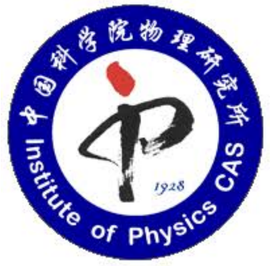
A Hitchhiker's Guide to Deep Learning

Lei Wang (王磊)
https://wangleiphy.github.io

Institute of Physics, Beijing
Chinese Academy of Sciences



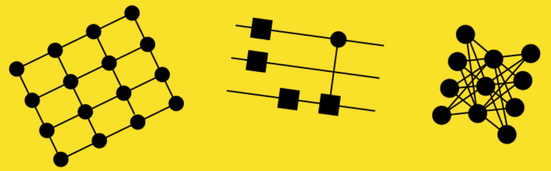
欢迎！

感谢！

祝有收获！

Spring School

深度学习&量子编程



王磊

深度学习：从理论到实践

以微分编程和表示学习为重点
介绍深度学习技术，并讲解它们在
统计物理和量子多体计算中的应用实例

张潘

从机器学习角度理解张量网络

从表述，优化，学习与泛化这
四个角度介绍张量网络及其
在应用数学和机器学习中的应用

罗秀哲

面向物理学家的Julia编程实践

以量子物理的工程实践为重点介绍
Julia语言，量子计算的基础概念，Julia
语言中的CUDA编程和量子物理工具链

刘金国

量子编程实践

介绍量子机器学习，量子优化算法和
量子化学中的研究前沿，基于Julia量子
计算库Yao.jl实现这些算法，介绍自动
微分与GPU编程在量子编程中的应用

报名方式： https://bit.ly/2CE5J8H

教学资料： https://github.com/QuantumBFS/SSSS

授课形式：

中文授课+程序演示+Hackathon (有奖品)

时间：2019年5月6-10日

地点：广东东莞

松山湖材料实验室

粤港澳交叉科学中心

Quantum Hackathon：

学员将通过组队的形式，完成
一个量子物理相关的编程挑战。
我们将评出表现突出的团队，
给予奖励。

Contact: wanglei@iphy.ac.cn

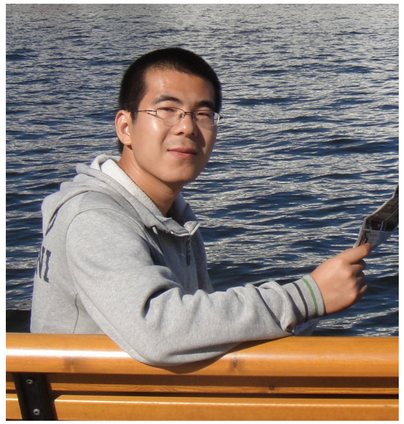
QuantumBFS
Yao Framework

julia

SONGSHAN LAKE
MATERIALS LABORATORY
松山湖材料实验室

The team

Lei Wang, IOP CAS



Pan Zhang, ITP CAS

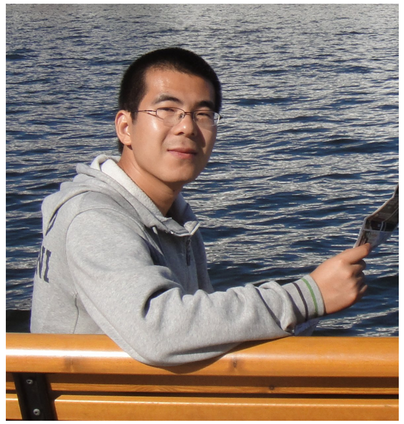
Hai-Jun Liao, IOP CAS

Jin-Guo Liu, IOP CAS

Xiu-Zhe Luo, Univ Waterloo

The team

Lei Wang, IOP CAS



Pan Zhang, ITP CAS

Hai-Jun Liao, IOP CAS

Jin-Guo Liu, IOP CAS

and YOU!

Xiu-Zhe Luo, Univ Waterloo

Spirit of the school

interactive

- First time we do it, so please expect some sharp edges

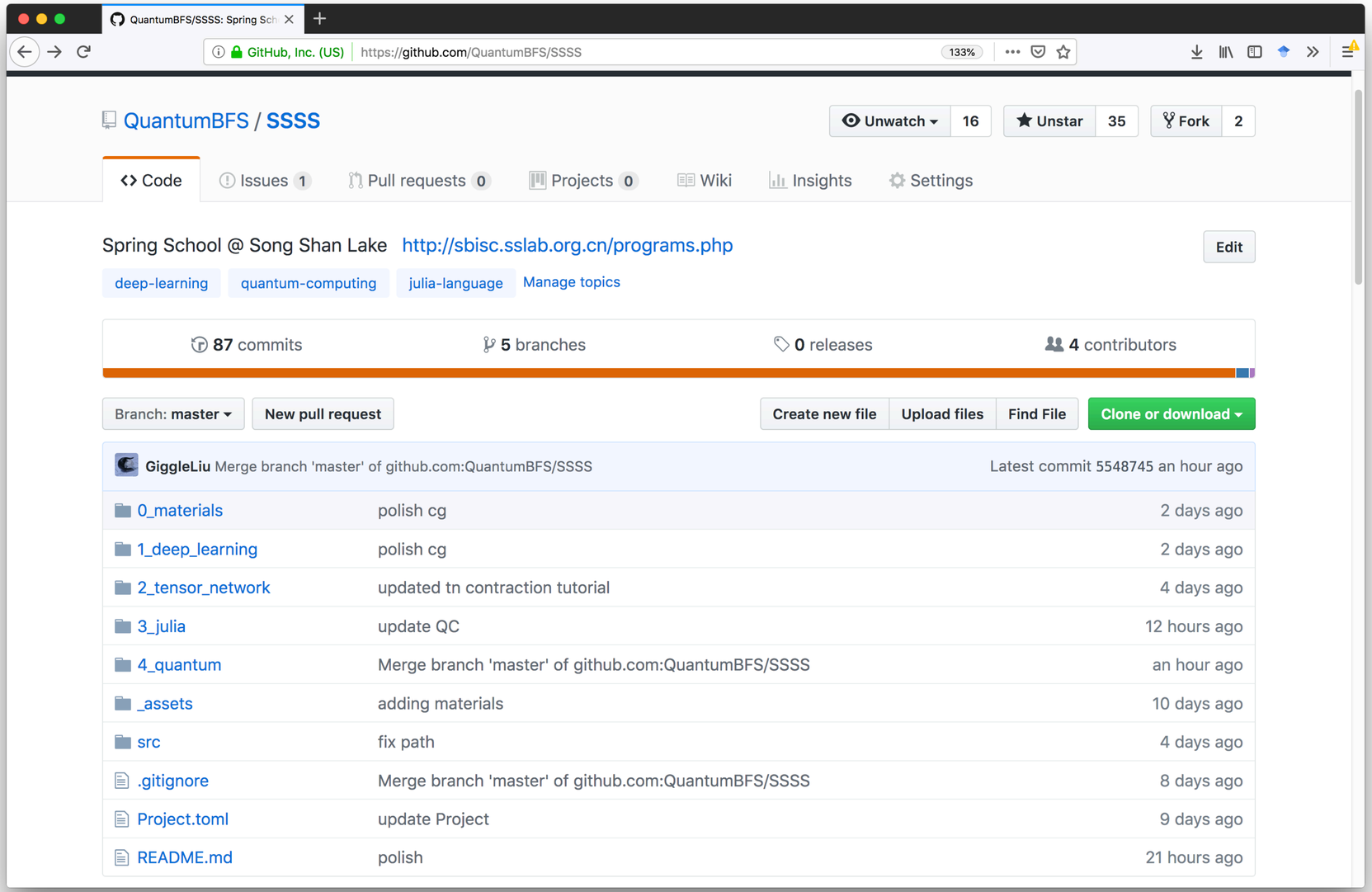
- To benefit more, please
  - Ask MANY questions
  - Think about the quiz

  - Get your hands dirty:
    - Play with the demo codes and notebooks
    - Participate in the code challenge

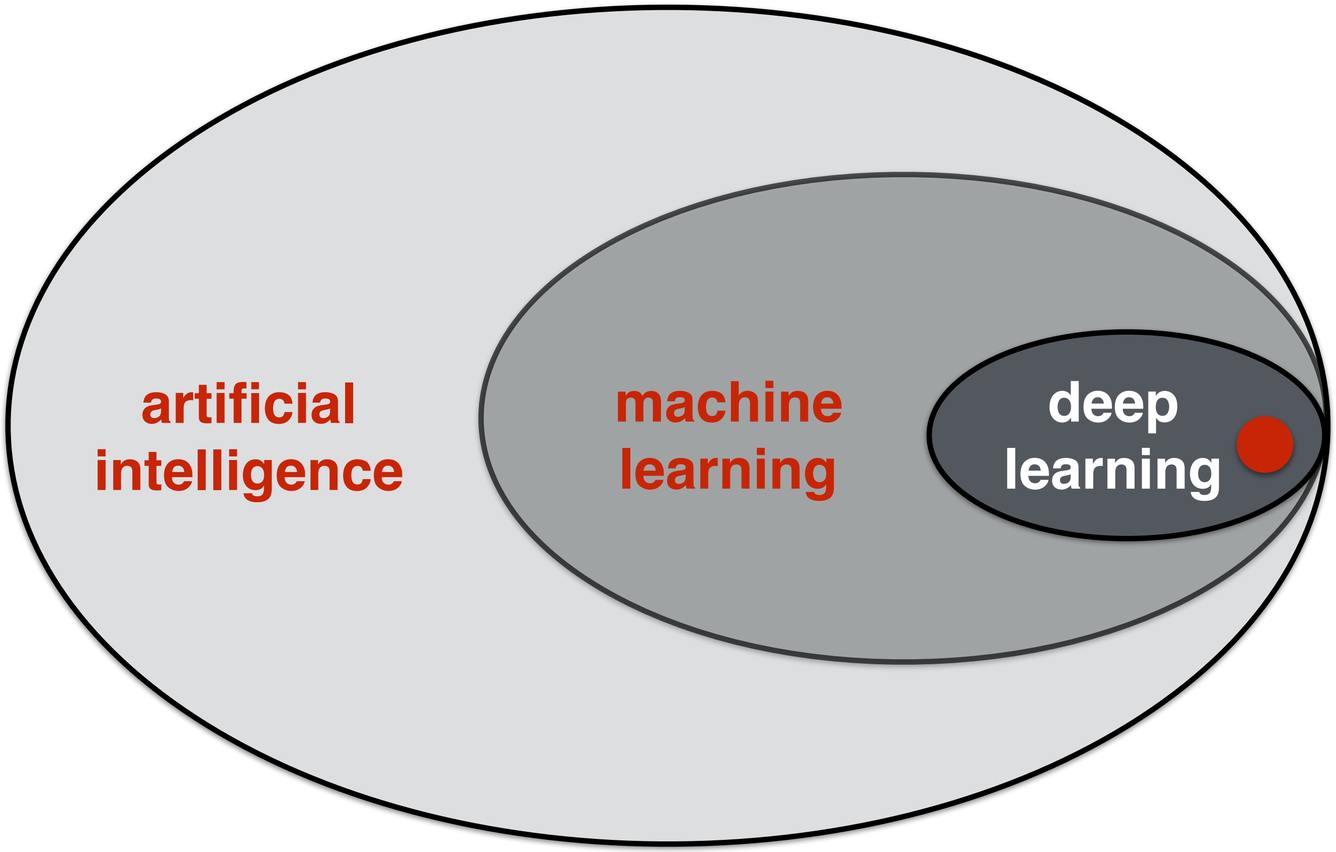
- We love feedbacks!

Resources, notes, demos, challenges

https://github.com/QuantumBFS/SSSS



Why deep learning ?



Game changing technology for scientific research

09:44 4G

Tweet

Kyle Cranmer and 6 others liked

Natasha Jaques
@natashajaques

.@demishassabis gave an absolutely packed talk yesterday about how @DeepMindAI is shifting its focus from games -> scientific benefit 🎉🎉🎉

05:33 · 2019/3/22 · Twitter for iPhone

Demis Hassabis, Google DeepMind
2019.3.20@MIT Center for
Brains, Minds and Machines
https://www.youtube.com/watch?v=cEOAerVz3UU

09:44 4G

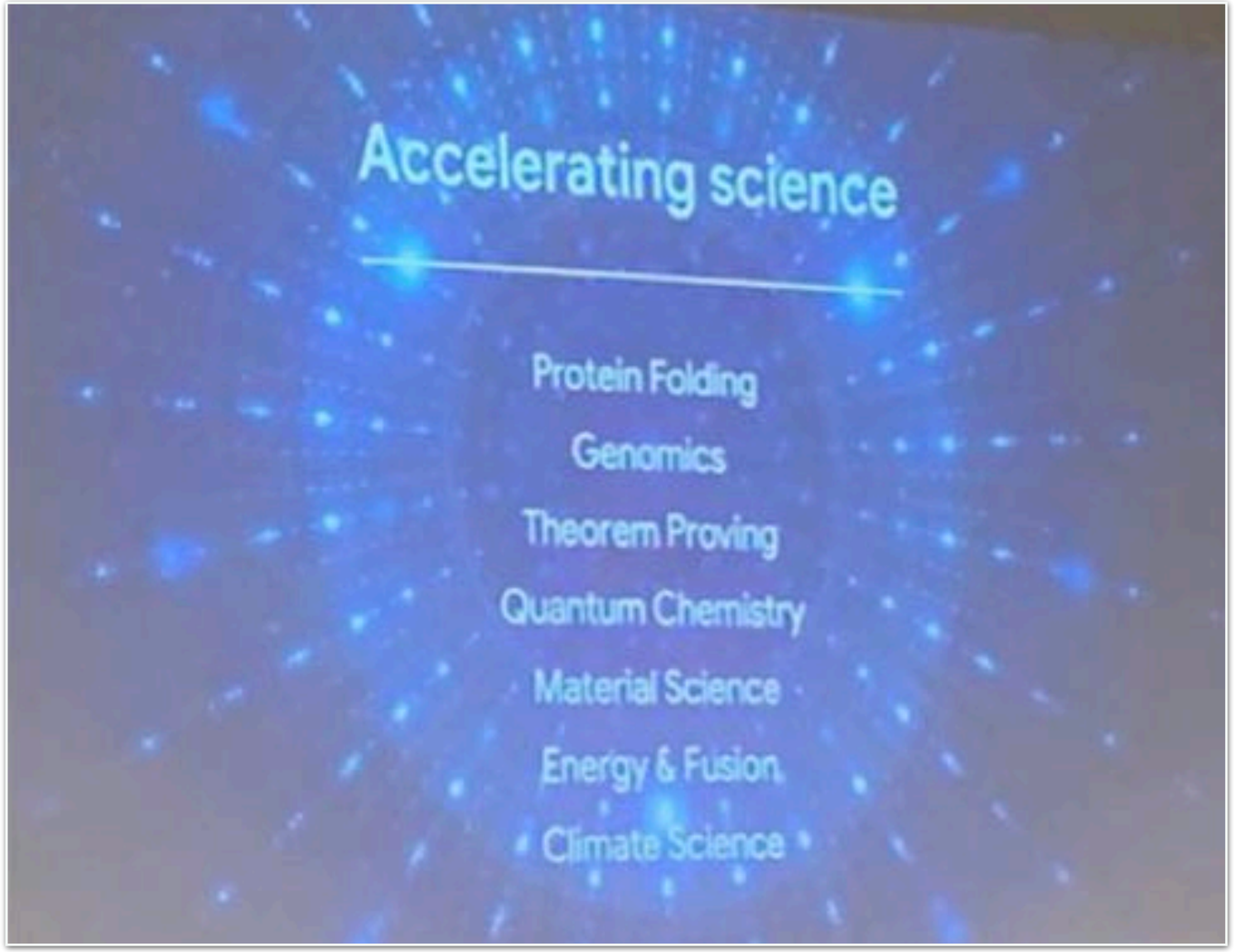
Tweet

Kyle Cranmer and 6 others liked

Natasha Jaques
@natashajaques

.@demishassabis gave an absolutely packed talk yesterday about how @DeepMindAI is shifting its focus from games -> scientific benefit 🎉🎉🎉

05:33 · 2019/3/22 · Twitter for iPhone

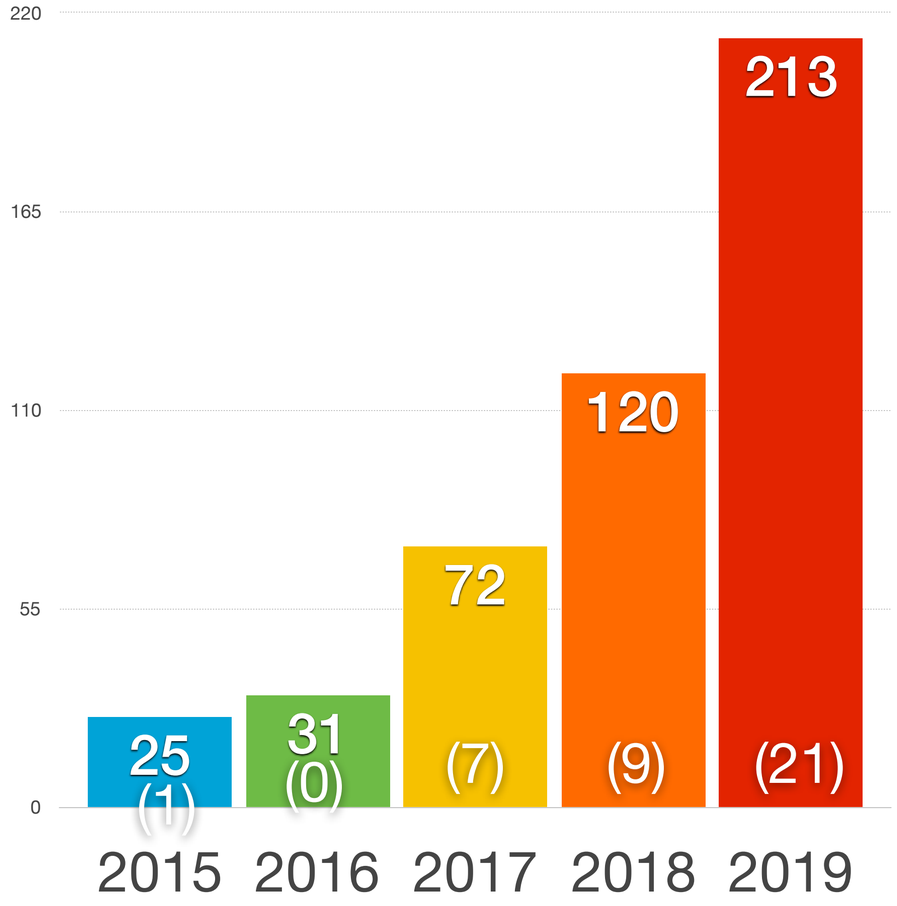


Demis Hassabis, Google DeepMind
2019.3.20@MIT Center for
Brains, Minds and Machines
https://www.youtube.com/watch?v=cEOAerVz3UU

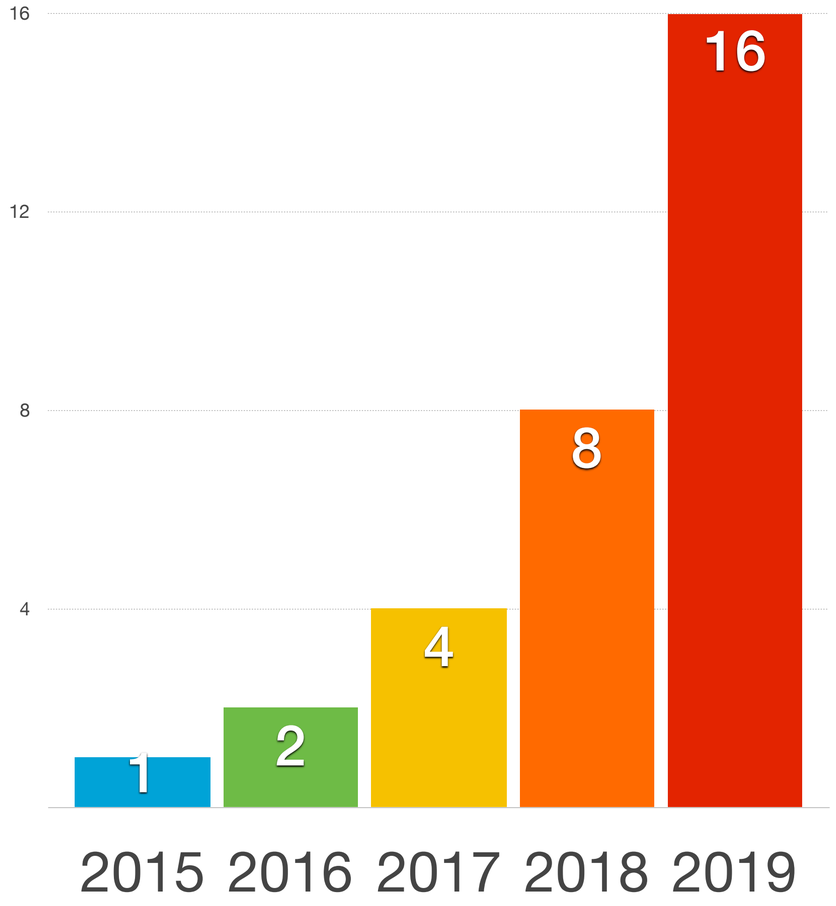
At APS March Meetings

with “machine learning” in the title/abstract

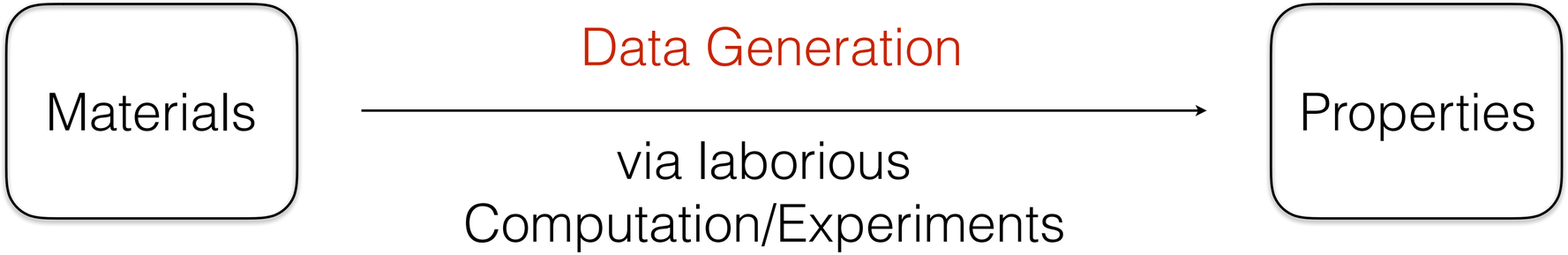
of (invited) talks



of sessions

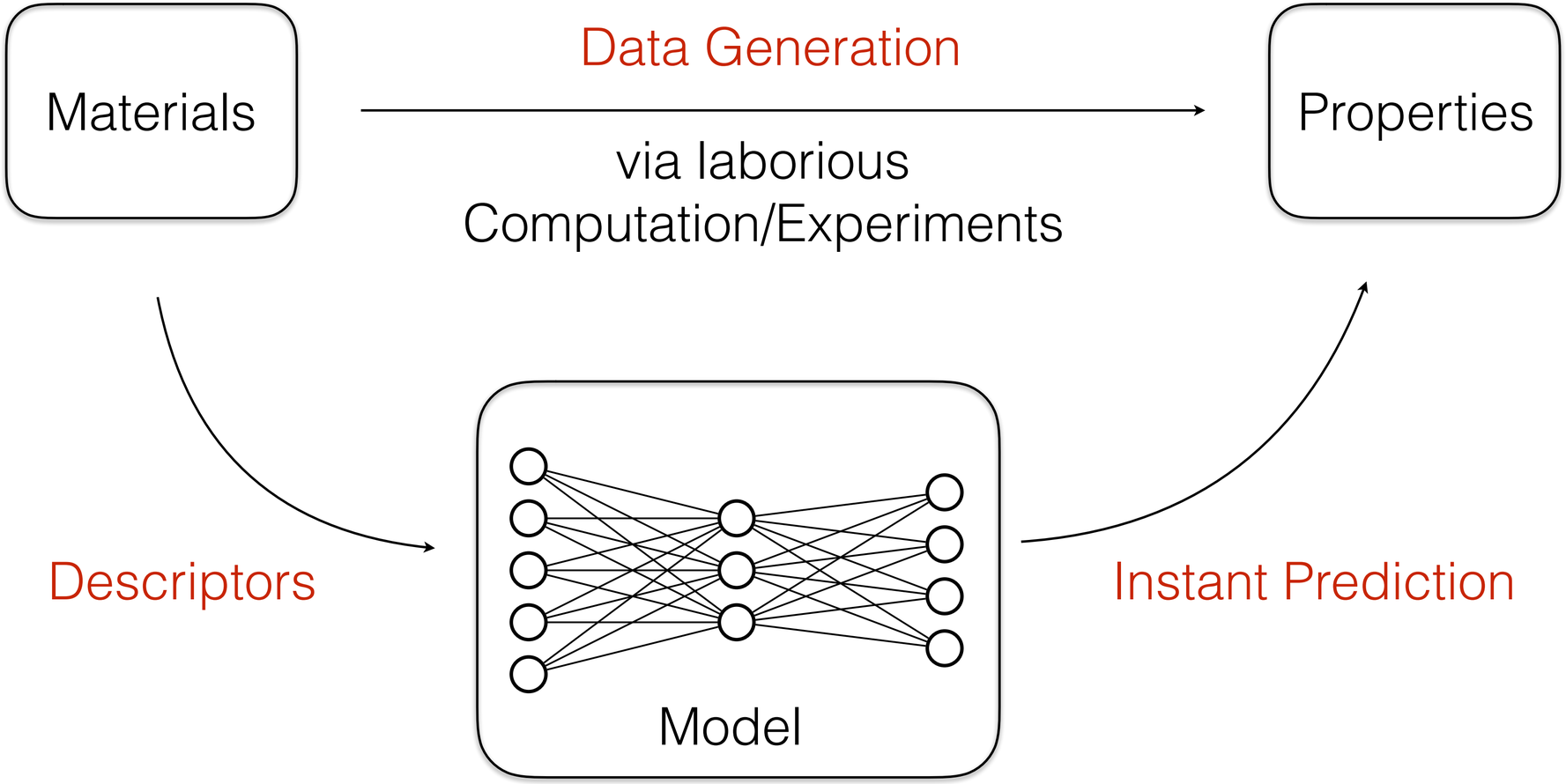


Materials Discovery/Ab Initio/DFT



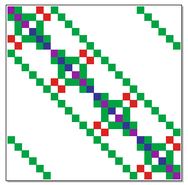
"Machine Learning in Materials Science: Recent Progress and Critical Next Steps"
Rampi Ramprasad @ IPAM program on Understanding Many-Particle Systems with ML, 2016

Materials Discovery/Ab Initio/DFT



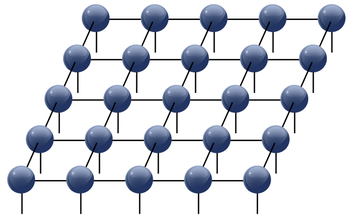
"Machine Learning in Materials Science: Recent Progress and Critical Next Steps"
Rampi Ramprasad @ IPAM program on Understanding Many-Particle Systems with ML, 2016

Quantum Many-Body Computation

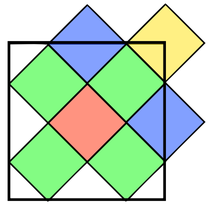


exact diagonalization

quantum Monte Carlo

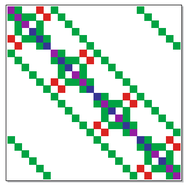


tensor network states



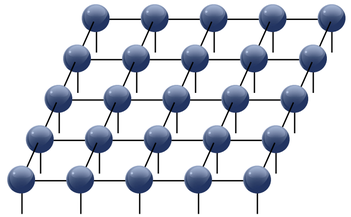
dynamical mean field theories

Quantum Many-Body Computation

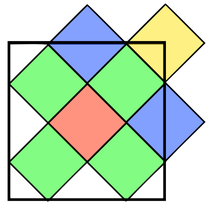


exact diagonalization

quantum Monte Carlo



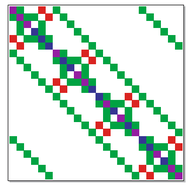
tensor network states



dynamical mean field theories

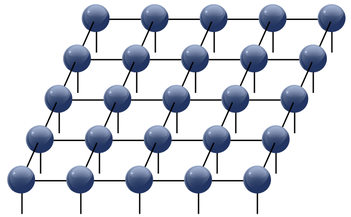
Algorithmic improvement in past 20 years outperformed Moore's law

Quantum Many-Body Computation

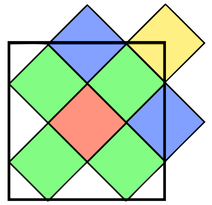


exact diagonalization

quantum Monte Carlo



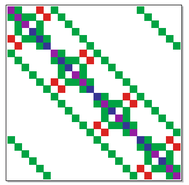
tensor network states



dynamical mean field theories

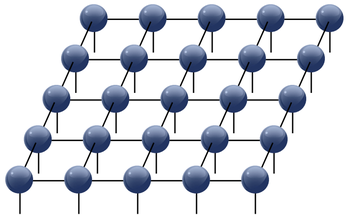
is faster than

Quantum Many-Body Computation

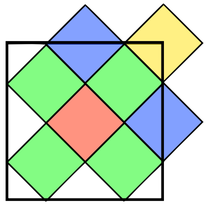


exact diagonalization

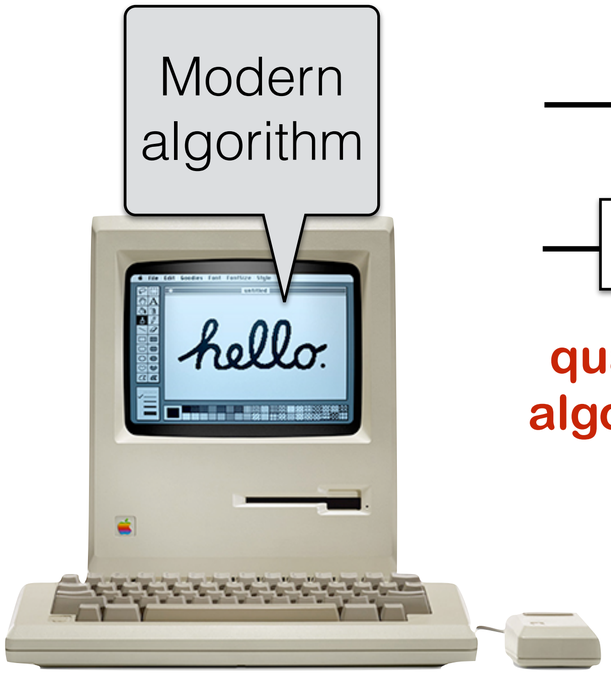
quantum Monte Carlo

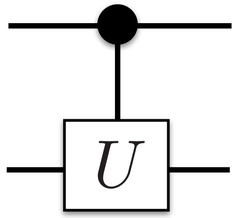


tensor network states

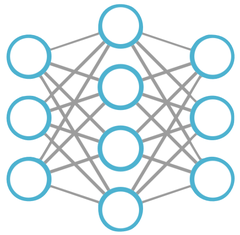


dynamical mean field theories

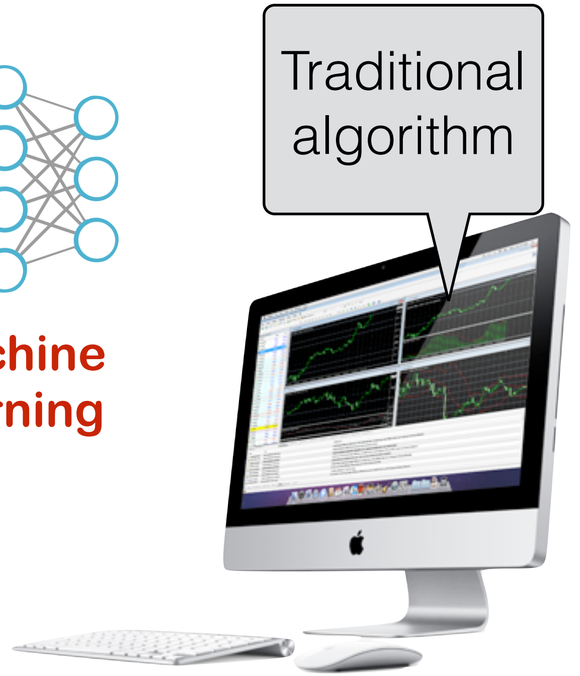




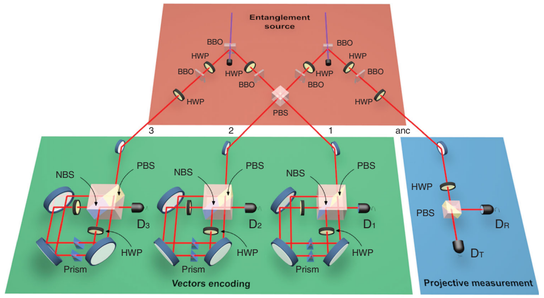
quantum algorithms



machine learning

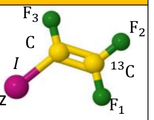


Quantum Information & Computation

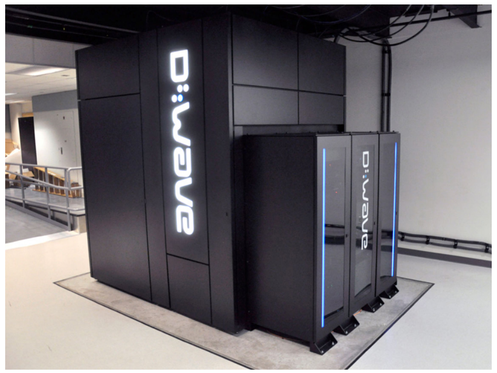


Cai et al, PRL 114, 110504 (2015)

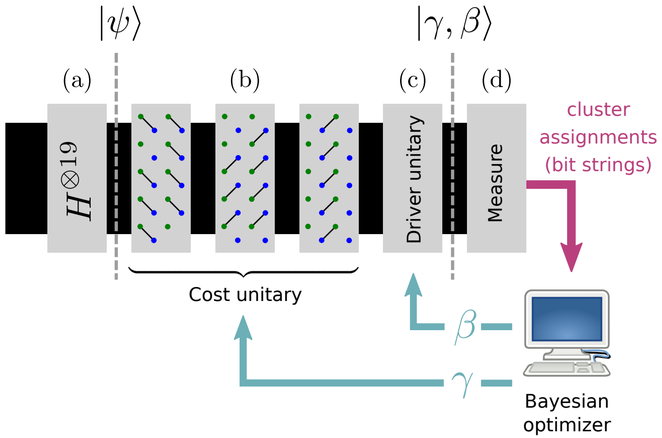
	^{13}C	F_1	F_2	F_3
^{13}C	15479.9Hz			
F_1	-297.7Hz	-33130.1Hz		
F_2	-275.7Hz	64.6Hz	-42681.4Hz	
F_3	39.1Hz	51.5Hz	-129.0Hz	-56443.5Hz
T_2^*	1.22s	0.66s	0.63s	0.61s
T_2	7.9s	4.4s	6.8s	4.8s



Li et al, PRL 114, 140504 (2015)



Perdomo-Ortiz et al, 1708.09757

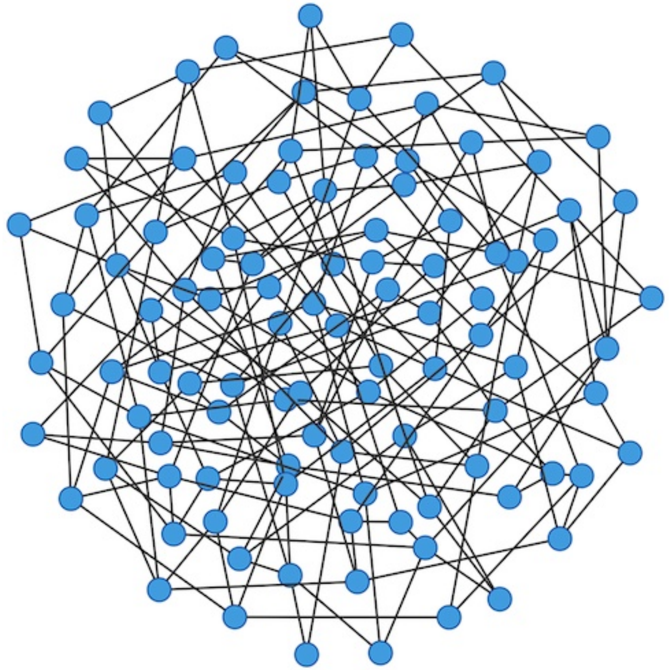


Rigetti Computing,1712.05771

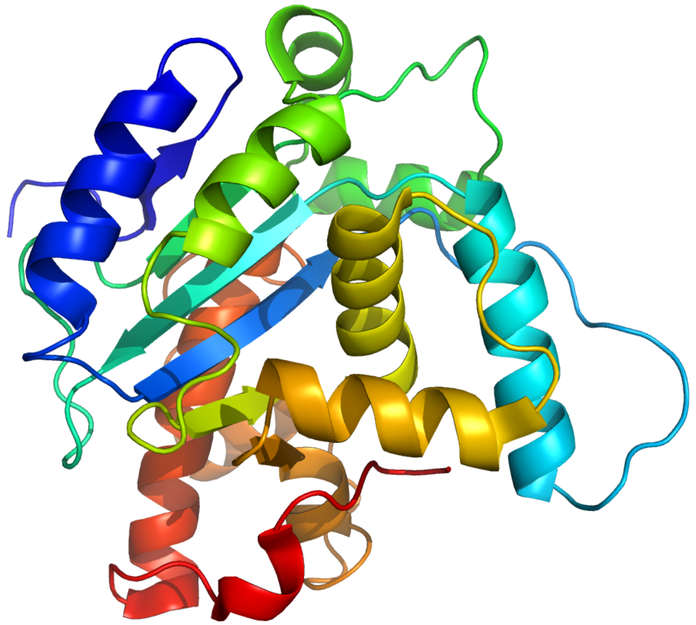
Review “Quantum machine learning”, Biamonte, Wittek et al, Nature 2017

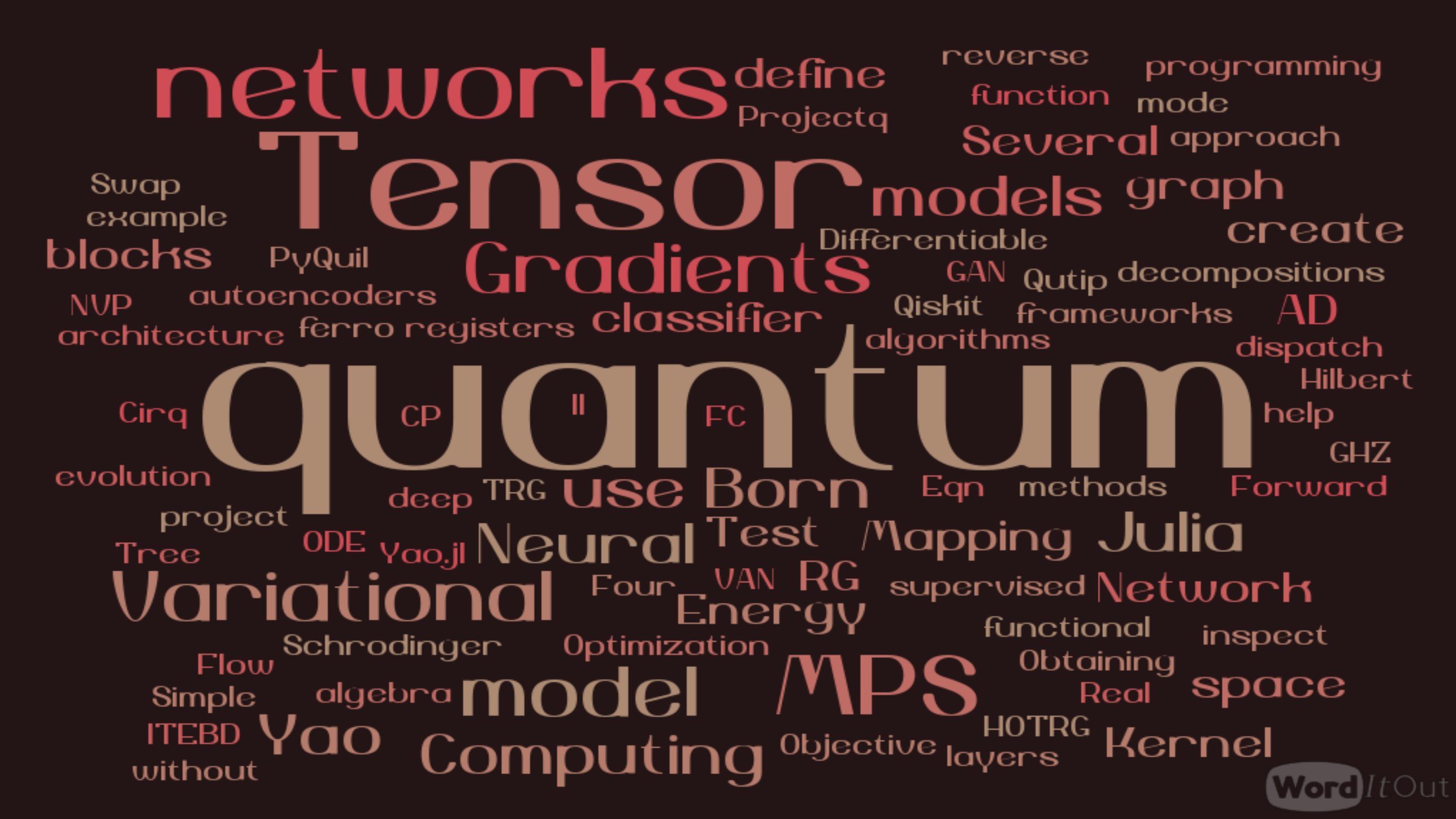
Statistical Physics

Spin Glasses/Complex Networks



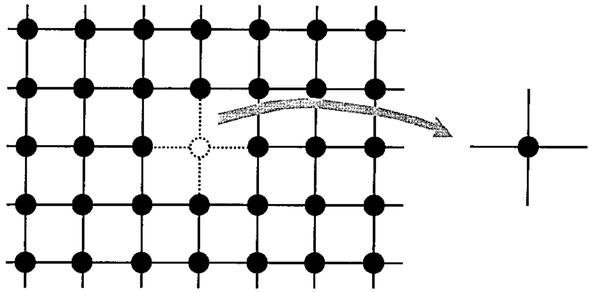
Soft Matter/Biophysics



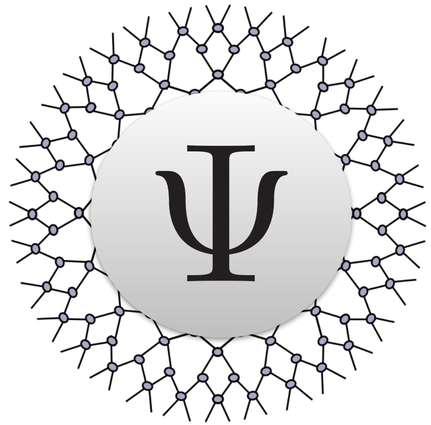

networks
define
reverse
programming
function
mode
Projectq
Several
approach
Tensor
models
graph
Swap
example
create
Differentiable
blocks
PyQuil
Gradients
GAN
Qutip
decompositions
NVP
autoencoders
Qiskit
frameworks
AD
classifier
architecture
ferro registers
algorithms
dispatch
quantum
Hilbert
Cirq
CP
II
FC
help
GHZ
evolution
deep
TRG
use
Born
Eqn
methods
Forward
project
Test
Mapping
Julia
Tree
ODE
Yao.jl
Neural
Variational
Four
VAN
RG
supervised
Network
Energy
functional
inspect
Schrodinger
Optimization
Obtaining
Flow
MPS
space
Simple
algebra
model
Real
ITEBD
Yao
HOTRG
Kernel
Computing
Objective
layers
without
WordItOut

Physicists' gifts to Machine Learning

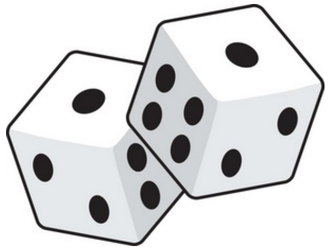
Mean Field Theory



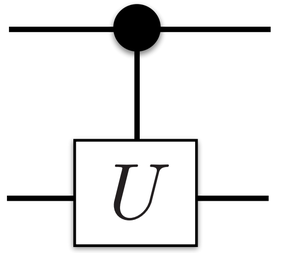
Tensor Networks



Monte Carlo Methods

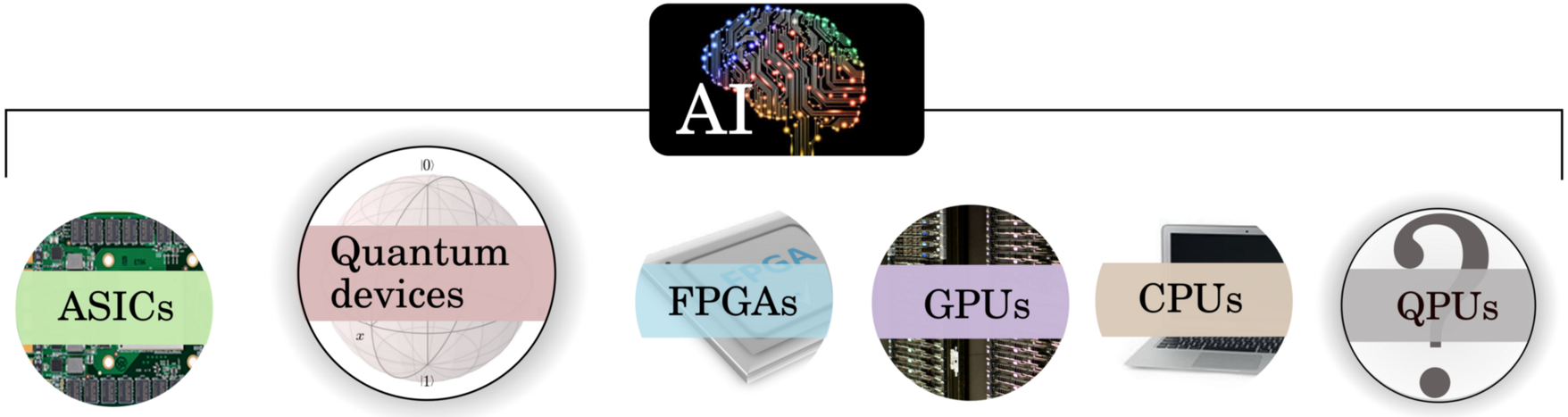


Quantum Computing

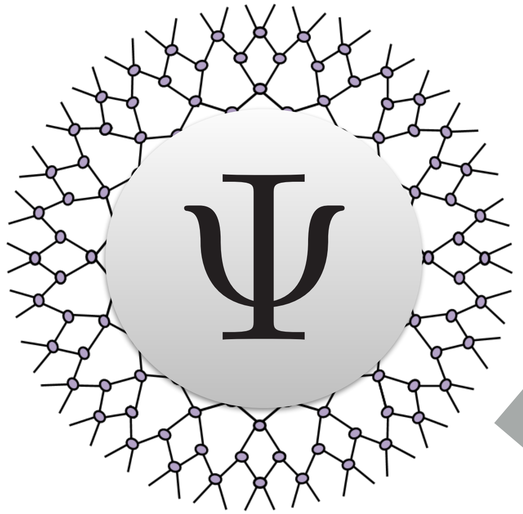


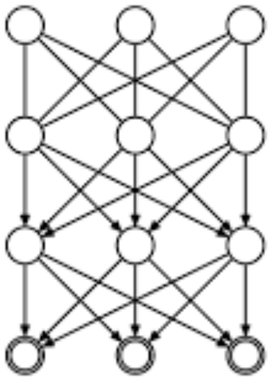
Physicists' gifts to Machine Learning

- Tensor network is 21 century's matrix (Mario Szegedy)
- Started as a technique for statistical and quantum many-body physics. Now, penetrating into machine learning.
- Is QPU the next generation coprocessor that powers AI ?

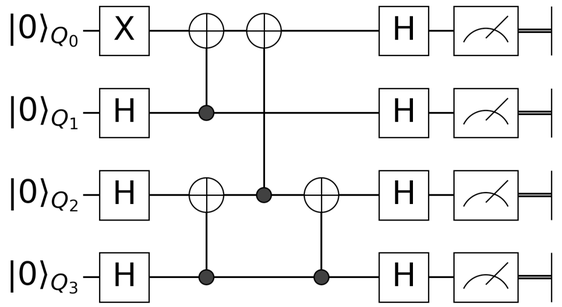


Tensor network as a bridge



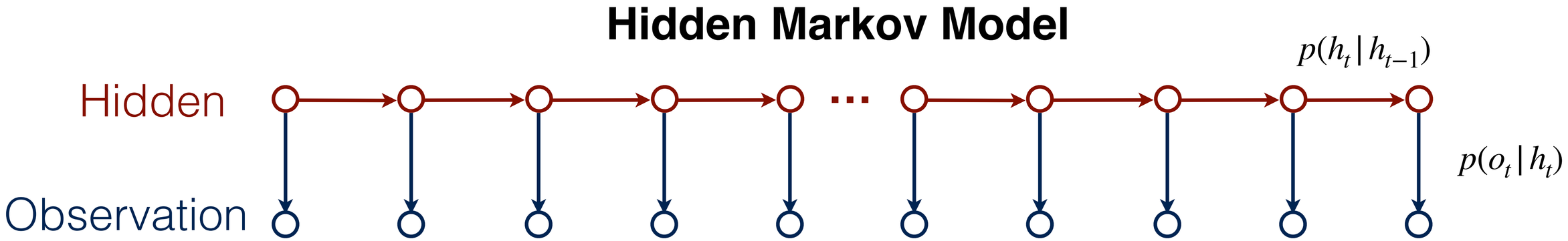


Neural networks and Probabilistic graphical models



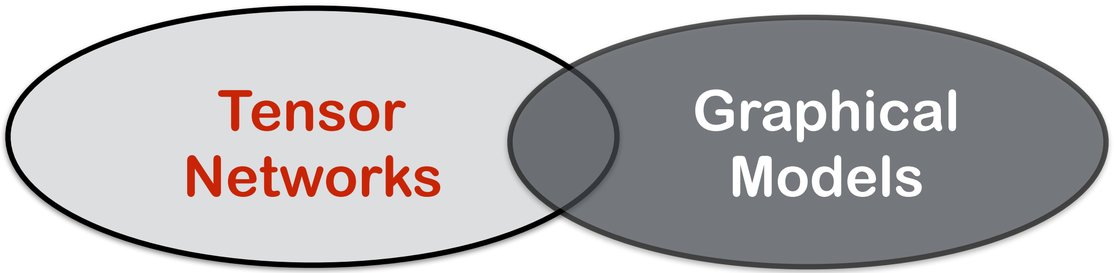
Quantum circuits architecture and parametrization

Tensor network for machine learning



- Widely used in speech recognition, bioinformatics, cryptography…
- Learning, inference, and sampling algorithms (Forward-backward message passing, Viterbi, Baum-Welch) => **Think of positive MPS**

Temme, Verstraete,1003.2545
Critch, Morton, 1210.2812



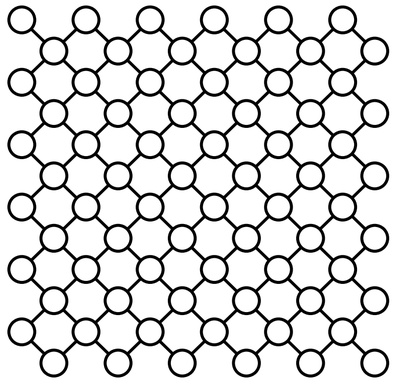
Zhao, Jaeger, Neural Comput. 2010 Bailly, J. Mach. Learn. Res. 2011 Robeva, Seigal, 1710.01437 Chen, Cheng, Xie, LW, Xiang, 1701.04831

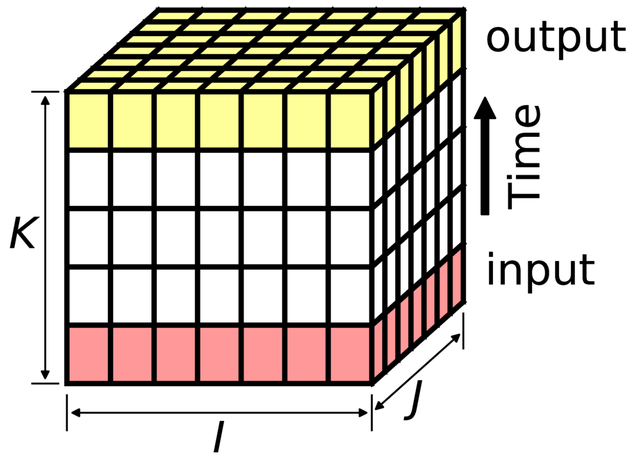
Tensor network for quantum computing

Simulation and validation of quantum circuits

Villalonga et al, 1811.09599

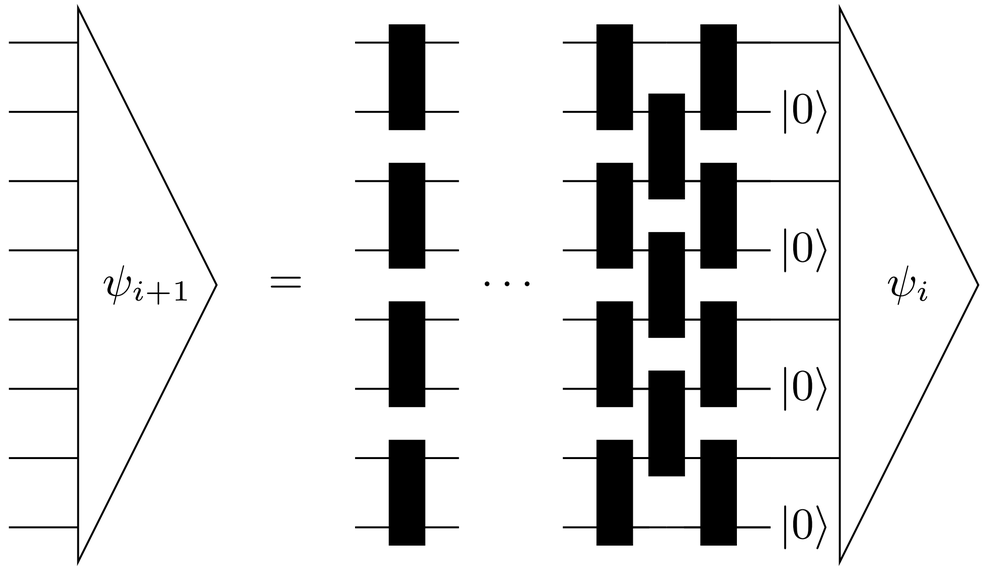
Bristlecone-72





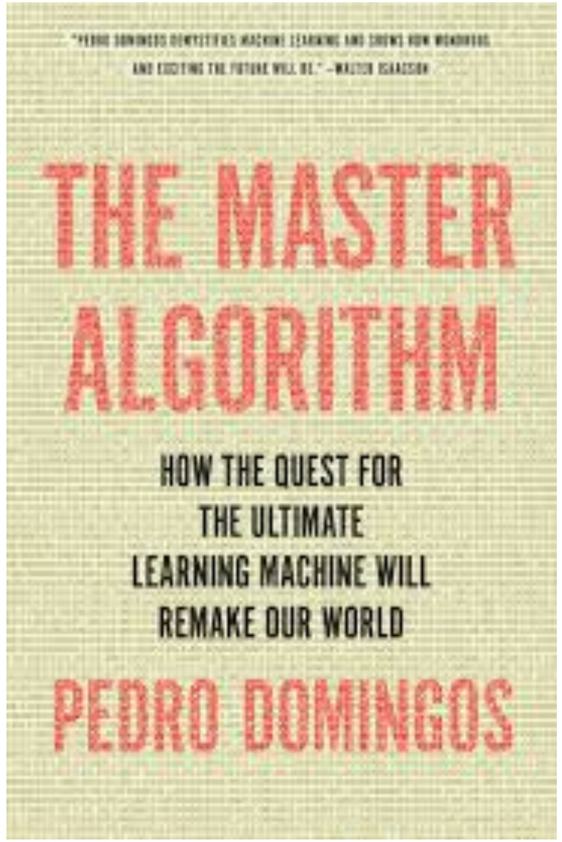
Noise-resilient quantum circuits with TNS architecture

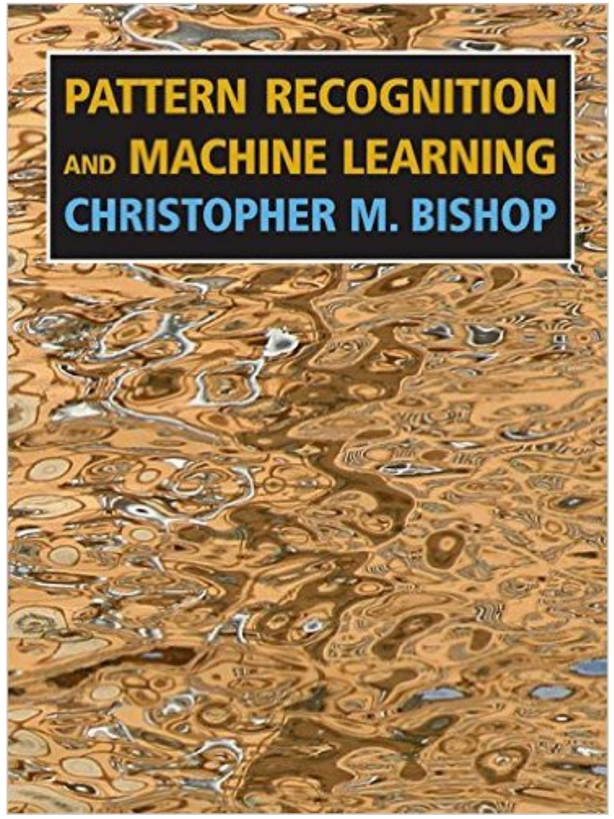
Kim and Swingle, 1711.07500

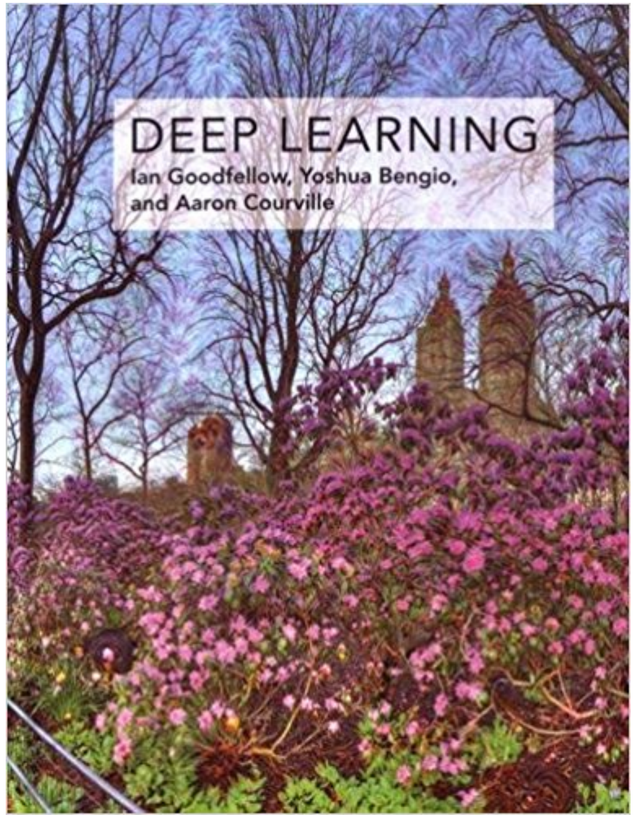


+ quantum tomography, quantum annealing, quantum error correction, holographic duality, and more

Introductory Books







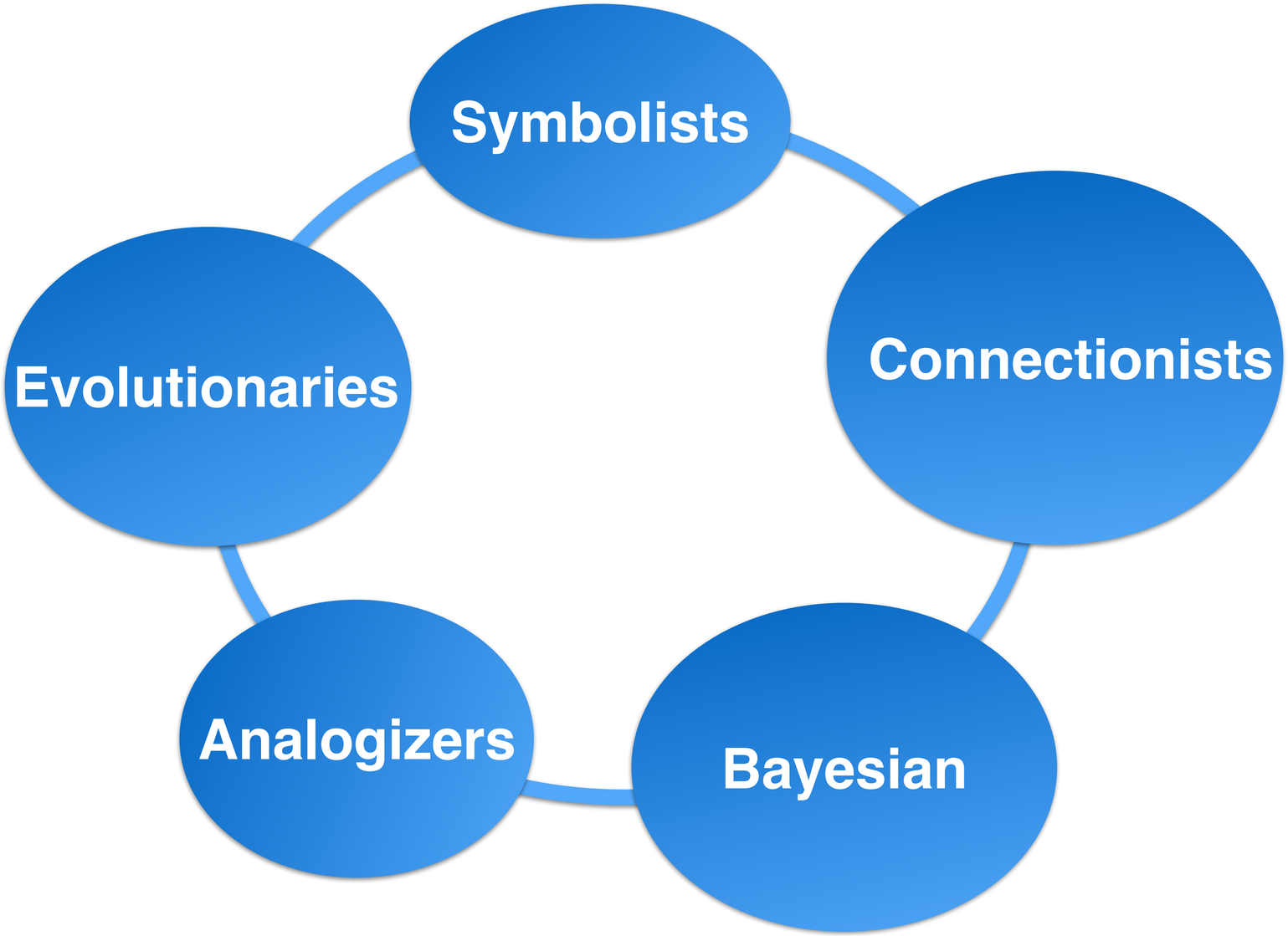
Popular overview

Solid math

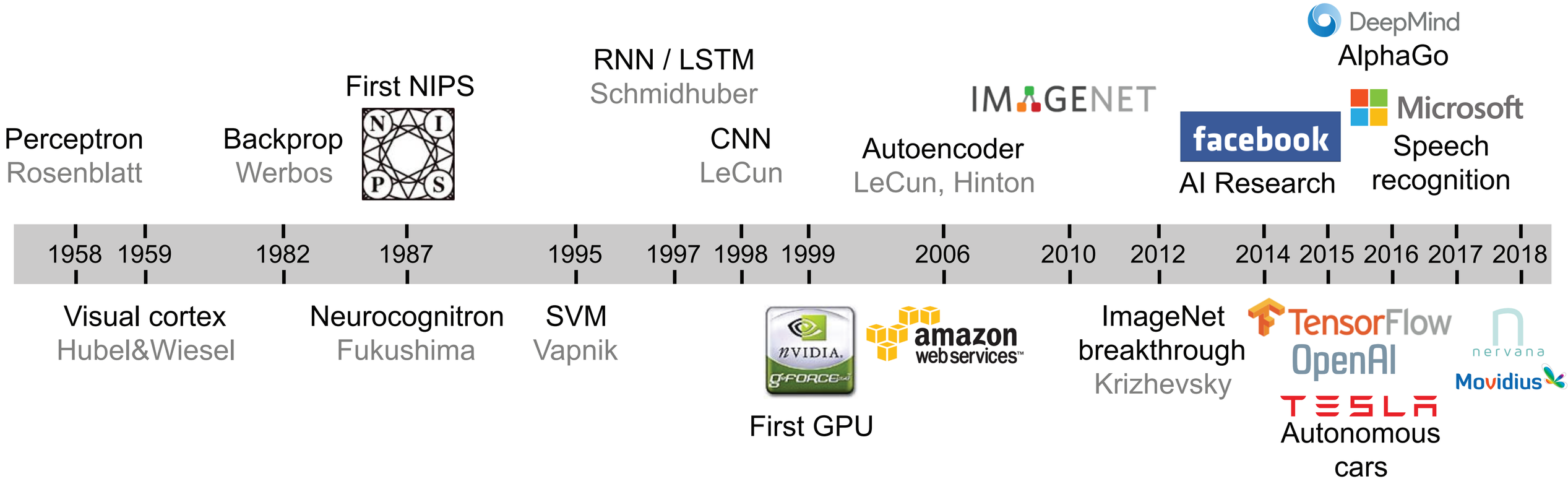
“Modern” topics

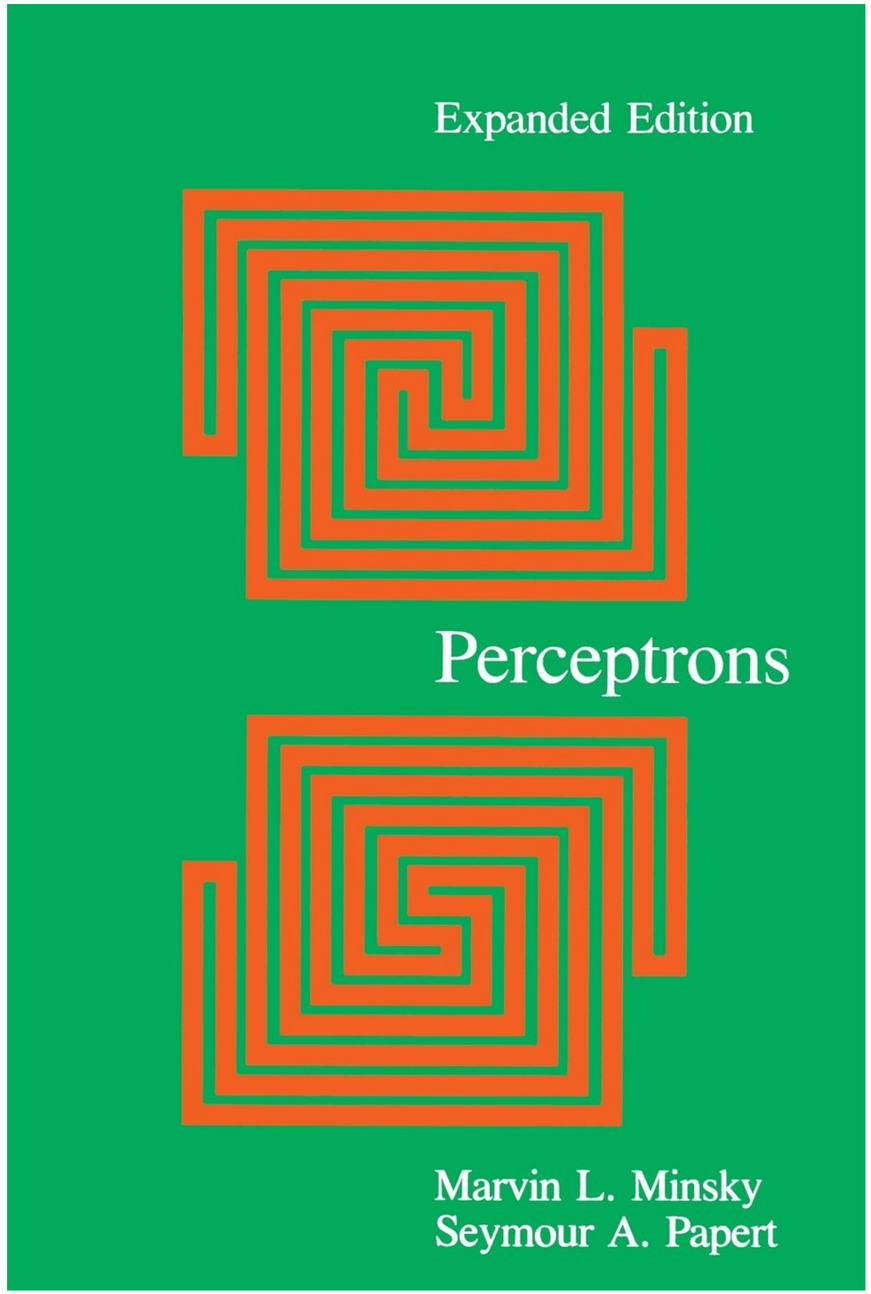
+ many excellent online courses, see the GitHub repo for a collection

Five schools of ML



A brief history





1st edition 1969
expanded 1972
commented 1988

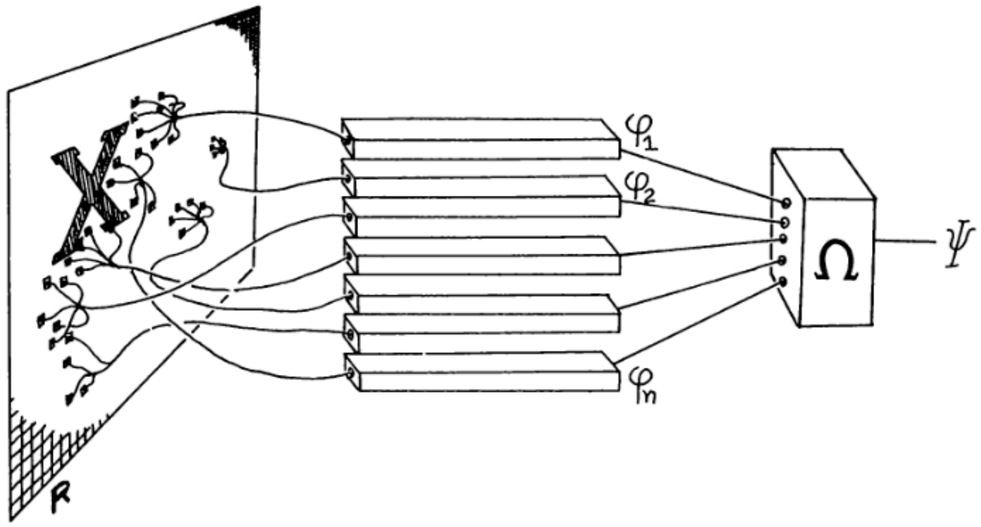
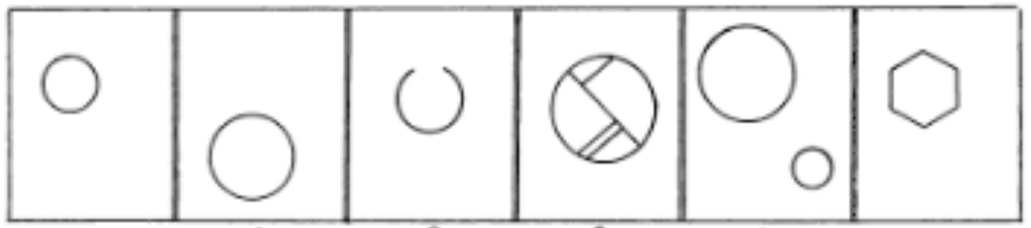
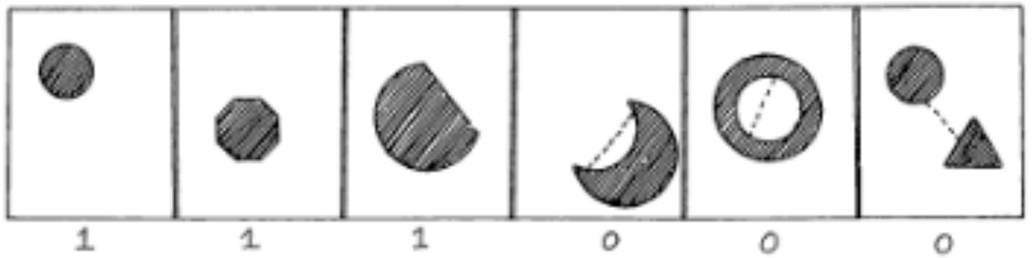


Figure 0.1

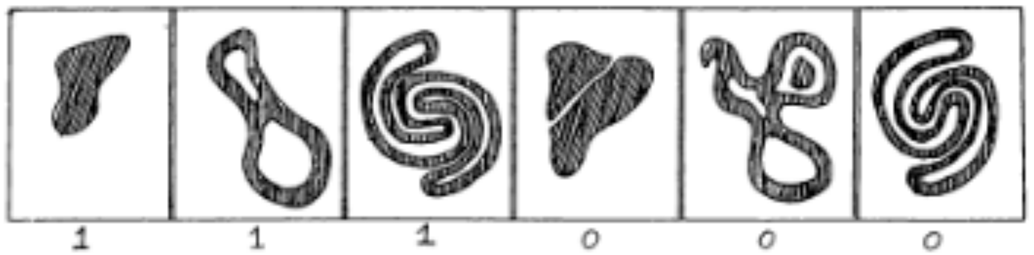
$$\psi_{\text{CIRCLE}}(X) = \begin{cases} 1 \text{ if the figure } X \text{ is a circle,} \\ 0 \text{ if the figure is not a circle;} \end{cases}$$

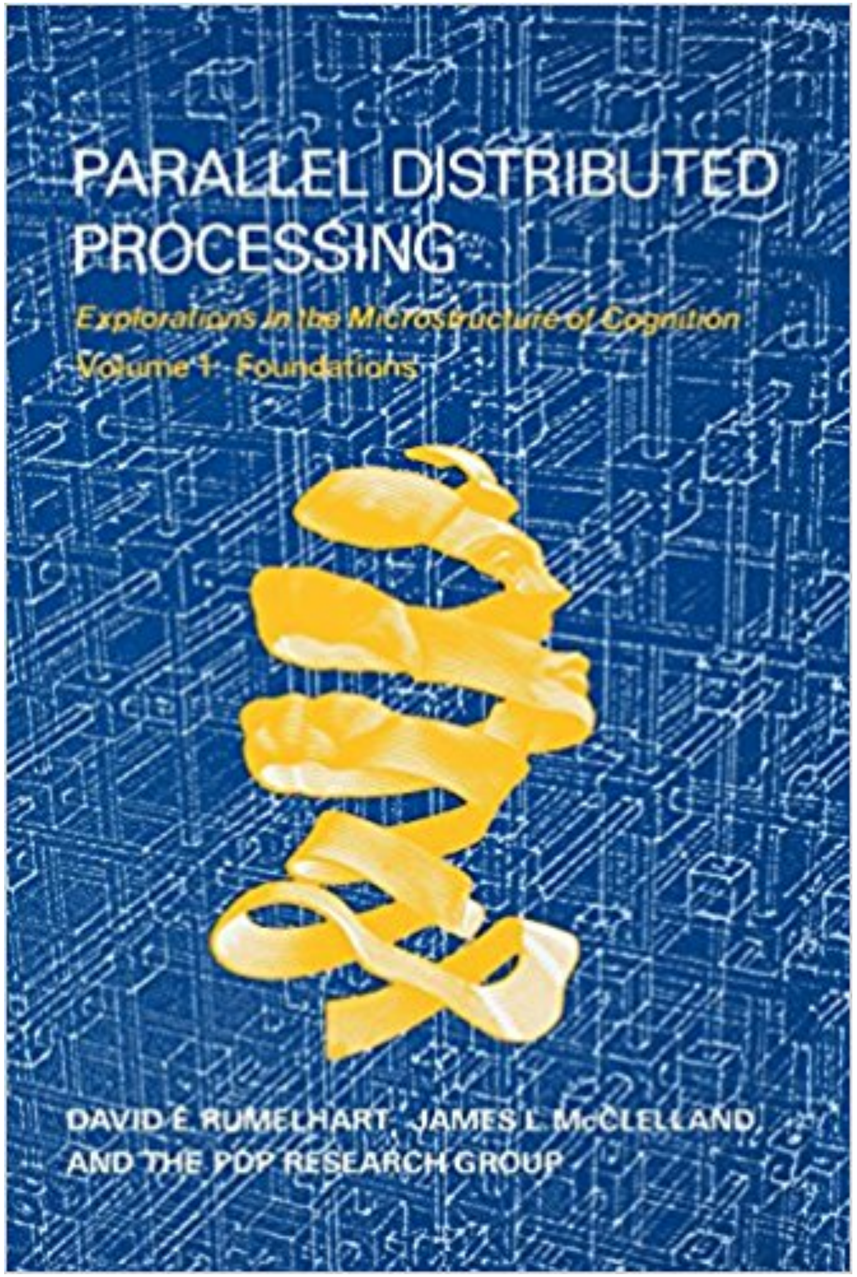


$$\psi_{\text{CONVEX}}(X) = \begin{cases} 1 \text{ if } X \text{ is a convex figure,} \\ 0 \text{ if } X \text{ is not a convex figure;} \end{cases}$$



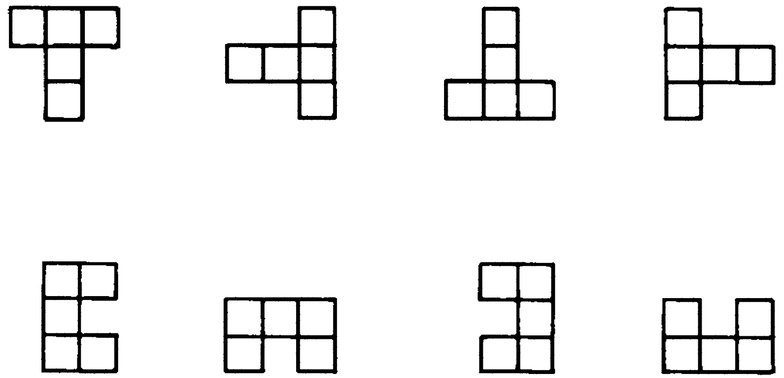
$$\psi_{\text{CONNECTED}}(X) = \begin{cases} 1 \text{ if } X \text{ is a connected figure,} \\ 0 \text{ otherwise.} \end{cases}$$





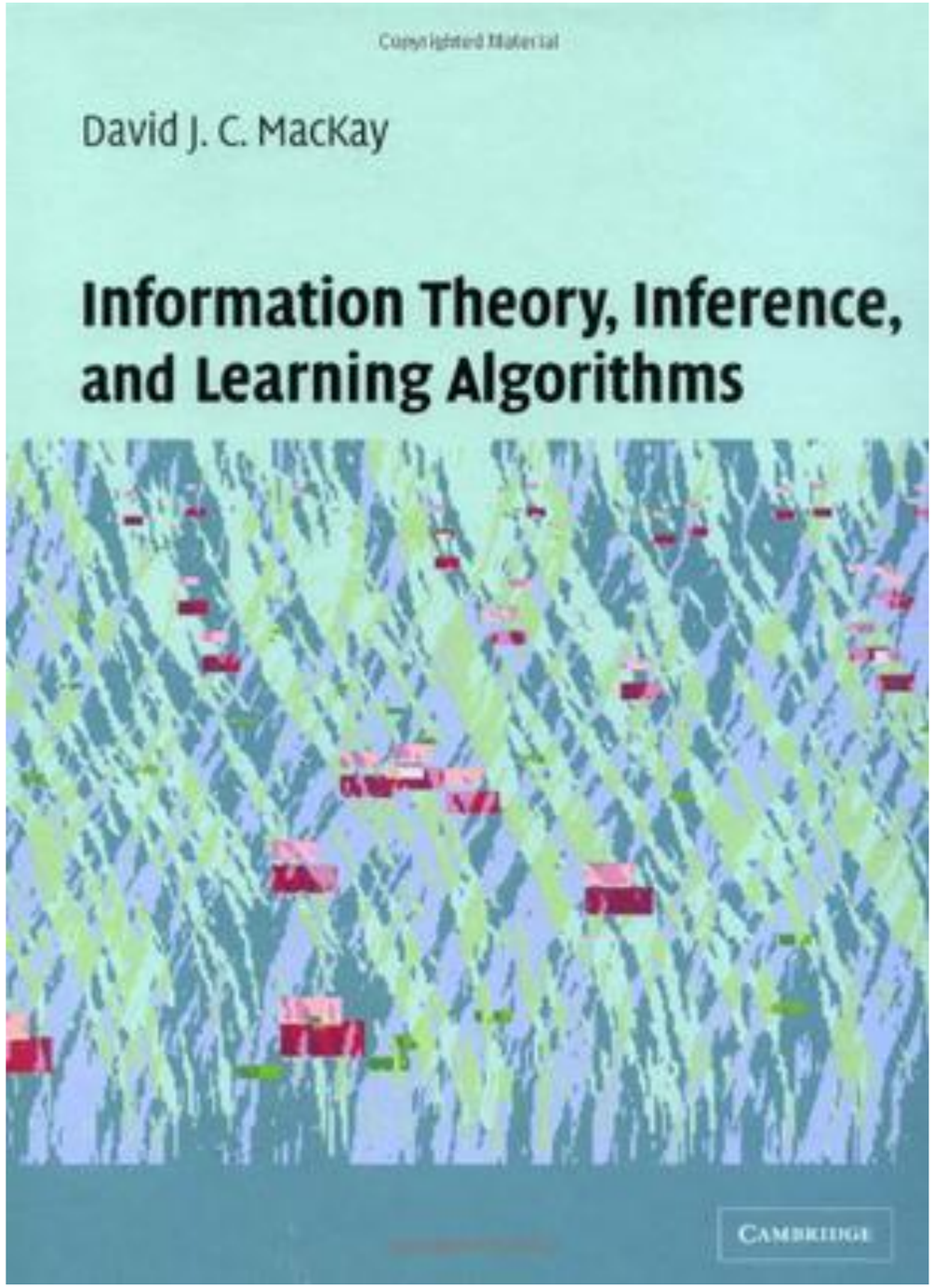
“Manifesto of Connectionism”

1988

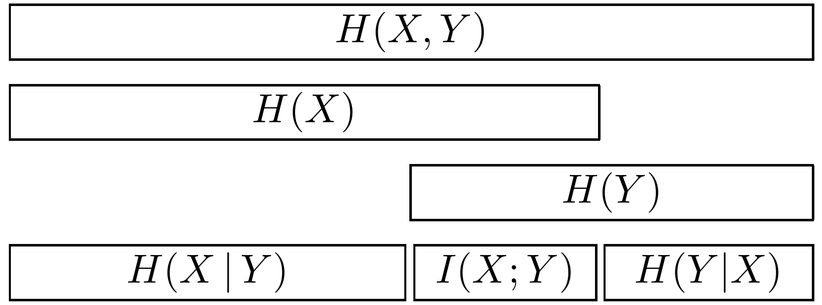


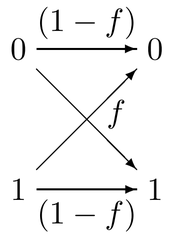
“T-C Problem”

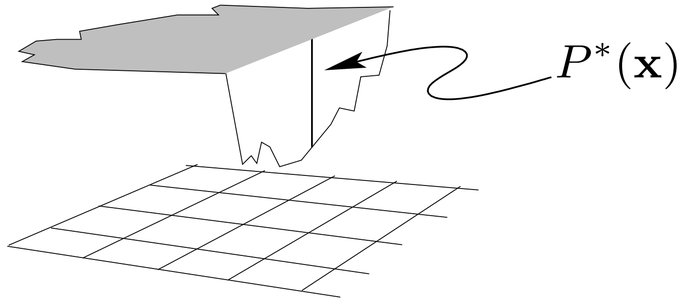
Contains many interesting experiments & theories on toy problems and foundational thoughts by the pioneers



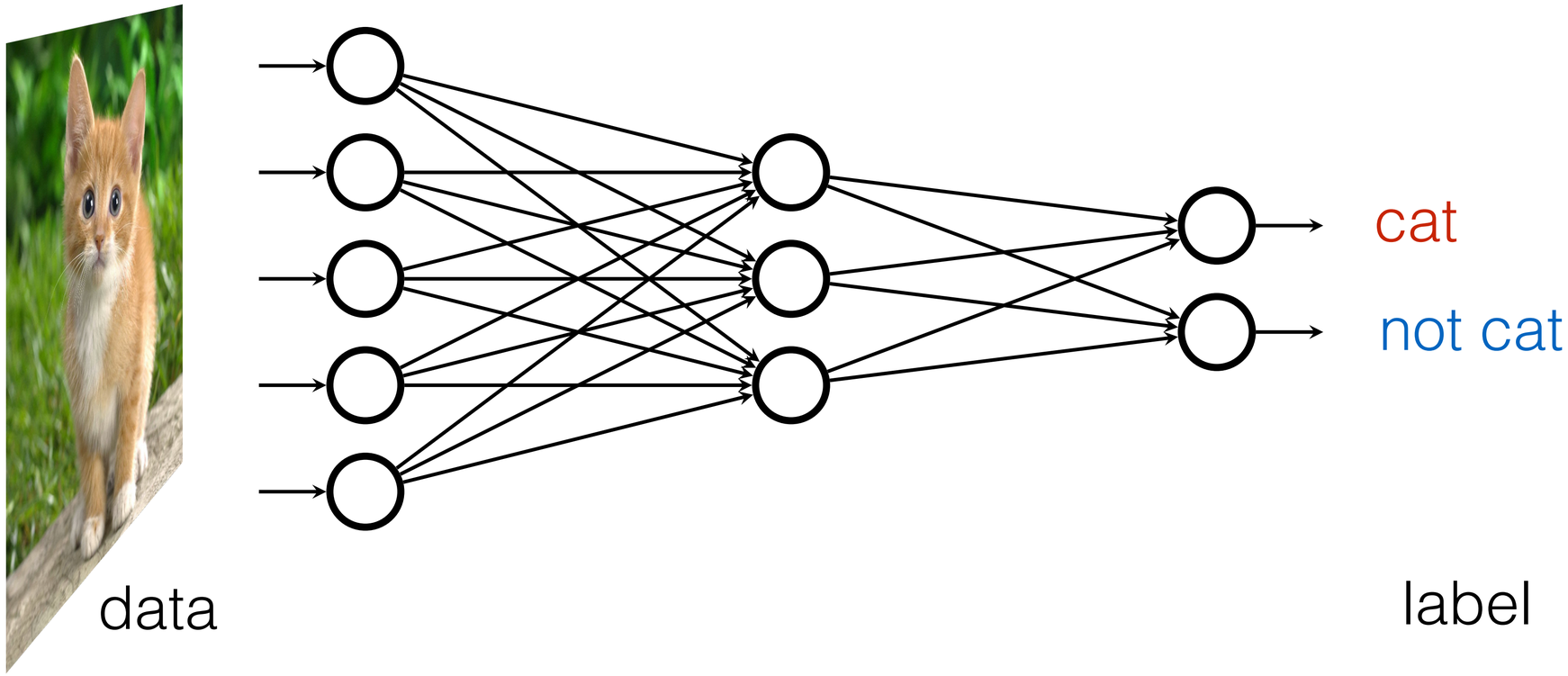
Insightful, fun to read



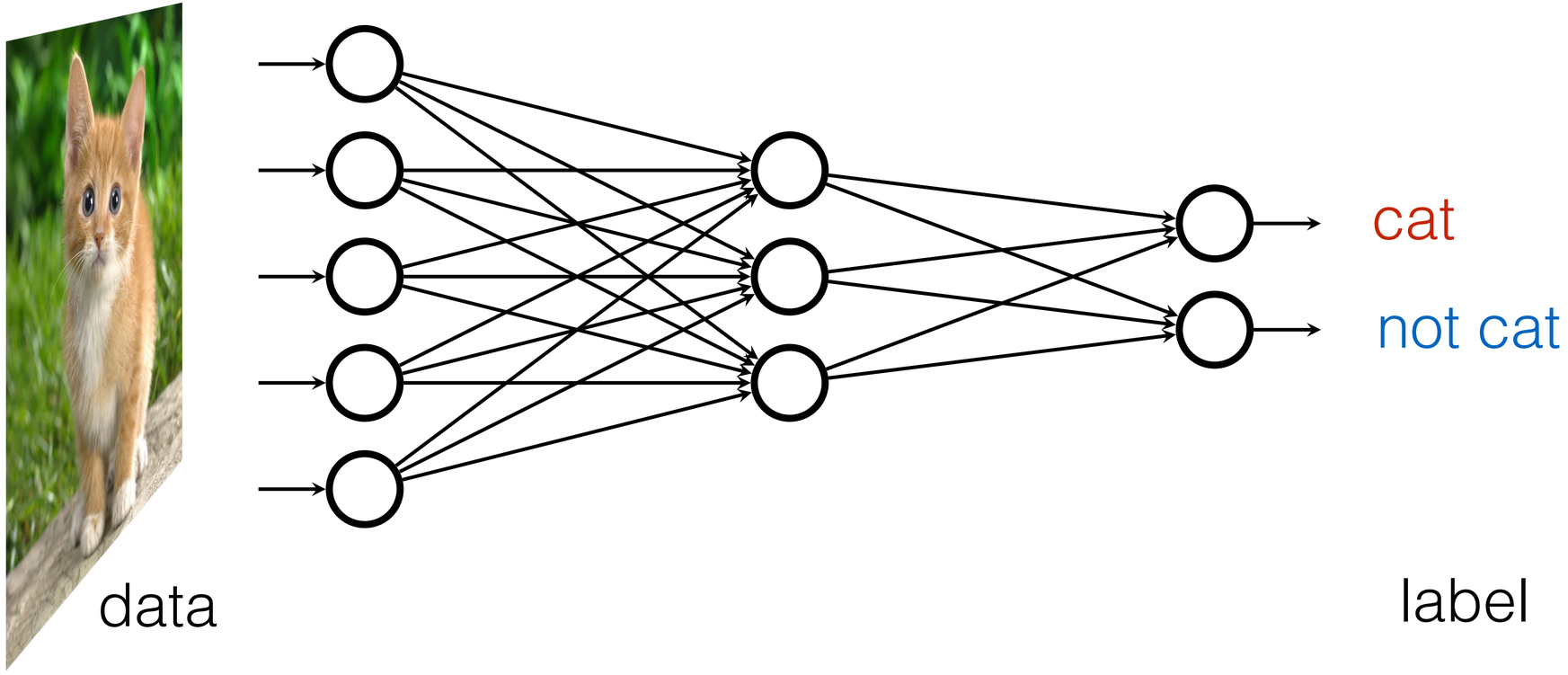




Pattern recognition and beyond



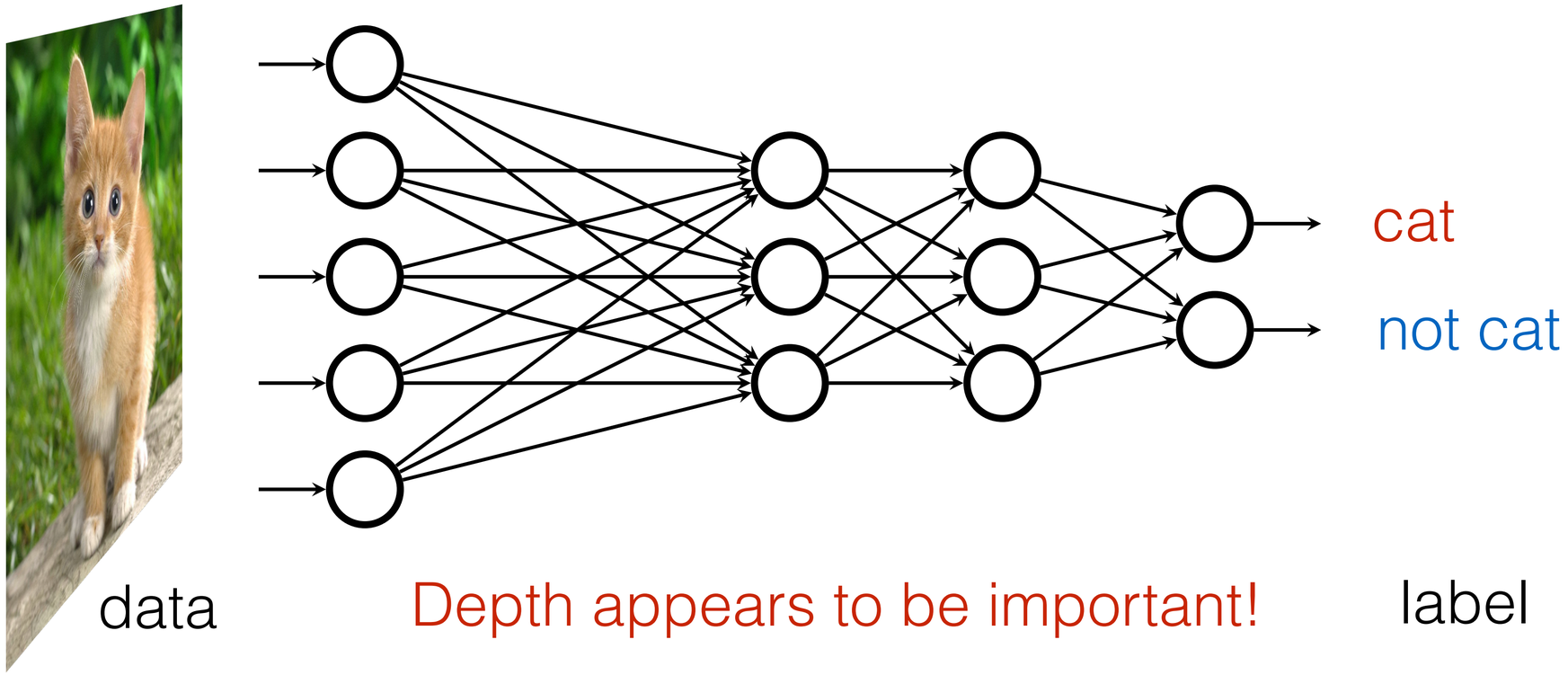
Pattern recognition and beyond



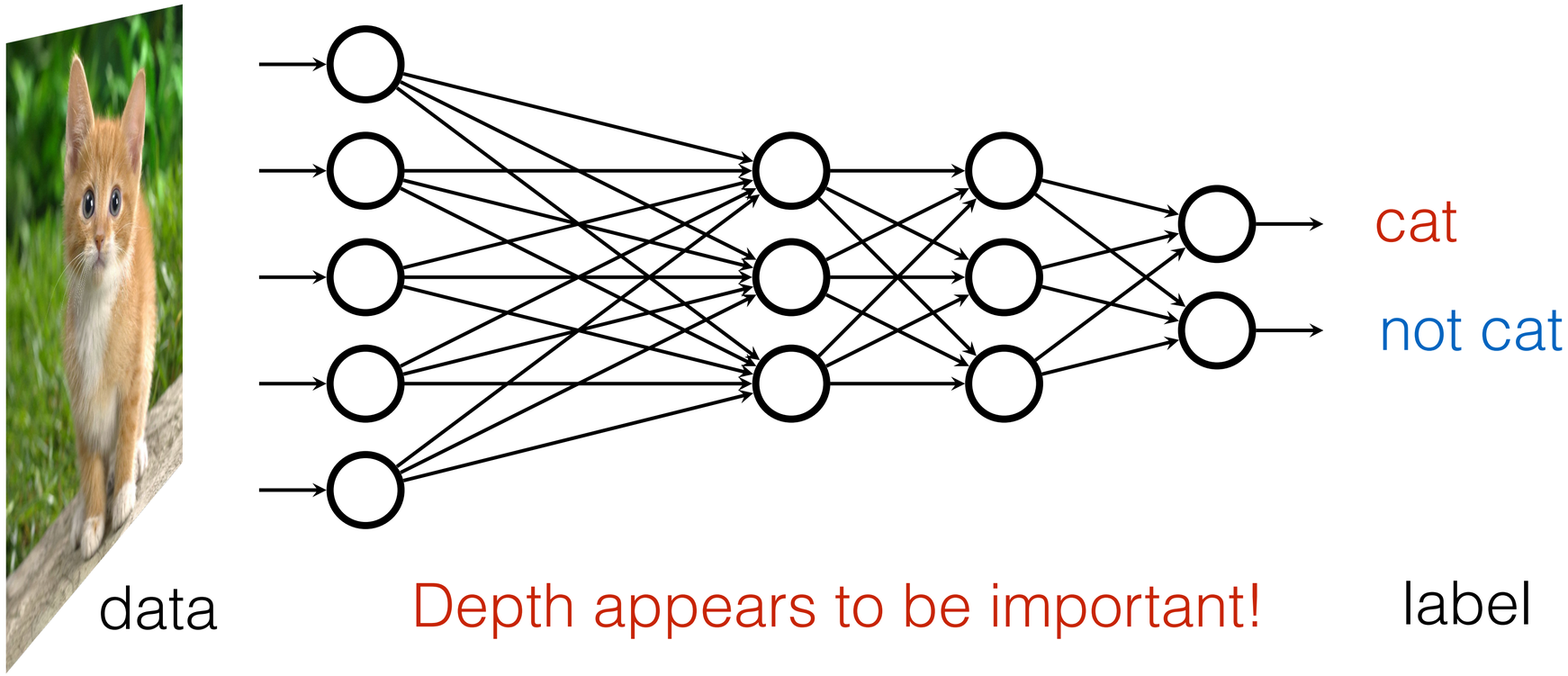
Universal Function Approximator

Cybenko 1989
Hornik, Stinchcombe, White 1989

Pattern recognition and beyond

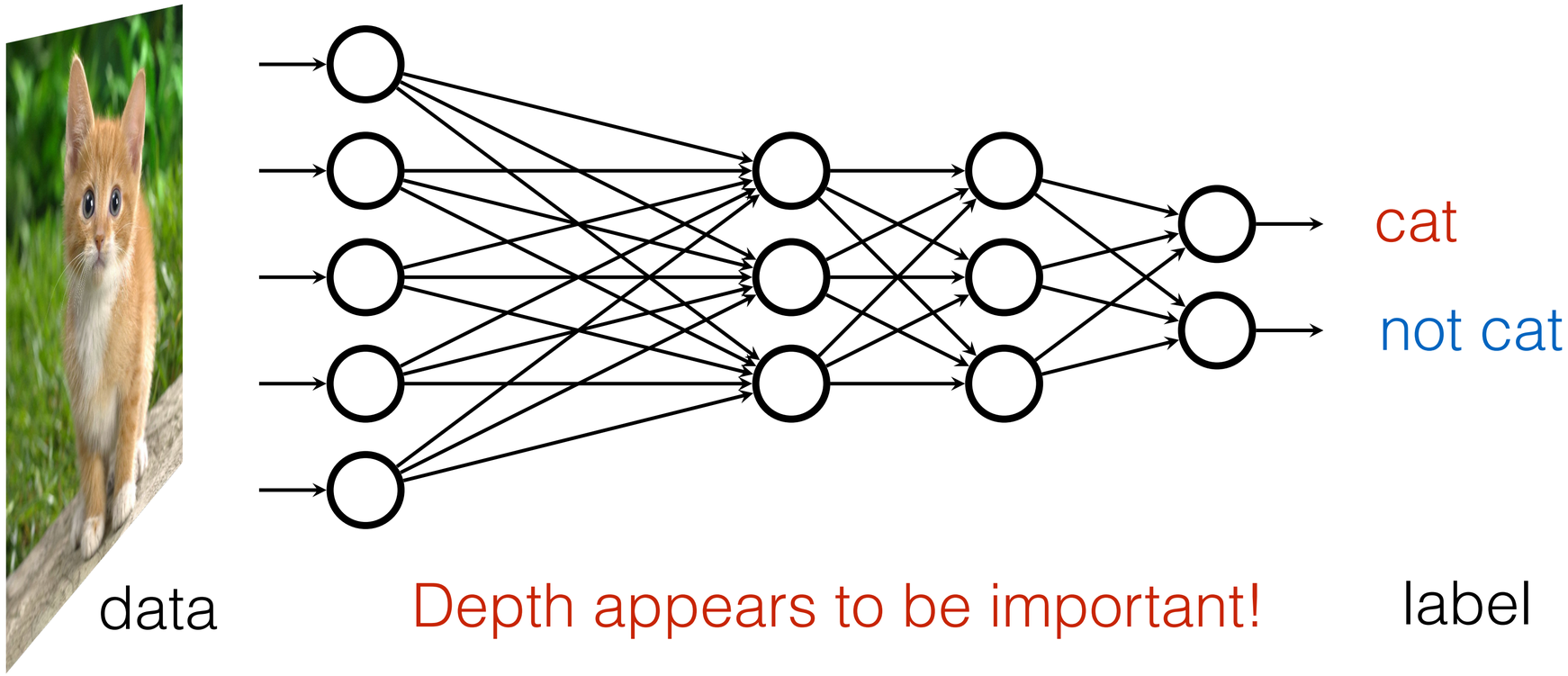


Pattern recognition and beyond



Q: Why does deep learning work?

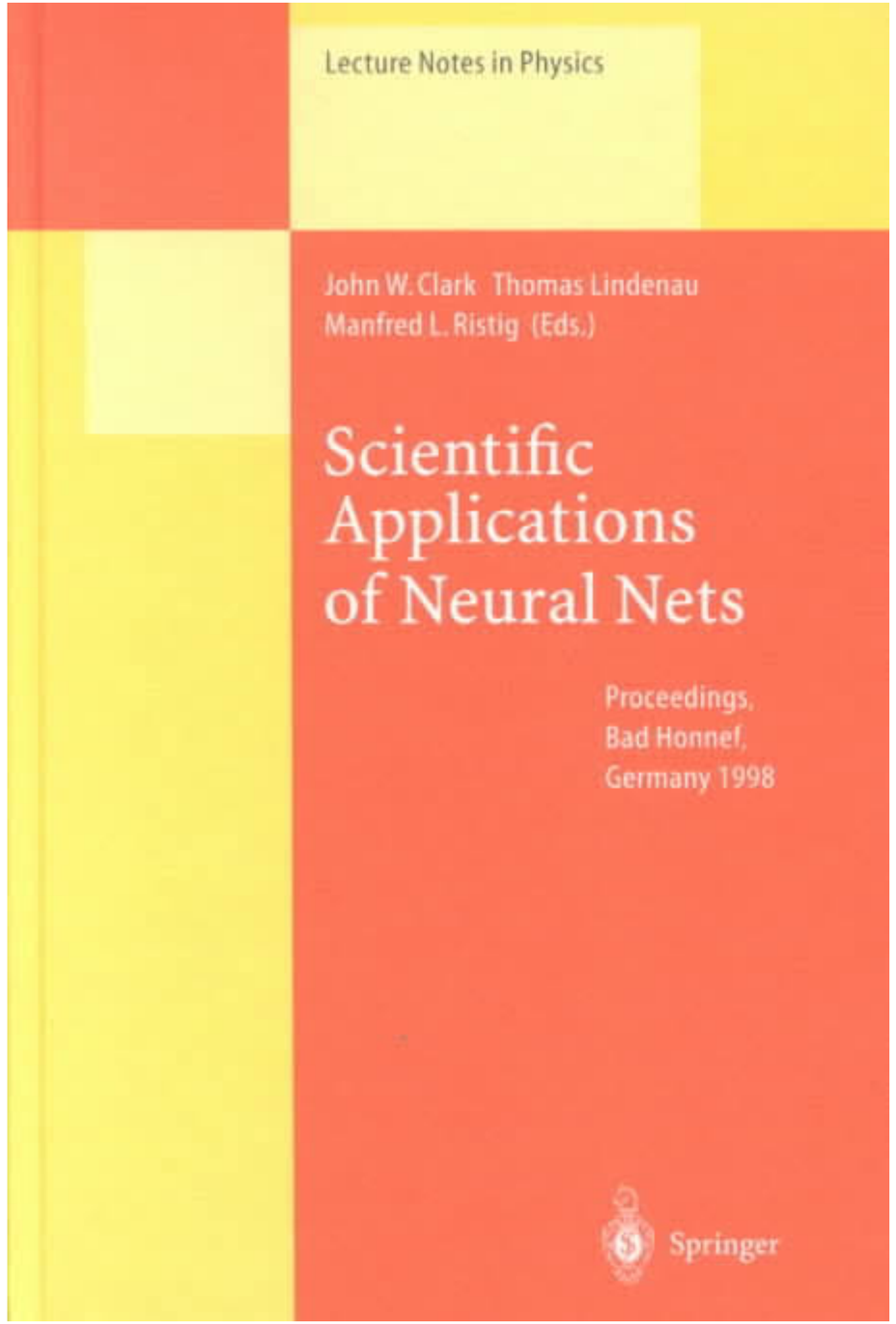
Pattern recognition and beyond



Q: Why does deep learning work?

A: Law of physics: symmetry, locality, compositionality, renormalization group, and quantum entanglement.

Lin, Tegmark, Rolnick ,1608.08225 Mehta, Schwab, 1410.3831 Levine et al, 1704.01552 …



Gem in between this and last hype cycles

8 Doing Science With Neural Nets: Pride and Prejudice

When neural networks re-emerged on the scene in the mid-80s as a new and glamorous computational paradigm, the initial reaction in some sectors of the scientific community was perhaps too enthusiastic and not sufficiently critical. There was a tendency on the part of practitioners to oversell the powers of neural-network or "connectionist" solutions relative to conventional techniques – where conventional techniques can include both traditional theory-rich modeling and established statistical methods. The last five years have seen a correction phase, as some of the practical limitations of neural-network approaches have become apparent, and as scientists have become better acquainted with the wide array of advanced statistical tools that are currently available.

Why now, again ?
What has changed ?
What has not ?

Deep learning is more than fitting functions

Discriminative learning

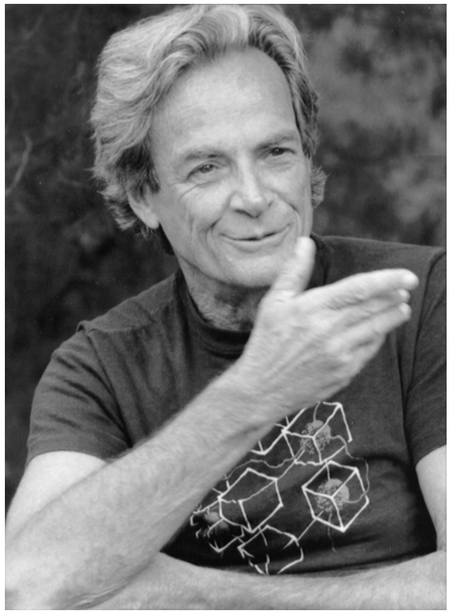
$$y = f(\boldsymbol{x})$$

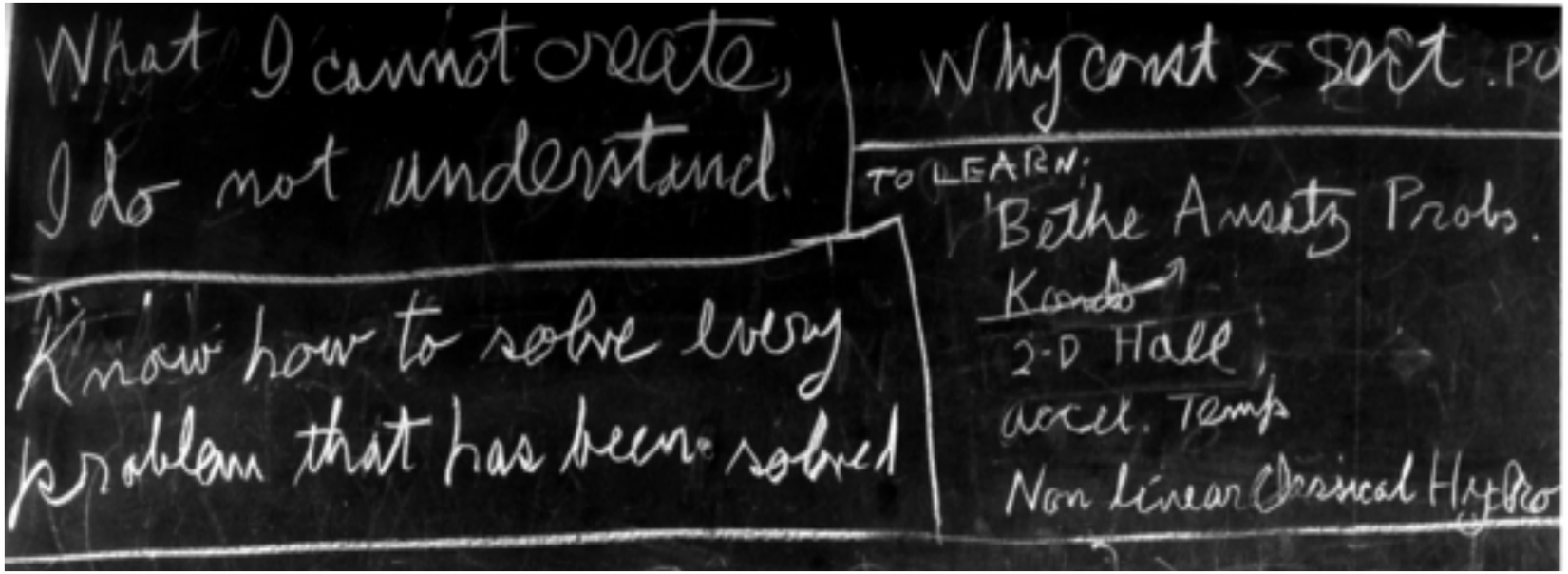
or $p(y \mid \boldsymbol{x})$

Generative learning

$$p(\boldsymbol{x})$$

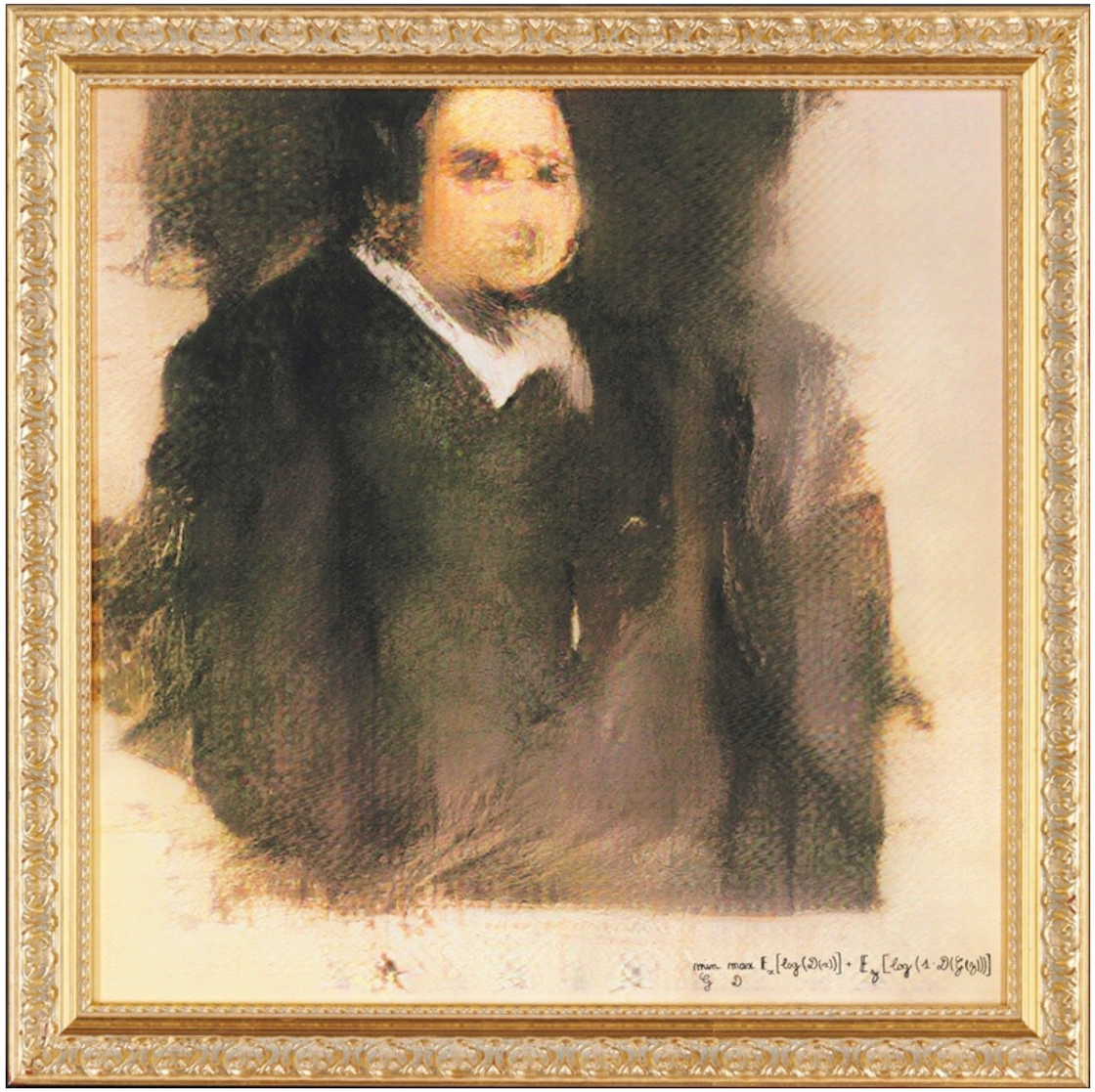
Deep learning is more than fitting functions





"What I can not create, I do not understand"

Generated Arts

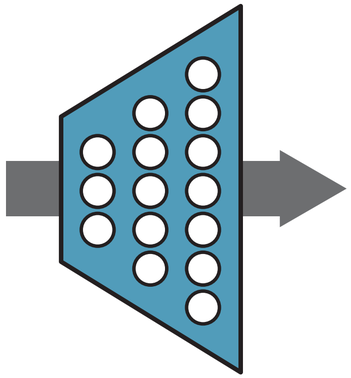


$432,500
25 October 2018
Christie's New York

https://www.christies.com/Features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx

Generated Arts

Gaussian Noise



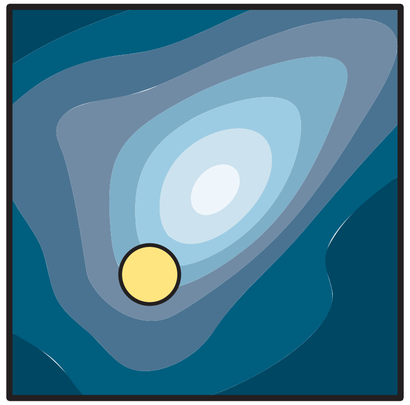
Generative Network

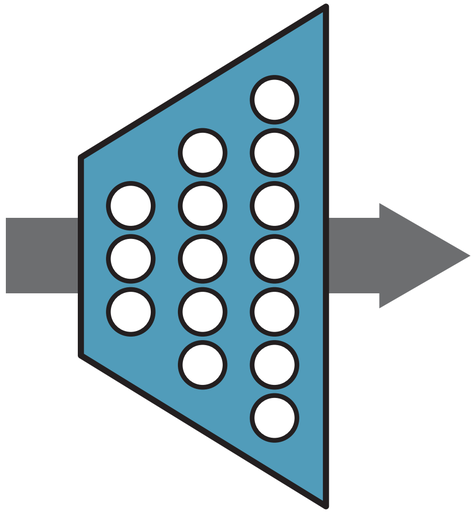
$432,500
25 October 2018
Christie's New York

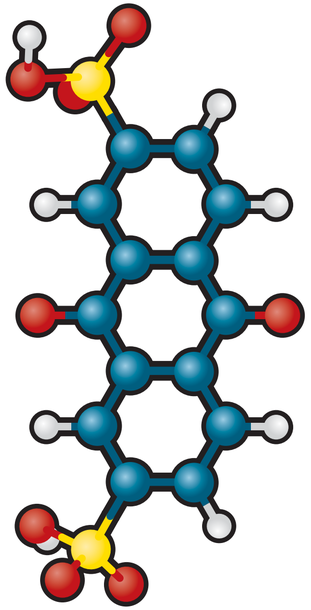
https://www.christies.com/Features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx

Generate Molecules

Latent attributes







Physical configurations

Simple Distributions

Generate

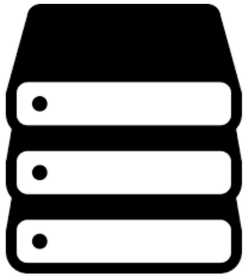
Inference

Complex Distribution

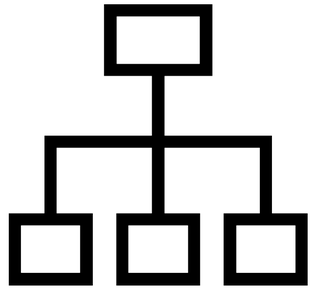
Sanchez-Lengeling & Aspuru-Guzik, Science 2018

Four components of ML

Data



Model



Cost function

Optimization

$$\partial$$

Switch to blackboard

Deep learning tools

HIPS/autograd