists

Pankaj Mehta, Ching-Hao Wang, Alexandre G. R. Day, and Clint Richardson

Department of Physics,
Boston University,
Boston, MA 02215,
*USA**

Marin Bukov

Department of Physics,
University of California,
Berkeley, CA 94720,
USA

Charles K. Fisher

Unlearn.AI, San Francisco,
CA 94108

David J. Schwab

Initiative for the Theoretical Sciences,
The Graduate Center,
City University of New York,
365 Fifth Ave., New York,
NY 10016

(Dated: March 26, 2018)

Machine Learning (ML) is one of the most exciting and dynamic areas of modern research and application. The purpose of this review is to provide an introduction to the core concepts and tools of machine learning in a manner easily understood and intuitive to physicists. The review begins by covering fundamental concepts in ML and modern statistics such as the bias-variance tradeoff, overfitting, regularization, and generalization before moving on to more advanced topics in both supervised and unsupervised learning. Topics covered in the review include ensemble models, deep learning and neural networks, clustering and data visualization, energy-based models (including MaxEnt models and Restricted Boltzmann Machines), and variational methods. Throughout, we emphasize the many natural connections between ML and statistical physics. A notable aspect of the review is the use of Jupyter notebooks to introduce modern ML/statistical packages to readers using physics-inspired datasets (the Ising Model and Monte-Carlo simulations of proton-proton collisions). We conclude with an extended

Already (at least) two reviews

1

A high-bias, low-variance introduction to Machine Learning for physicists

Pankaj Mehta, Ching-Hao Wang, Alexandre G. R. Day, and Clint Richardson

Department of Physics,
Boston University,

comp-ph] 23 Mar 2018

2

Machine learning & artificial intelligence in the quantum domain

Vedran Dunjko

Institute for Theoretical Physics, University of Innsbruck, Innsbruck 6020, Austria
Max Planck Institute of Quantum Optics, Garching 85748, Germany
Email: vedran.dunjko@mpq.mpg.de

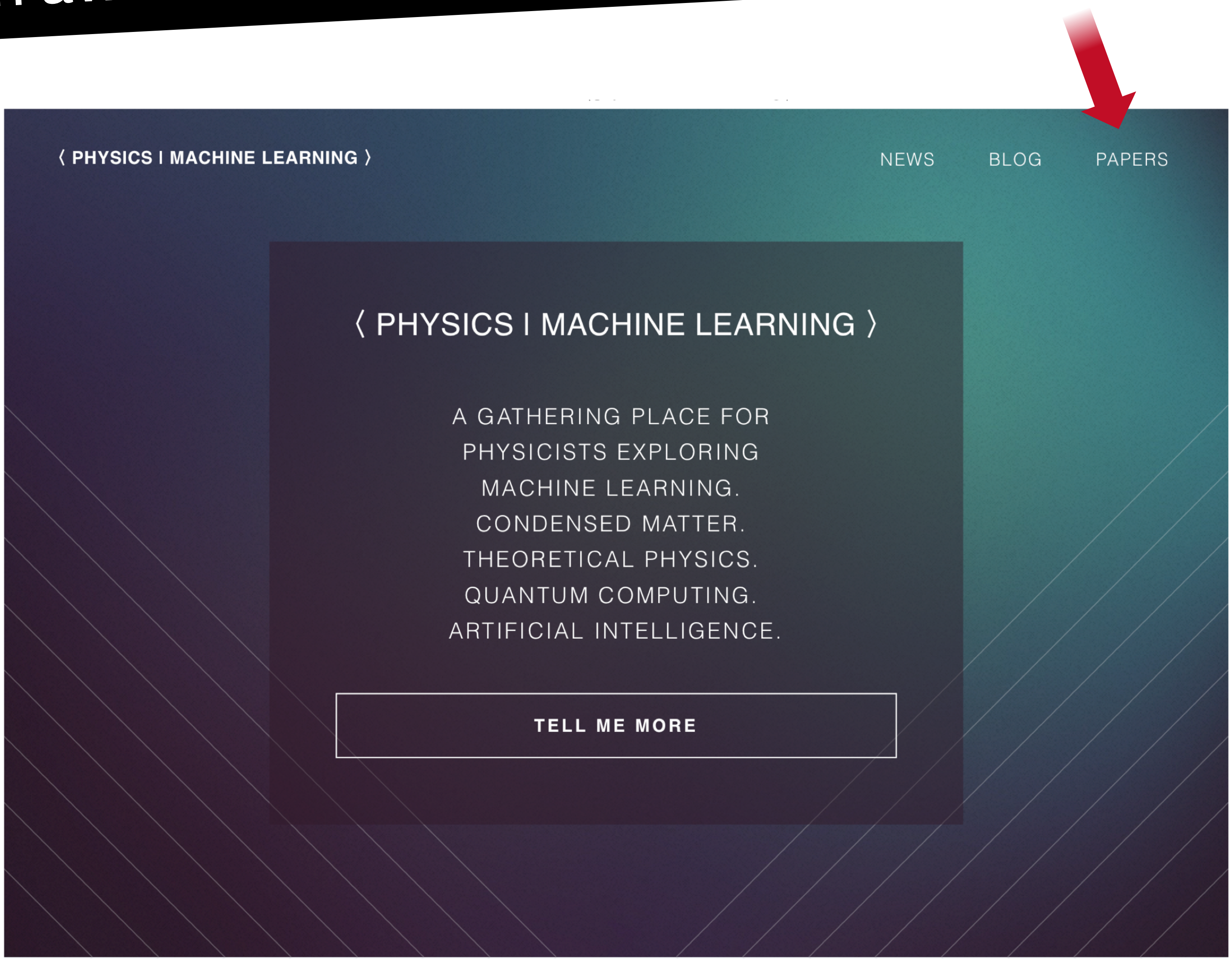
Hans J. Briegel

Institute for Theoretical Physics, University of Innsbruck Innsbruck 6020, Austria
Department of Philosophy, University of Konstanz, Konstanz 78457, Germany
Email: hans.briegel@uibk.ac.at

[quant-ph] 8 Sep 2017

Abstract. Quantum information technologies, on the one side, and intelligent learning systems, on the other, are both emergent technologies that will likely have a transforming impact on our society in the future. The respective underlying fields of basic research – quantum information (QI) versus machine learning and artificial intelligence (AI) – have their own specific questions and challenges, which have hitherto been investigated largely independently. However, in a growing body of recent work, researchers have been probing the question to what extent these fields can indeed learn and benefit from each other. QML explores the interaction between quantum computing and machine learning, investigating how results and techniques from one field can be used to solve the problems of the other. In recent time, we have witnessed significant breakthroughs in both directions of influence. For instance, quantum computing is finding a vital application in providing speed-ups for machine learning problems, critical in our "big data" world. Conversely, machine learning already permeates many cutting-edge technologies, and may become instrumental in advanced quantum technologies. Aside from quantum speed-up in data analysis, or classical machine learning optimization used in quantum experiments, quan-

And an awesome blog tracking new papers



https://physicsml.github.io

Condensed matter physics applications

1 **Identifying phase transitions**

With local order parameters

Carrasquilla, Melko,
Nature Physics 13, 431 (2017)

Broecker, Carrasquilla, Melko, Trebst
Sci. Rep. 7, 8823 (2017)

Etc.

Without local order parameters (topological)

Zhang , Kim,
Phys. Rev. Lett. 118, 216401 (2017)

Zhang, Melko, Kim,
Phys. Rev. B 96, 245119 (2017)

Etc.

Many-body localized

Schindler, Regnault, Neupert,
Phys. Rev. B 95, 245134 (2017)

Etc.

More general methods

Van Nieuwenburg, Liu, Huber,
Nature Physics 13, 435 (2017)

Broecker, Assaad, Trebst,
arXiv:1707.00663 (2017)

Van Nieuwenburg, Bairey, Refael,
arXiv:1712.00450 (2018)

Condensed matter physics applications

1 **Identifying phase transitions**

With local order parameters

Carrasquilla, Melko,
Nature Physics 13, 431 (2017)

Broecker, Carrasquilla, Melko, Trebst
Sci. Rep. 7, 8823 (2017)

Etc.

Without local order parameters (topological)

Zhang , Kim,
Phys. Rev. Lett. 118, 216401 (2017)

Zhang, Melko, Kim,
Phys. Rev. B 96, 245119 (2017)

Etc.

Many-body localized

Schindler, Regnault, Neupert,
Phys. Rev. B 95, 245134 (2017)

Etc.

More general methods

Van Nieuwenburg, Liu, Huber,
Nature Physics 13, 435 (2017)

Broecker, Assaad, Trebst,
arXiv:1707.00663 (2017)

Van Nieuwenburg, Bairey, Refael,
arXiv:1712.00450 (2018)

Condensed matter physics applications

nature physics

LETTERS

PUBLISHED ONLINE: 13 FEBRUARY 2017 | DOI: 10.1038/NPHYS4035

Machine learning phases of matter

Juan Carrasquilla[1*] and Roger G. Melko[1,2]

Condensed-matter physics is the study of the collective behaviour of infinitely complex assemblies of electrons, nuclei, magnetic moments, atoms or qubits[1]. This complexity is reflected in the size of the state space, which grows exponentially with the number of particles, reminiscent of the 'curse of dimensionality' commonly encountered in machine learning[2]. Despite this curse, the machine learning community has developed techniques with remarkable abilities to recognize, classify, and characterize complex sets of data. Here, we show that modern machine learning architectures, such as fully connected and convolutional neural networks[3], can identify phases and phase transitions in a variety of condensed-matter Hamiltonians. Readily programmable through modern software libraries[4,5], neural networks can be trained to detect multiple types of order parameter, as well as highly non-trivial states with no conventional order, directly from raw state configurations sampled with Monte Carlo[6,7].

Conventionally, the study of phases in condensed-matter systems is performed with the help of tools that have been carefully designed to elucidate the underlying physical structures of various states. Among the most powerful are Monte Carlo simulations, which consist of two steps: a stochastic importance sampling over state

is composed of an input layer with values determined by the spin configurations, a 100-unit hidden layer of sigmoid neurons, and an analogous output layer. When trained on a broad range of data at temperatures above and below T_c, the neural network is able to correctly classify data in a test set. Finite-size scaling is capable of systematically narrowing in on the thermodynamic value of T_c in a way analogous to measurements of the magnetization: a data collapse of the output layer (Fig. 1b) leads to an estimate of the critical exponent $\nu \simeq 1.0 \pm 0.2$, while a size scaling of the crossing temperature T^*/J estimates $T_c/J \simeq 2.266 \pm 0.002$ (Fig. 1c). One can understand the training of the network through a simple toy model involving a hidden layer of only three analytically 'trained' perceptrons, representing the possible combinations of high- and low-temperature magnetic states exclusively on the basis of their magnetization. Similarly, our 100-unit neural network relies on the magnetization of the configurations in the classification task. Details about the toy model, the 100-unit neural network, as well as a low-dimensional visualization of the training data, which may be used as a preprocessing step to generate the labels if they are not available a priori, are discussed in the Supplementary Figs 1, 2, and 4. We note that in a recent development, a closely related neural-network-based approach allows for the determination of critical points using

Condensed matter physics applications

www.nature.com/scientificreports

SCIENTIFIC REPORTS

OPEN

Machine learning quantum phases of matter beyond the fermion sign problem

Received: 15 May 2017
Accepted: 21 July 2017
Published online: 18 August 2017

Peter Broecker[1], Juan Carrasquilla[2], Roger G. Melko[2,3] & Simon Trebst[1]

State-of-the-art machine learning techniques promise to become a powerful tool in statistical mechanics via their capacity to distinguish different phases of matter in an automated way. Here we demonstrate that convolutional neural networks (CNN) can be optimized for quantum many-fermion systems such that they correctly identify and locate quantum phase transitions in such systems. Using auxiliary-field quantum Monte Carlo (QMC) simulations to sample the many-fermion system, we show that the Green's function holds sufficient information to allow for the distinction of different fermionic phases via a CNN. We demonstrate that this QMC + machine learning approach works even for systems exhibiting a severe fermion sign problem where conventional approaches to extract information from the Green's function, e.g. in the form of equal-time correlation functions, fail.

In quantum statistical physics, the sign problem refers to the generic inability of quantum Monte Carlo (QMC) approaches to tackle fermionic systems with the same unparalleled efficiency it exhibits for unfrustrated bosonic

Condensed matter physics applications

1

Identifying phase transitions

With local order parameters

Carrasquilla, Melko,
Nature Physics 13, 431 (2017)

Broecker, Carrasquilla, Melko, Trebst
Sci. Rep. 7, 8823 (2017)

Etc.

Without local order parameters (topological)

Zhang , Kim,
Phys. Rev. Lett. 118, 216401 (2017)

Zhang, Melko, Kim,
Phys. Rev. B 96, 245119 (2017)

Etc.

Many-body localized

Schindler, Regnault, Neupert,
Phys. Rev. B 95, 245134 (2017)

Etc.

More general methods

Van Nieuwenburg, Liu, Huber,
Nature Physics 13, 435 (2017)

Broecker, Assaad, Trebst,
arXiv:1707.00663 (2017)

Van Nieuwenburg, Bairey, Refael,
arXiv:1712.00450 (2018)

Condensed matter physics applications

Selected for a Viewpoint in *Physics*

PRL **118,** 216401 (2017) PHYSICAL REVIEW LETTERS week ending 26 MAY 2017

Quantum Loop Topography for Machine Learning

Yi Zhang* and Eun-Ah Kim†

Department of Physics, Cornell University, Ithaca, New York 14853, USA
and Kavli Institute for Theoretical Physics, University of California, Santa Barbara, California 93106, USA

(Received 15 November 2016; revised manuscript received 13 February 2017; published 22 May 2017)

Despite rapidly growing interest in harnessing machine learning in the study of quantum many-body systems, training neural networks to identify quantum phases is a nontrivial challenge. The key challenge is in efficiently extracting essential information from the many-body Hamiltonian or wave function and turning the information into an image that can be fed into a neural network. When targeting topological phases, this task becomes particularly challenging as topological phases are defined in terms of nonlocal properties. Here, we introduce quantum loop topography (QLT): a procedure of constructing a multidimensional image from the "sample" Hamiltonian or wave function by evaluating two-point operators that form loops at independent Monte Carlo steps. The loop configuration is guided by the characteristic response for defining the phase, which is Hall conductivity for the cases at hand. Feeding QLT to a fully connected neural network with a single hidden layer, we demonstrate that the architecture can be effectively trained to distinguish the Chern insulator and the fractional Chern insulator from trivial insulators with high fidelity. In addition to establishing the first case of obtaining a phase diagram with a topological quantum phase transition with machine learning, the perspective of bridging traditional condensed matter theory with machine learning will be broadly valuable.

DOI: 10.1103/PhysRevLett.118.216401

Introduction.—Machine learning techniques have been enabling neural networks to recognize and interpret big data sets of images and speeches [1]. Through supervised

presence of translational symmetry, targeting a single topological phase at a time [7,10]. Another approach was to detect the topological edge states [13]. In addition,

Condensed matter physics applications

PHYSICAL REVIEW B **96**, 245119 (2017)

Machine learning $\mathbb{Z}_2$ quantum spin liquids with quasiparticle statistics

Yi Zhang,[1,*] Roger G. Melko,[2,3] and Eun-Ah Kim[1,†]
[1]*Department of Physics, Cornell University, Ithaca, New York 14853, USA*
[2]*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*
[3]*Department of Physics and Astronomy, University of Waterloo, Ontario, N2L 3G1, Canada*
(Received 20 May 2017; revised manuscript received 30 October 2017; published 13 December 2017; corrected 12 February 2018)

After decades of progress and effort, obtaining a phase diagram for a strongly correlated topological system still remains a challenge. Although in principle one could turn to Wilson loops and long-range entanglement, evaluating these nonlocal observables at many points in phase space can be prohibitively costly. With growing excitement over topological quantum computation comes the need for an efficient approach for obtaining topological phase diagrams. Here we turn to machine learning using quantum loop topography (QLT), a notion we have recently introduced. Specifically, we propose a construction of QLT that is sensitive to quasiparticle statistics. We then use mutual statistics between the spinons and visons to detect a $\mathbb{Z}_2$ quantum spin liquid in a multiparameter phase space. We successfully obtain the quantum phase boundary between the topological and trivial phases using a simple feed-forward neural network. Furthermore, we demonstrate advantages of our approach for the evaluation of phase diagrams relating to speed and storage. Such statistics-based machine learning of topological phases opens new efficient routes to studying topological phase diagrams in strongly correlated systems.

DOI: 10.1103/PhysRevB.96.245119

I. INTRODUCTION

Despite much interest in topological phases of matter, the search for and detection of the finite regions of phase space that support topological order has been a longstanding challenge. This is a nontrivial challenge because microscopic models of

specific heat is an effective indicator of a phase transition, it has the drawback that it does not reveal any information regarding the topological aspects of the associated phases. Hence, in addition to these standard techniques, developing a cost-effective approach that can map out a phase diagram with topological

Condensed matter physics applications

1 **Identifying phase transitions**

With local order parameters

Carrasquilla, Melko,
Nature Physics 13, 431 (2017)

Broecker, Carrasquilla, Melko, Trebst
Sci. Rep. 7, 8823 (2017)

Etc.

Without local order parameters (topological)

Zhang , Kim,
Phys. Rev. Lett. 118, 216401 (2017)

Zhang, Melko, Kim,
Phys. Rev. B 96, 245119 (2017)

Etc.

Many-body localized

Schindler, Regnault, Neupert,
Phys. Rev. B 95, 245134 (2017)

Etc.

More general methods

Van Nieuwenburg, Liu, Huber,
Nature Physics 13, 435 (2017)

Broecker, Assaad, Trebst,
arXiv:1707.00663 (2017)

Van Nieuwenburg, Bairey, Refael,
arXiv:1712.00450 (2018)

Condensed matter physics applications

PHYSICAL REVIEW B **95**, 245134 (2017)

Probing many-body localization with neural networks

Frank Schindler,[1] Nicolas Regnault,[2] and Titus Neupert[1]

[1]*Department of Physics, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland*

[2]*Laboratoire Pierre Aigrain, Département de physique de l'ENS, Ecole normale supérieure, PSL Research University, Université Paris Diderot, Sorbonne Paris Cité, Sorbonne Universités, UPMC Univ. Paris 06, CNRS, 75005 Paris, France*

(Received 11 April 2017; published 26 June 2017)

We show that a simple artificial neural network trained on entanglement spectra of individual states of a many-body quantum system can be used to determine the transition between a many-body localized and a thermalizing regime. Specifically, we study the Heisenberg spin-1/2 chain in a random external field. We employ a multilayer perceptron with a single hidden layer, which is trained on labeled entanglement spectra pertaining to the fully localized and fully thermal regimes. We then apply this network to classify spectra belonging to states in the transition region. For training, we use a cost function that contains, in addition to the usual error and regularization parts, a term that favors a confident classification of the transition region states. The resulting phase diagram is in good agreement with the one obtained by more conventional methods and can be computed for small systems. In particular, the neural network outperforms conventional methods in classifying individual eigenstates pertaining to a single disorder realization. It allows us to map out the structure of these eigenstates across the transition with spatial resolution. Furthermore, we analyze the network operation using the dreaming technique to show that the neural network correctly learns by itself the power-law structure of the entanglement spectra in the many-body localized regime.

DOI: 10.1103/PhysRevB.95.245134

I. INTRODUCTION

Artificial neural networks are routinely employed for data classification. They are useful when features distinguishing one class of data from another are unknown or unwieldy. A neural network can learn such features from examples, i.e.,

robust quantum memories [29]. Here, we study the Heisenberg chain in a random field as a simple model for MBL. At strong disorder, the model is in the MBL regime, whereas it satisfies the ETH if disorder is weak. Several measures or quantities allow a well-controlled quantitative distinction of

Condensed matter physics applications

1 **Identifying phase transitions**

With local order parameters

Carrasquilla, Melko,
Nature Physics 13, 431 (2017)

Broecker, Carrasquilla, Melko, Trebst
Sci. Rep. 7, 8823 (2017)

Etc.

Without local order parameters (topological)

Zhang , Kim,
Phys. Rev. Lett. 118, 216401 (2017)

Zhang, Melko, Kim,
Phys. Rev. B 96, 245119 (2017)

Etc.

Many-body localized

Schindler, Regnault, Neupert,
Phys. Rev. B 95, 245134 (2017)

Etc.

More general methods

Van Nieuwenburg, Liu, Huber,
Nature Physics 13, 435 (2017)

Broecker, Assaad, Trebst,
arXiv:1707.00663 (2017)

Van Nieuwenburg, Bairey, Refael,
arXiv:1712.00450 (2018)

Condensed matter physics applications

nature physics

LETTERS

PUBLISHED ONLINE: 13 FEBRUARY 2017 | DOI: 10.1038/NPHYS4037

Learning phase transitions by confusion

Evert P. L. van Nieuwenburg*, Ye-Hua Liu and Sebastian D. Huber

Classifying phases of matter is key to our understanding of many problems in physics. For quantum-mechanical systems in particular, the task can be daunting due to the exponentially large Hilbert space. With modern computing power and access to ever-larger data sets, classification problems are now routinely solved using machine-learning techniques[1]. Here, we propose a neural-network approach to finding phase transitions, based on the performance of a neural network after it is trained with data that are deliberately labelled incorrectly. We demonstrate the success of this method on the topological phase transition in the Kitaev chain[2], the thermal phase transition in the classical Ising model[3], and the many-body-localization transition in a disordered quantum spin chain[4]. Our method does not depend on order parameters, knowledge of the topological content of the phases, or any other specifics of the transition at hand. It therefore paves the way to the development of a generic tool for identifying unexplored phase transitions.

Machine learning as a tool for analysing data is becoming more and more prevalent in an increasing number of fields. This is due to a combination of availability of large amounts of data and the of the machine learner. We will base our method on NNs, which are capable of fitting arbitrary nonlinear functions[11]. Indeed, if a linear feature extraction method worked, there would have been no need to explicitly find labels in the first place.

We emphasize the main result in this work is that with the proposed method we are able to find a consistent labelling for data that have distinct patterns. A change in the pattern of some observable is not necessarily correlated with a physical phase transition. Our method is capable of recognizing the change of pattern, after which it is up to the user to investigate whether the change corresponds to a crossover or a phase transition. We remark that we do not exclude the possibility that linear methods would be able to perform some of the tasks we describe below. Nor do we exclude the possibility that other methods such as latent-variable models or other maximum likelihood algorithms would be able to perform the same task. Finding the correct method or transformation of the data may be a prohibitive task however, and so using a (possibly overpowered) method such as NNs provides a useful starting point. Our method boils down to bootstrapping a supervised learning method to an unsupervised one, at the expense of computational time.

Additionally, but not less important, we propose the use of the

Condensed matter physics applications

Quantum phase recognition via unsupervised machine learning

Peter Broecker,[1] Fakher F. Assaad,[2] and Simon Trebst[1]
[1]*Institute for Theoretical Physics, University of Cologne, 50937 Cologne, Germany*
[2]*Institut für Theoretische Physik und Astrophysik, Universität Würzburg, 97074 Würzburg, Germany*
(Dated: July 4, 2017)

The application of state-of-the-art machine learning techniques to statistical physic problems has seen a surge of interest for their ability to discriminate phases of matter by extracting essential features in the many-body wavefunction or the ensemble of correlators sampled in Monte Carlo simulations. Here we introduce a generalization of supervised machine learning approaches that allows to accurately map out phase diagrams of interacting many-body systems without any prior knowledge, e.g. of their general topology or the number of distinct phases. To substantiate the versatility of this approach, which combines convolutional neural networks with quantum Monte Carlo sampling, we map out the phase diagrams of interacting boson and fermion models both at zero and finite temperatures and show that first-order, second-order, and Kosterlitz-Thouless phase transitions can all be identified. We explicitly demonstrate that our approach is capable of identifying the phase transition to non-trivial many-body phases such as superfluids or topologically ordered phases without supervision.

cond-mat.str-el] 3 Jul 2017

In statistical physics, a continuous stream of computational and conceptual advances has been directed towards attacking the quantum many-body problem – the identification of the ground state of a macroscopic number of interacting bosons, spins or fermions. Pivotal steps forward have included the development of numerical many-body techniques such as quantum Monte Carlo simulations [1] and the density matrix renormalization group [2, 3] along with conceptual advances such as the formulation of an entanglement perspective [4, 5] on the quantum many-body problem arising from the interplay of quantum information theory and quantum statistical physics. Currently, machine learning (ML) approaches are entering this field as new players. Their core functions, dimensional reduction and feature extraction, are a perfect match to the goal of identifying essential characteristics of a quantum many-

any prior knowledge, e.g. regarding the overall topology or number of distinct phases present in a phase diagram. The essential ingredient of our approach are convolutional neural networks (CNN) [15] that combine a preprocessing step using convolutional filters with a conventional neural network (typically involving multiple layers itself). In previous work [10–14] such CNNs have been used in a *supervised* learning setting where a (quantum) many-body Hamiltonian is considered that, as a function of some parameter λ, exhibits a phase transition between two phases – such as the thermal phase transition in the classical Ising model [11] or the zero-temperature quantum phase transition as a function of some coupling parameter [10]. In such a setting where one has prior knowledge about the existence of two distinct phases in some parameter range, one can train the CNN with *labeled* configurations or

Condensed matter physics applications

Learning phase transitions from dynamics

Evert van Nieuwenburg,[1,*] Eyal Bairey,[2,*] and Gil Refael[1]
[1] Institute for Quantum Information and Matter, Caltech, Pasadena, California 91125, USA
[2] Physics Department, Technion, 3200003, Haifa, Israel

We propose the use of recurrent neural networks for classifying phases of matter based on the dynamics of experimentally accessible observables. We demonstrate this approach by training recurrent networks on the magnetization traces of two distinct models of one-dimensional disordered and interacting spin chains. The obtained phase diagram for a well-studied model of the many-body localization transition shows excellent agreement with previously known results obtained from time-independent entanglement spectra. For a periodically-driven model featuring an inherently dynamical time-crystalline phase, the phase diagram that our network traces in a previously-unexplored regime coincides with an order parameter for its expected phases.

Introduction - Machine learning is emerging as a novel tool for identifying phases of matter [1–15]. At its core, this problem can be cast as a classification problem in which data obtained from physical systems are assigned a class (i.e. a phase) using machine learning methods. This approach has enabled autonomous detection of order parameters [2, 5, 6], phase transitions [1, 3] and entire phase diagrams [4, 7, 16, 17]. Simultaneous reserach effort at the interface between machine learning and many-body physics has focussed on the use of neural networks for efficient representations of quantum wavefunctions [18–26], drawing a parallel between deep networks and the renormalization group [27–29]. Overall, these studies exemplify the power of machine learning for extracting information from physical data without detailed physical input. In particular, it shows potential for identifying novel phases through automatic processing of large-scale data; possibly identifying features that may have been

of the same model [11], as well as on a slightly different model featuring two distinct MBL phases [17]. Here, we insist on using only experimentally relevant (i.e. measureable) quantities such as the magnetization of individual spins. We find that the network succeeds at distinguishing between the ergodic and localized phases of this model, recovering phase boundaries similar to those obtained by previous methods.

We then apply our method to a periodically driven model, featuring among its three phases one which is unique to the time-dependent setting, namely a time crystal [44–50]. Indeed the method distinguishes between the time-crystalline, Floquet-ergodic and Floquet-MBL [51–53] phases of this model.

In the following section, we first introduce the essentials of recurrent neural networks. We refer the reader to Ref. [54] for an extensive introduction to the non-

Condensed matter physics applications

Learning phase transitions from dynamics

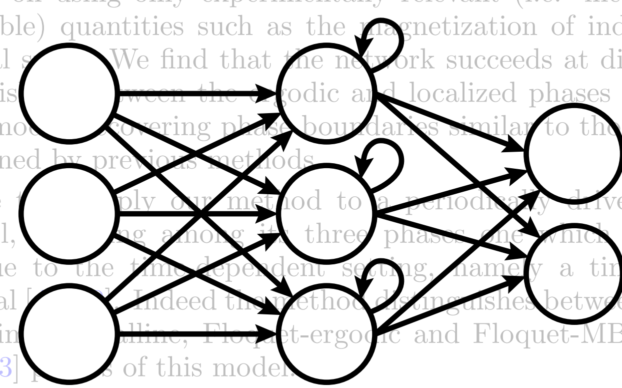
Evert van Nieuwenburg,[1,*] Eyal Bairey,[2,*] and Gil Refael[1]
[1] *Institute for Quantum Information and Matter, Caltech, Pasadena, California 91125, USA*
[2] *Physics Department, Technion, 3200003, Haifa, Israel*

We propose the use of recurrent neural networks for classifying phases of matter based on the dynamics of experimentally accessible observables. We demonstrate this approach by training recurrent networks on the magnetization traces of two distinct models of one-dimensional disordered and interacting spin chains. The obtained phase diagram for a well-studied model of the many-body localization transition shows excellent agreement with previously known results obtained from time-independent entanglement spectra. For a periodically-driven model featuring an inherently dynamical time-crystalline phase, the phase diagram that our network traces in a previously-unexplored regime coincides with an order parameter for its expected phases.

Introduction - Machine learning is emerging as a novel tool for identifying phases of matter [1–15]. At its core, this problem can be cast as a classification problem in which data obtained from physical systems are assigned a class (i.e. a phase) using machine learning methods. This approach has enabled autonomous detection of order parameters [2, 5, 6], phase transitions [1, 3] and entire phase diagrams [4, 7, 16, 17]. Simultaneous reserach effort at the interface between machine learning and many-body physics has focussed on the use of neural networks for efficient representations of quantum wavefunctions [18–26], drawing a parallel between deep networks and the renormalization group [27–29]. Overall, these studies exemplify the power of machine learning for extracting information from physical data without detailed physical input. In particular, it shows potential for identifying novel phases through automatic processing of large-scale data; possibly identifying features that may have been

of the same model [11], as well as on a slightly different

Recurrent neural network (RNN)



In the following section, we first introduce the essentials of recurrent neural networks. We refer the reader to Ref. [54] for an extensive introduction to the non-

Condensed matter physics applications

2

Learning a representation of a many-body wavefunction (e.g., ground states)

With Restricted Boltzmann machines (RMBs)

Carleo, Troyer,
Science 355, 602 (2017)

With Deep Boltzmann machines (DBMs)

Gao, Duan,
Nature Comm. 8, 662 (2017)

G. Carleo, Y. Nomura, and M. Imada,
arXiv:1802.09558 (2018)

Many more to come!

Condensed matter physics applications

2

Learning a representation of a many-body wavefunction (e.g., ground states)

With Restricted Boltzmann machines (RMBs)

Carleo, Troyer,
Science 355, 602 (2017)

With Deep Boltzmann machines (DBMs)

Gao, Duan,
Nature Comm. 8, 662 (2017)

G. Carleo, Y. Nomura, and M. Imada,
arXiv:1802.09558 (2018)

Many more to come!

Condensed matter physics applications

Science 355, 602–606 (2017)

RESEARCH ARTICLE

MANY-BODY PHYSICS

Solving the quantum many-body problem with artificial neural networks

Giuseppe Carleo[1*] and Matthias Troyer[1,2]

The challenge posed by the many-body problem in quantum physics originates from the difficulty of describing the nontrivial correlations encoded in the exponential complexity of the many-body wave function. Here we demonstrate that systematic machine learning of the wave function can reduce this complexity to a tractable computational form for some notable cases of physical interest. We introduce a variational representation of quantum states based on artificial neural networks with a variable number of hidden neurons. A reinforcement-learning scheme we demonstrate is capable of both finding the ground state and describing the unitary time evolution of complex interacting quantum systems. Our approach achieves high accuracy in describing prototypical interacting spins models in one and two dimensions.

The wave function Ψ is a fundamental object in quantum physics and possibly the hardest to grasp in the classical world. Ψ is a monolithic mathematical quantity that contains all of the information on a quantum state, be it a single particle or a complex a large number of unexplored regimes exist, including many open problems. These encompass fundamental questions ranging from the dynamical properties of high-dimensional systems (*11*, *12*) to the exact ground-state properties of strongly interacting fermions (*13*, *14*). At the heart techniques to attack these problems, artificial neural networks play a prominent role (*16*). They can perform exceedingly well in a variety of contexts ranging from image and speech recognition (*17*) to game playing (*18*). Very recently, applications of neural networks to the study of physical phenomena have been introduced (*19*–*23*). These have so far focused on the classification of complex phases of matter, when exact sampling of configurations from these phases is possible. The challenging goal of solving a many-body problem without prior knowledge of exact samples is nonetheless still unexplored, and the potential benefits of artificial intelligences in this task are, at present, substantially unknown. Therefore, it is of fundamental and practical interest to understand whether an artificial neural network can modify and adapt itself to describe and analyze such a quantum system. This ability could then be used to solve the quantum many-body problem in regimes that have traditionally been inaccessible to existing exact numerical approaches.

Here we introduce a representation of the wave function in terms of artificial neural networks specified by a set of internal parameters $\mathcal{W}$. We present a stochastic framework for reinforcement learning of the parameters $\mathcal{W}$, allowing for the best possible representation of both ground-state and time-dependent physical states of a given quantum Hamiltonian $\mathcal{H}$. The parameters of the neural network are then optimized (trained, in the language of neural networks),

Condensed matter physics applications

Science 355, 602–606 (2017)

RESEARCH ARTICLE

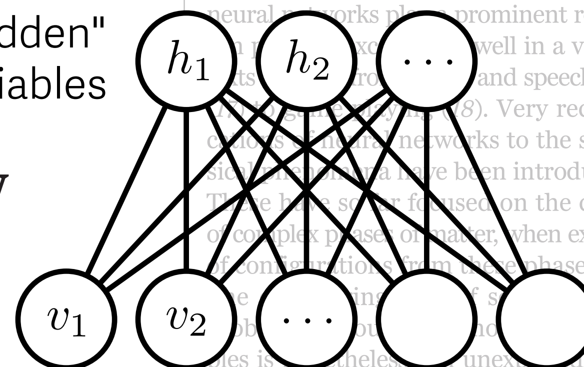
MANY-BODY PHYSICS

Solving the quantum many-body problem with artificial neural networks

Giuseppe Carleo[1*] and Matthias Troyer[1,2]

The challenge posed by the many-body problem in quantum physics originates from the difficulty of describing the nontrivial correlations encoded in the exponential complexity of the many-body wave function. Here we demonstrate that systematic machine learning of the wave function can reduce this complexity to a tractable computational form for some notable cases of physical interest. We introduce a variational representation of quantum states based on artificial neural networks with a variable number of hidden neurons. A reinforcement-learning scheme we demonstrate is capable of both finding the ground state and describing the unitary time evolution of complex interacting quantum systems. Our approach achieves high accuracy in describing prototypical interacting spins models in one and two dimensions.

Restricted Boltzmann Machine (RBM)



Encodes $P_\Lambda(\mathcal{H}|\mathcal{V}) = \frac{1}{Z_\Lambda} e^{-E_\Lambda(\mathcal{V},\mathcal{H})}$

$$E_\Lambda(\mathcal{V},\mathcal{H}) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_{i,j} v_i \lambda_{ij} h_j$$

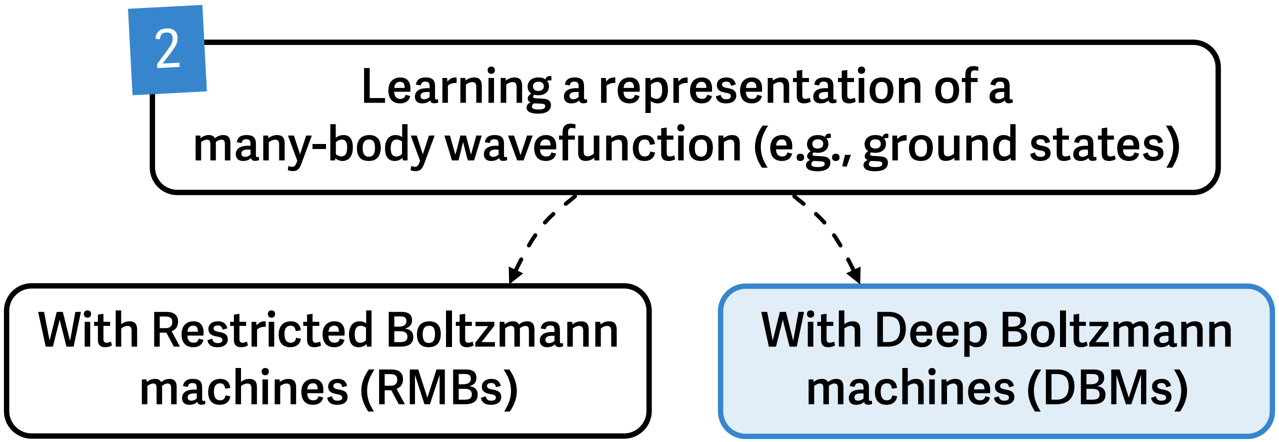
Weights (or "interactions" between hidden and visible variables)

The wave function Ψ is a fundamental object in quantum physics and possibly the hardest to grasp in the classical world. Ψ is a monolithic mathematical quantity that contains all of the information on a quantum state, be it a single particle or a complex

a large number of unexplored regimes exist, including many open problems. These encompass fundamental questions ranging from the dynamical properties of high-dimensional systems (11, 12) to the exact ground-state properties of strongly interacting fermions (13, 14). At the heart

specified by a set of internal parameters $\mathcal{W}$. We present a stochastic framework for reinforcement learning of the parameters $\mathcal{W}$, allowing for the best possible representation of both ground-state and time-dependent physical states of a given quantum Hamiltonian $\mathcal{H}$. The parameters of the neural network are then optimized (trained, in the language of neural networks),

Condensed matter physics applications



Carleo, Troyer,
Science 355, 602 (2017)

Gao, Duan,
Nature Comm. 8, 662 (2017)

G. Carleo, Y. Nomura, and M. Imada,
arXiv:1802.09558 (2018)

Many more to come!

Condensed matter physics applications

nature COMMUNICATIONS

ARTICLE

DOI: 10.1038/s41467-017-00705-2 OPEN

Efficient representation of quantum many-body states with deep neural networks

Xun Gao[1] & Lu-Ming Duan[1,2]

Part of the challenge for quantum many-body problems comes from the difficulty of representing large-scale quantum states, which in general requires an exponentially large number of parameters. Neural networks provide a powerful tool to represent quantum many-body states. An important open question is what characterizes the representational power of deep and shallow neural networks, which is of fundamental interest due to the popularity of deep learning methods. Here, we give a proof that, assuming a widely believed computational complexity conjecture, a deep neural network can efficiently represent most physical states, including the ground states of many-body Hamiltonians and states generated by quantum dynamics, while a shallow network representation with a restricted Boltzmann machine cannot efficiently represent some of those states.

Condensed matter physics applications

nature COMMUNICATIONS

ARTICLE

DOI: 10.1038/s41467-017-00705-2 OPEN

Efficient representation of quantum many-body states with deep neural networks

Xun Gao[1] & Lu-Ming Duan[1,2]

COOL

Part of the challenge for quantum many-body problems comes from the difficulty of representing large-scale quantum states, which in general requires an exponentially large number of parameters. Neural networks provide a powerful tool to represent quantum many-body states. An important open question is what characterizes the representational power of deep and shallow neural networks, which is of fundamental interest due to the popularity of deep learning methods. Here, we give a proof that, assuming a widely believed computational complexity conjecture, a deep neural network can efficiently represent most physical states, including the ground states of many-body Hamiltonians and states generated by quantum dynamics, while a shallow network representation with a restricted Boltzmann machine cannot efficiently represent some of those states.

Condensed matter physics applications

Constructing exact representations of quantum many-body systems with deep neural networks

Giuseppe Carleo
Center for Computational Quantum Physics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA and Institute for Theoretical Physics, ETH Zurich, Wolfgang-Pauli-Str. 27, 8093 Zurich, Switzerland

Yusuke Nomura and Masatoshi Imada
Department of Applied Physics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

We develop a constructive approach to generate artificial neural networks representing the exact ground states of a large class of many-body lattice Hamiltonians. It is based on the deep Boltzmann machine architecture, in which two layers of hidden neurons mediate quantum correlations among physical degrees of freedom in the visible layer. The approach reproduces the exact imaginary-time Hamiltonian evolution, and is completely deterministic. In turn, compact and exact network representations for the ground states are obtained without stochastic optimization of the network parameters. The number of neurons grows linearly with the system size and total imaginary time, respectively. Physical quantities can be measured by sampling configurations of both physical and neuron degrees of freedom. We provide specific examples for the transverse-field Ising and Heisenberg models by implementing efficient sampling. As a compact, classical representation for many-body quantum systems, our approach is an alternative to the standard path integral, and it is potentially useful also to systematically improve on numerical approaches based on the restricted Boltzmann machine architecture.

INTRODUCTION

A tremendous amount of successful developments in quantum physics builds upon the mapping between many-body quantum systems and effective classical theories. The probably most well known mapping is due

metric representations of quantum states, where the effective parameters are determined by means of the variational principle [16–19]. In matrix-product and tensor-network-states the ground-state is expressed as a classical network [20, 21]. In general, finding alternative, efficient classical representations of quantum states can help establishing novel numerical and analytical techniques to

Condensed matter physics applications

What else?

3

Identifying the relevant degrees of freedom (in a RG sense)

Koch-Janusz, Ringel,
Nat. Phys. 14, 578 (2018)

4

Quantum state tomography

Rocchetto et al.,
arXiv:1712.00127 (2017)

Torlai et al.,
Nature Physics 14, 447 (2018)

5

Using tensor networks, DMRG, MERA, etc. for traditional (classical) machine-learning tasks

Stoudenmire, Schwab,
Adv. Neural Inf. Proc. Sys. 29, 4799 (2016)

+ Follow-up papers

+ **More efficient Monte-Carlo samplings** (Huang and Wang, Liu et al., 2017),
Electronic structure calculations (Grisafi et al., 2017),
Design of materials by ML combined with DMFT (Arsenault et al., 2014),
Etc.

Condensed matter physics applications

ARTICLES
https://doi.org/10.1038/s41567-018-0081-4

nature physics

Mutual information, neural networks and the renormalization group

Maciej Koch-Janusz[1]* and Zohar Ringel[2]

Physical systems differing in their microscopic details often display strikingly similar behaviour when probed at macroscopic scales. Those universal properties, largely determining their physical characteristics, are revealed by the powerful renormalization group (RG) procedure, which systematically retains 'slow' degrees of freedom and integrates out the rest. However, the important degrees of freedom may be difficult to identify. Here we demonstrate a machine-learning algorithm capable of identifying the relevant degrees of freedom and executing RG steps iteratively without any prior knowledge about the system. We introduce an artificial neural network based on a model-independent, information-theoretic characterization of a real-space RG procedure, which performs this task. We apply the algorithm to classical statistical physics problems in one and two dimensions. We demonstrate RG flow and extract the Ising critical exponent. Our results demonstrate that machine-learning techniques can extract abstract physical concepts and consequently become an integral part of theory- and model-building.

Machine learning has been captivating public attention lately due to groundbreaking advances in automated translation, image and speech recognition[1], game-playing[2] and achieving super-human performance in tasks in which humans excelled while more traditional algorithmic approaches struggled[3]. The applications of those techniques in physics are very recent, initially

a Boltzmann distribution; no further knowledge about the microscopic details of the system is provided. The internal parameters of the network, which ultimately encode the degrees of freedom of interest at each step, are optimized ('learned', in neural network parlance) by a training algorithm based on evaluating real-space mutual information (RSMI) between spatially separated regions.

Condensed matter physics applications

ARTICLES
https://doi.org/10.1038/s41567-018-0081-4

nature physics

Mutual information, neural networks and the renormalization group

Maciej Koch-Janusz[1*] and Zohar Ringel[2]

Physical systems differing in their microscopic details often display strikingly similar behaviour when probed at macroscopic scales. Those universal properties, largely determining their physical characteristics, are revealed by the powerful renormalization group (RG) procedure, which systematically retains 'slow' degrees of freedom and integrates out the rest. However, the important degrees of freedom may be difficult to identify. Here we demonstrate a machine-learning algorithm capable of identifying the relevant degrees of freedom and executing RG steps iteratively without any prior knowledge about the system. We introduce an artificial neural network based on a model-independent, information-theoretic characterization of a real-space RG procedure, which performs this task. We apply the algorithm to classical statistical physics problems in one and two dimensions. We demonstrate RG flow and extract the Ising critical exponent. Our results demonstrate that machine-learning techniques can extract abstract physical concepts and consequently become an integral part of theory- and model-building.

Machine learning has been captivating public attention lately due to groundbreaking advances in automated translation, image and speech recognition[1], game-playing[2] and achieving super-human performance in tasks in which humans excelled while more traditional algorithmic approaches struggled[3]. The applications of those techniques in physics are very recent, initially

a Boltzmann distribution; no further knowledge about the microscopic details of the system is provided. The internal parameters of the network, which ultimately encode the degrees of freedom of interest at each step, are optimized ('learned', in neural network parlance) by a training algorithm based on evaluating real-space mutual information (RSMI) between spatially separated regions.

Machine-learned RG

Koch-Janusz, Ringel,
Nat. Phys. 14, 578 (2018)

Data

Configurations sampled from: $P(\mathcal{X}) = \frac{1}{Z} e^{-\beta H(\mathcal{X})}$

E.g., classical Ising variables $x_i \in \{0, 1\}$

Known Hamiltonian

RG procedure

Identify "relevant" degrees of freedom, throw away (integrate out) "irrelevant" ones

Relevant = capturing long-distance physics, carrying **information about the system at large** (as opposed to local fluctuations)

Iterate to get a universal, low-energy effective theory

Neural network

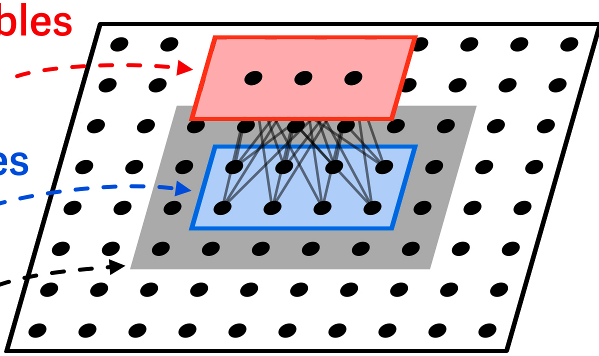
Restricted Boltzmann Machine (RBM)

"Hidden" variables $\mathcal{H} = \{h_j\}$

"Visible" variables $\mathcal{V} = \{v_i\}$

Buffer region

"Environment" variables $\mathcal{E} = \{e_i\}$



Encodes $P_\Lambda(\mathcal{H}|\mathcal{V}) = \frac{1}{Z_\Lambda} e^{-E_\Lambda(\mathcal{V},\mathcal{H})}$

$$E_\Lambda(\mathcal{V}, \mathcal{H}) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_{i,j} v_i \lambda_{ij} h_j$$

Cost function

Maximize the mutual information:

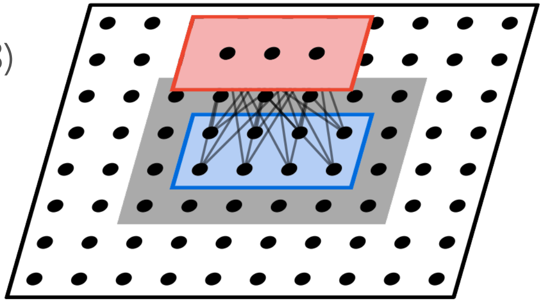
$$I_\Lambda(\mathcal{H} : \mathcal{E}) = \sum_{\mathcal{H},\mathcal{E}} P_\Lambda(\mathcal{H}, \mathcal{E}) \log \left(\frac{P_\Lambda(\mathcal{H}, \mathcal{E})}{P_\Lambda(\mathcal{H}) P(\mathcal{E})} \right)$$

To estimate it, **need to estimate** $P(\mathcal{V}, \mathcal{E})$ **and** $P(\mathcal{V})$ **first** (using, e.g., 2 other RBMs)

Machine-learned RG

Illustrative examples

Koch-Janusz, Ringel,
Nat. Phys. 14, 578 (2018)



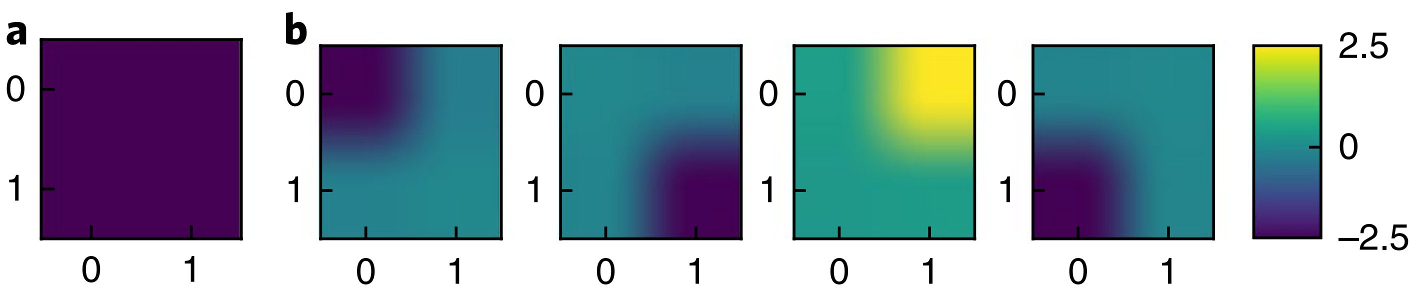
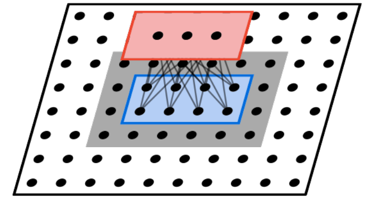


Fig. 2 | The weights of the RSMI network trained on the Ising model. Visualization of the weights of the RSMI network trained on the Ising model for a visibile area $\mathcal{V}$ of 2×2 spins. The ANN couples strongly to areas with large absolute value of the weights. **a**, The weights for $N_h = 1$ hidden neuron: the ANN discovers Kadanoff blocking. **b**, The weights for $N_h = 4$ hidden neurons: each neuron tracks one original spin.

Machine-learned RG

Koch-Janusz, Ringel, Nat. Phys. 14, 578 (2018)



Illustrative examples

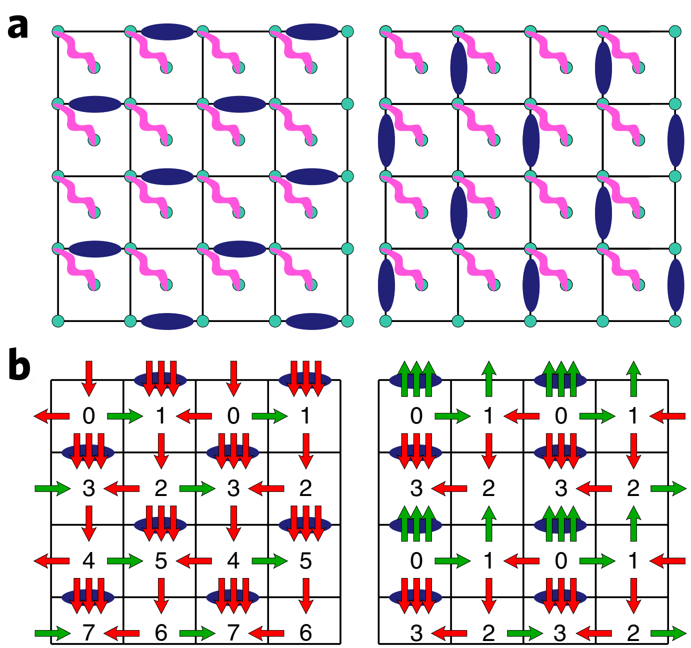
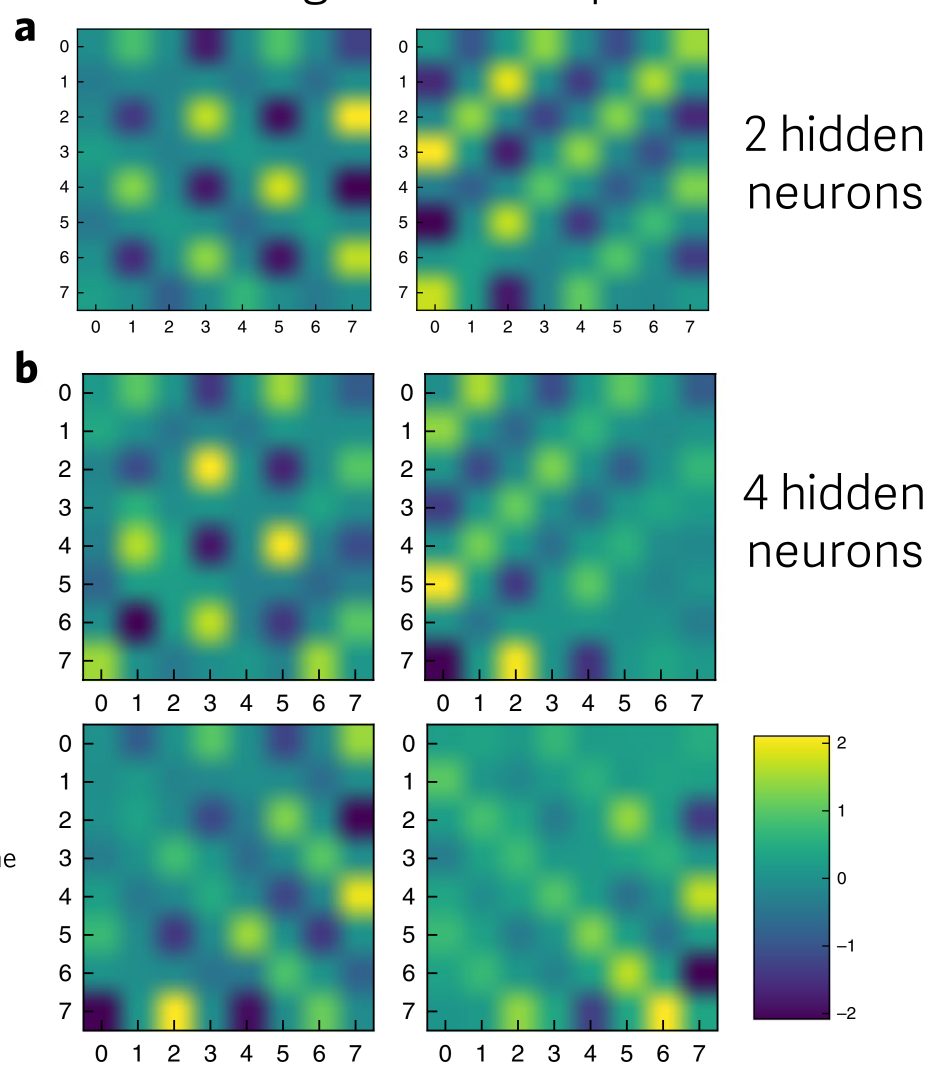


Fig. 3 | The dimer model. a, Two sample dimer configurations (blue links), corresponding to the E_y and E_x electrical fields, respectively. The coupled pairs of additional spin degrees of freedom on vertices and faces of the lattice (wiggly lines) are decoupled from the dimers and from each other. Their fluctuations constitute irrelevant noise. **b**, An example of mapping the dimer model to local electric fields. The so-called staggered configuration on the left maps to uniform non-vanishing field in the vertical direction: $\langle E_y \rangle \neq 0$. The 'columnar' configuration on the right produces both E_x and E_y that are zero on average (see ref. [36] for details of the mapping).

Weights of hidden neurons for visible region of 8 x 8 spins



Take-home messages

WIKIPEDIA

A **physical theory** is a model of physical events. It is judged by the extent to which its predictions agree with empirical observations. The quality of a physical theory is also judged on its ability to make new predictions which can be verified by new observations. A physical theory differs from a mathematical theorem in that while both are based on some form of axioms, judgment of mathematical applicability is not based on agreement with any experimental results.[2][3]

1 **Machine learning is awesome**

It really is

2 It is a set of tools for **learning useful representations from complex data**

Representations of a cat,
of a quantum phase,
of a wavefunction, etc.

3 It is **a lot like physics**

It **comes from** (statistical) **physics**
(and statistical learning)

It is **partly empirical, aimed at** making new **predictions**

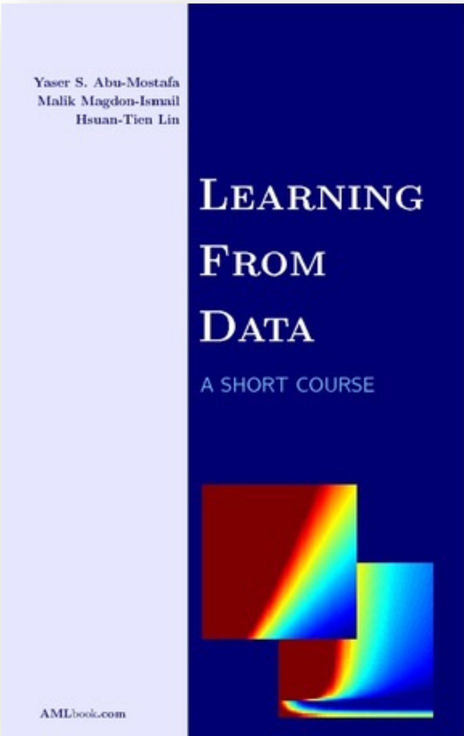
4 **Deep learning drives the recent advances** (and hype)

5 **Physicists are both ahead and behind**

State-of-the-art techniques from physics can be useful in machine learning, and vice versa

Exciting stuff ahead!

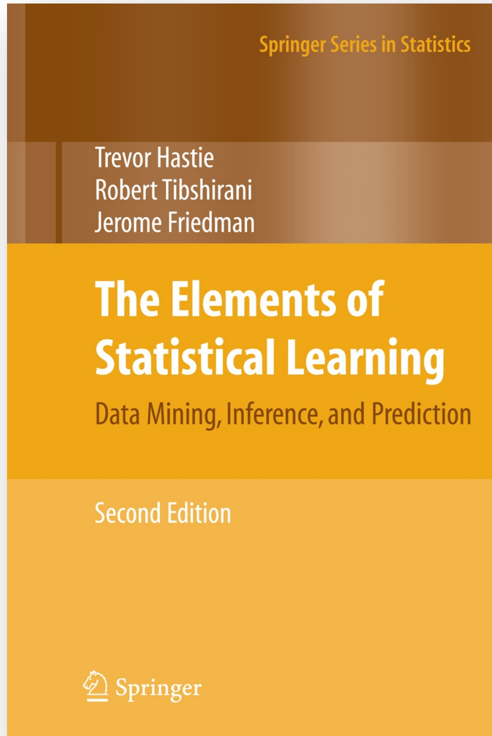
To machine-learn like a boss



Learning From Data

Abu-Mostafa et al. (2012)

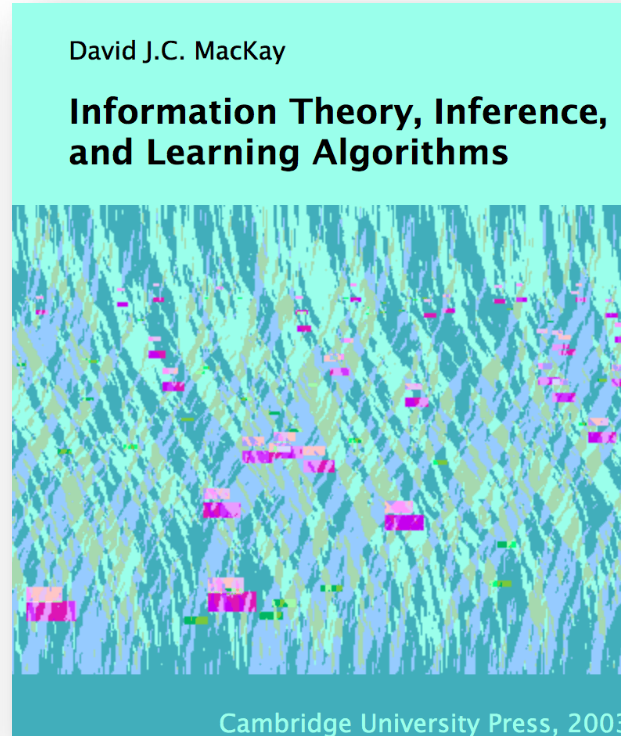
Basic concepts of statistical learning theory



The Elements of Statistical Learning

Hastie et al. (2001)

More advanced, more mathematical

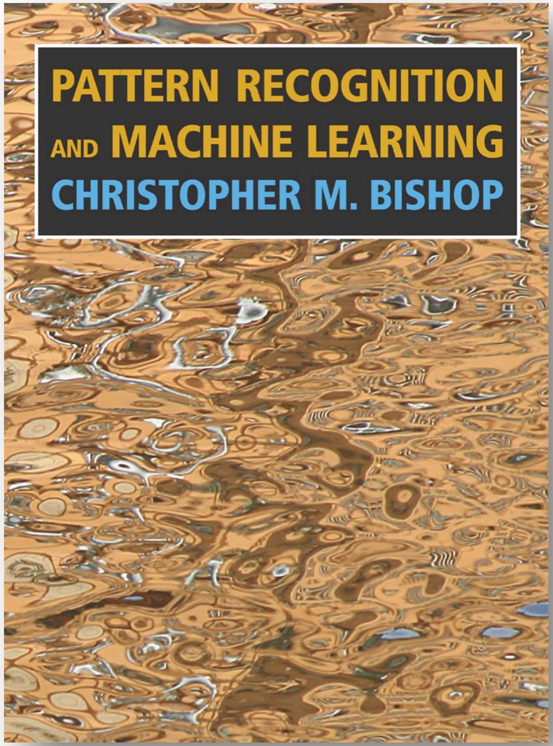


Information Theory, Inference, and Learning Algorithms

MacKay (2003)

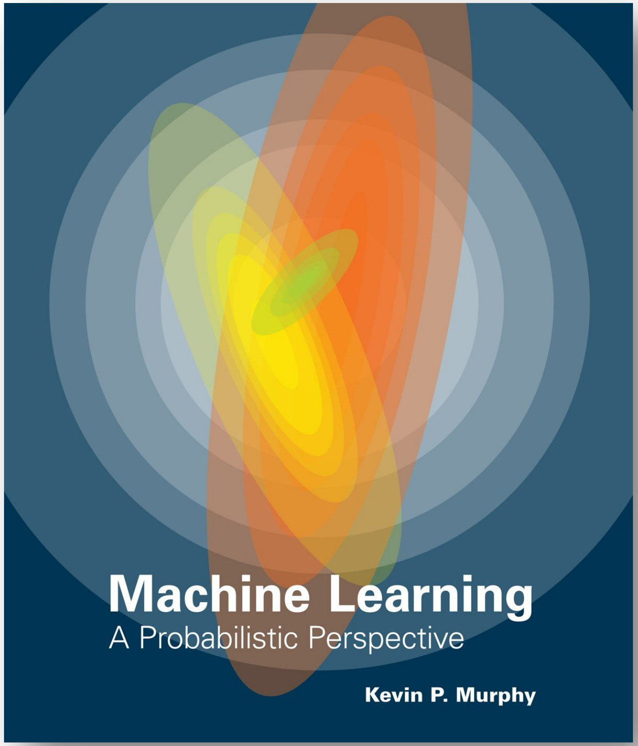
A classic for Bayesian inference and information theory

To machine-learn like a boss



Pattern Recognition and Machine Learning,
Bishop (2006)

Comprehensive book on modern machine-learning techniques



Machine Learning: A Probabilistic Perspective,
Murphy (2012)

More recent comprehensive book (seems good but I don't know)

To machine-learn like a boss

Neural Networks and Deep Learning

Neural Networks and Deep Learning is a free online book. The book will teach you about:

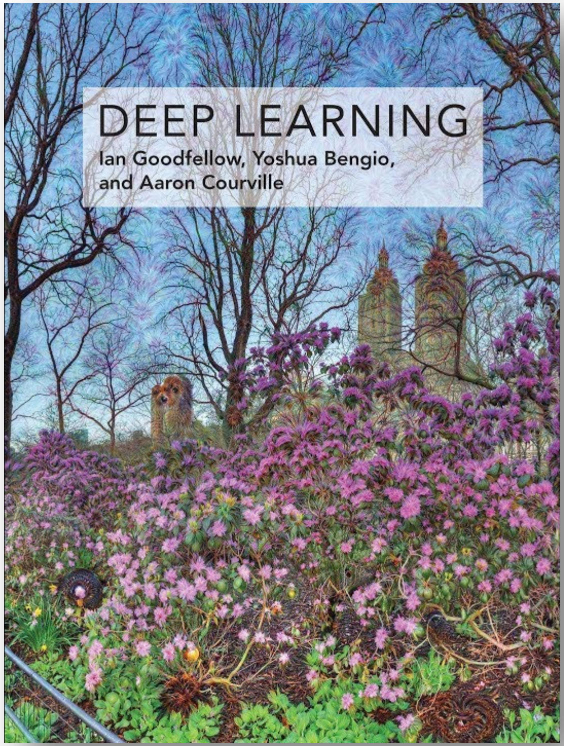
- Neural networks, a beautiful biologically-inspired programming paradigm which enables a computer to learn from observational data
- Deep learning, a powerful set of techniques for learning in neural networks

Neural networks and deep learning currently provide the best solutions to many problems in image recognition, speech recognition, and natural language processing. This book will teach you many of the core concepts behind neural networks and deep learning.

For more details about the approach taken in the book, see here. Or you can jump directly to Chapter 1 and get started.

Neural Networks and Deep Learning,
Nielsen (2015)

Great introduction to neural nets and deep learning, easy read



Deep Learning,
Goodfellow et al. (2016)

THE textbook for deep learning

Reviews aimed at physicists

1 **A high-bias, low-variance introduction to Machine Learning for physicists**

Pankaj Mehta, Ching-Hao Wang, Alexandre G. R. Day, and Clint Richardson

Department of Physics,

2 **Machine learning & artificial intelligence in the quantum domain**

Vedran Dunjko

Institute for Theoretical Physics, University of Innsbruck, Innsbruck 6020, Austria
Max Planck Institute of Quantum Optics, Garching 85748, Germany
Email: vedran.dunjko@mpq.mpg.de

Hans J. Briegel

Institute for Theoretical Physics, University of Innsbruck Innsbruck 6020, Austria
Department of Philosophy, University of Konstanz, Konstanz 78457, Germany
Email: hans.briegel@uibk.ac.at

Abstract. Quantum information technologies, on the one side, and intelligent learning systems, on the other, are both emergent technologies that will likely have a transforming impact on our society in the future. The respective underlying fields of basic research – quantum information (QI) versus machine learning and artificial intelligence (AI) – have their own specific questions and challenges, which have hitherto been investigated largely independently. However, in a growing body of recent work, researchers have been probing the question to what extent these fields can indeed learn and benefit from each other. QML explores the interaction between quantum computing and machine learning, investigating how results and techniques from one field can be used to solve the problems of the other. In recent time, we have witnessed significant breakthroughs in both directions of influence. For instance, quantum computing is finding a vital application in providing speed-ups for machine learning problems, critical in our "big data" world. Conversely, machine learning already permeates many cutting-edge technologies, and may become instrumental in advanced quantum technologies. Aside from quantum speed-up in data analysis, or classical machine learning optimization used in quantum experiments, quan-

Deep learning libraries

Library	Rank	Overall	Github	Stack Overflow	Google Results
tensorflow	1	10.87	4.25	4.37	2.24
keras	2	1.93	0.61	0.83	0.48
caffe	3	1.86	1.00	0.30	0.55
theano	4	0.76	-0.16	0.36	0.55
pytorch	5	0.48	-0.20	-0.30	0.98
sonnet	6	0.43	-0.33	-0.36	1.12
mxnet	7	0.10	0.12	-0.31	0.28
torch	8	0.01	-0.15	-0.01	0.17
cntk	9	-0.02	0.10	-0.28	0.17
dlib	10	-0.60	-0.40	-0.22	0.02
caffe2	11	-0.67	-0.27	-0.36	-0.04
chainer	12	-0.70	-0.40	-0.23	-0.07
paddlepaddle	13	-0.83	-0.27	-0.37	-0.20
deeplearning4j	14	-0.89	-0.06	-0.32	-0.51
lasagne	15	-1.11	-0.38	-0.29	-0.44
bigdl	16	-1.13	-0.46	-0.37	-0.30
dynet	17	-1.25	-0.47	-0.37	-0.42
apache singa	18	-1.34	-0.50	-0.37	-0.47
nvidia digits	19	-1.39	-0.41	-0.35	-0.64
matconvnet	20	-1.41	-0.49	-0.35	-0.58
tflearn	21	-1.45	-0.23	-0.28	-0.94
nervana neon	22	-1.65	-0.39	-0.37	-0.89
opennn	23	-1.97	-0.53	-0.37	-1.07

Source: The Data Incubator

Deep learning libraries

Library	Rank	Overall	Github	Stack Overflow	Google Results
tensorflow	1	10.87	4.25	4.37	2.24
keras	2	1.93	0.61	0.83	0.48

	Languages	Tutorials and training materials	CNN modeling capability	RNN modeling capability	Architecture: easy-to-use and modular front end	Speed	Multiple GPU support	Keras compatible
Theano	Python, C++	++	++	++	+	++	+	+
Tensor-Flow	Python	+++	+++	++	+++	++	++	+
Torch	Lua, Python (new)	+	+++	++	++	+++	++	
Caffe	C++	+	++		+	+	+	
MXNet	R, Python, Julia, Scala	++	++	+	++	++	+++	
Neon	Python	+	++	+	+	++	+	
CNTK	C++	+	+	+++	+	++	+	

Source: The Data Incubator

Deep learning libraries

Library	Rank	Overall	Github	Stack Overflow	Google Results
tensorflow	1	10.87	4.25	4.37	2.24
keras	2	1.93	0.61	0.83	0.48

	Languages	Tutorials and training materials	CNN modeling capability	RNN modeling capability	Architecture: easy-to-use and modular front end	Speed	Multiple GPU support	Keras compatible
Theano	Python, C++	++	++	++	+	++	+	+
Tensor-Flow	Python	+++	+++	++	+++	++	++	+
Torch	Lua, Python (new)	+	+++	++	++	+++	++	
Caffe	C++	+	++	[illegible]	[illegible]	+	+	
MXNet	R, Python, Julia, Scala	++	++	[illegible]	++	++	+++	
Neon	Python	+	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	
CNTK	C++	+	+	+++	+	++	+	

All you need is (love and):

TensorFlow

By the one and only Google Brain Team

+

Keras

High-level library for neural networks, using TensorFlow or Theano as backend

Source: The Data Incubator

Deep learning libraries

Library	Rank	Overall	Github	Stack Overflow	Google Results
tensorflow	1	10.87	4.25	4.37	2.24
keras	2	1.93	0.61	0.83	0.48

	Languages	Tutorials and training materials	CNN modeling capability	RNN modeling capability	Architecture: easy-to-use and modular front end	Speed	Multiple GPU support	Keras compatible
Theano	Python, C++	++	++	++	[illegible]	[illegible]	+	+
Tensor-Flow	Python	+++	+++	++	[illegible]	[illegible]	++	+
Torch	Lua, Python (new)	+	+++	++	++	+++	++	
Caffe	C++	+	++	[illegible]	+	+	+	
MXNet	R, Python, Julia, Scala	++	++	[illegible]	++	++	+++	
Neon	Python	+	++	[illegible]	+	++	+	
CNTK	C++	+	+	+++	+	++	+	

Machine learning with a minimal amount of (Python) code

All you need is (love and):

TensorFlow + Keras

By the one and only Google Brain Team

High-level library for neural networks, using TensorFlow or Theano as backend

Source: The Data Incubator

Only one way to write less code: pigeons

RESEARCH ARTICLE

Pigeons (*Columba livia*) as Trainable Observers of Pathology and Radiology Breast Cancer Images

Richard M. Levenson[1]*, Elizabeth A. Krupinski[3], Victor M. Navarro[2], Edward A. Wasserman[2]*

1 Department of Pathology and Laboratory Medicine, University of California Davis Medical Center, Sacramento, California, United States of America, **2** Department of Psychological and Brain Sciences, The University of Iowa, Iowa City, Iowa, United States of America, **3** Department of Radiology & Imaging Sciences, College of Medicine, Emory University, Atlanta, Georgia, United States of America

* levenson@ucdavis.edu (RML); ed-wasserman@uiowa.edu (EAW)

CrossMark
click for updates

OPEN ACCESS

Citation: Levenson RM, Krupinski EA, Navarro VM, Wasserman EA (2015) Pigeons (*Columba livia*) as Trainable Observers of Pathology and Radiology Breast Cancer Images. PLoS ONE 10(11): e0141357. doi:10.1371/journal.pone.0141357

Abstract

Pathologists and radiologists spend years acquiring and refining their medically essential visual skills, so it is of considerable interest to understand how this process actually unfolds and what image features and properties are critical for accurate diagnostic performance. Key insights into human behavioral tasks can often be obtained by using appropriate animal models. We report here that pigeons (*Columba livia*)—which share many visual system properties with humans—can serve as promising surrogate observers of medical images, a capability not previously documented. The birds proved to have a remarkable ability to distinguish benign from malignant human breast histopathology after training with differential food reinforcement; even more importantly, the pigeons were able to generalize what they had learned when confronted with novel image sets. The birds' histological accuracy, like

Only one way to write less code: pigeons

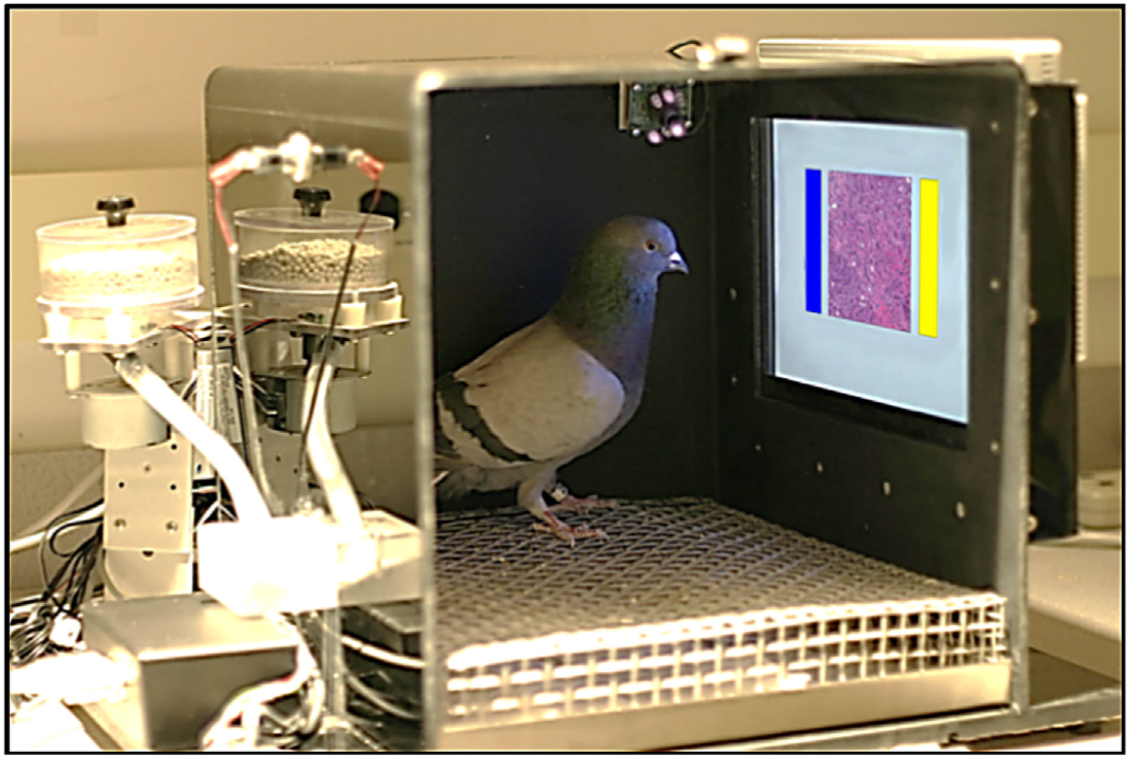


Fig 1. The pigeons' training environment. The operant conditioning chamber was equipped with a food pellet dispenser, and a touch-sensitive screen upon which the medical image (center) and choice buttons (blue and yellow rectangles) were presented.

It’s demo time!

Great Keras + TensorFlow tutorial

Introduction to Deep Neural Networks with Keras and Tensorflow

#tensorflow #python #tutorial #deep-learning #keras #keras-tutorials #keras-tensorflow

github.com/leriomaggio/deep-learning-keras-tensorflow

123 commits | 2 branches | 6 releases | 3 contributors | MIT

Branch: master | New pull request | Find file | Clone or download

leriomaggio added webvalley2017 logo among release logo images — Latest commit `ee1e0fe` on Aug 22, 2017

1. ANN	Fixed Path to data in notebook	a year ago
2. Deep Learning Frameworks	Repo Structure Refactoring + Merge from PyData London Version	a year ago
3. Fully Connected Networks and ...	Repo Structure Refactoring + Merge from PyData London Version	a year ago
4. Convolutional Neural Networks	Repo Structure Refactoring + Merge from PyData London Version	a year ago
5. HyperParameter Tuning and Tra...	Repo Structure Refactoring + Merge from PyData London Version	a year ago
6. AutoEncoders and Embeddings	Added Examples for Image Denoising and VAE	a year ago
7. Recurrent Neural Networks	Repo Structure Refactoring + Merge from PyData London Version	a year ago
8. Extra	Repo Structure Refactoring + Merge from PyData London Version	a year ago
data	Reorganised data for word embedding	a year ago
imgs	added webvalley2017 logo among release logo images	10 months ago
solutions	Fixed Typo in Solutions for ANN (ANN renamed as MLP class)	a year ago
0. Preamble.ipynb	Updated Outline in README and Preamble	a year ago
Conclusions.ipynb	Repo Structure Refactoring + Merge from PyData London Version	a year ago
LICENSE	New Materials for PyDataIt	a year ago


The End