第四届全国统计物理与复杂系统学术会议暨海峡两岸统计物理会议

2017.7.16-19 陕西师范大学

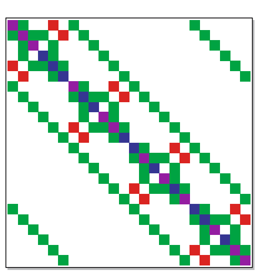
Machine Learning for Many-Body Physics

Lei Wang (王磊)
Institute of Physics, CAS
https://wangleiphy.github.io

About me

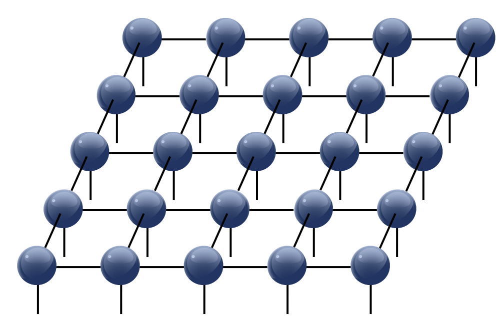
2006	Bachelor	Nanjing University
2011	PhD	IOP, CAS
2016	Postdoctor	ETH Zurich

Computational Quantum Physics

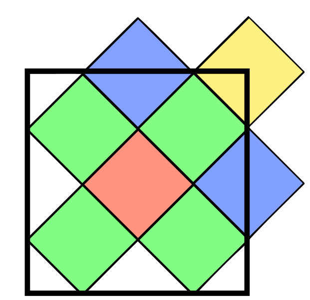


exact diagonalization

quantum Monte Carlo



tensor network states

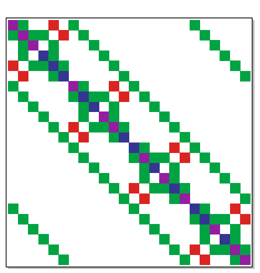


dynamical mean field theories

About me

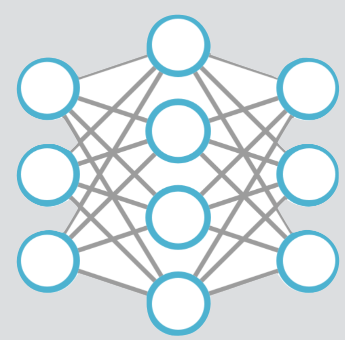
2006	Bachelor	Nanjing University
2011	PhD	IOP, CAS
2016	Postdoctor	ETH Zurich

Computational Quantum Physics

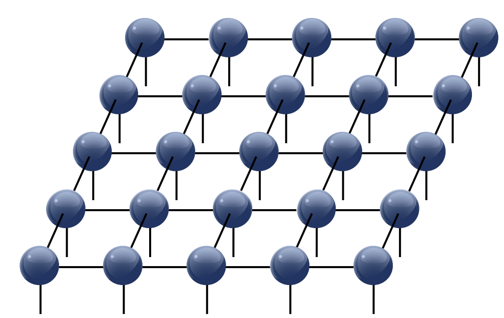


exact diagonalization

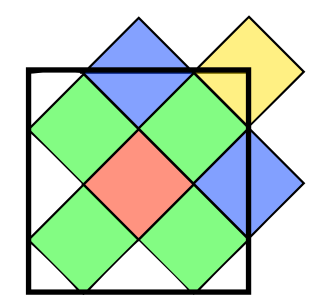
quantum Monte Carlo



machine learning



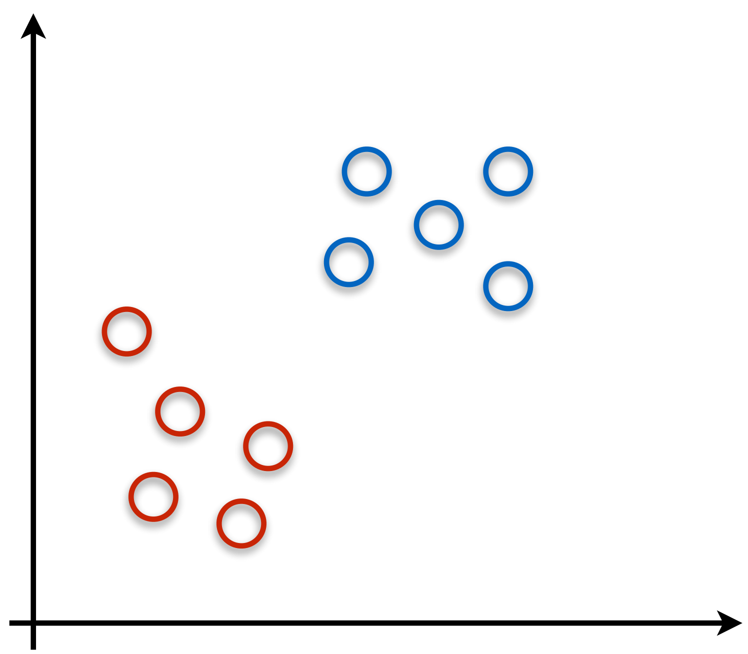
tensor network states



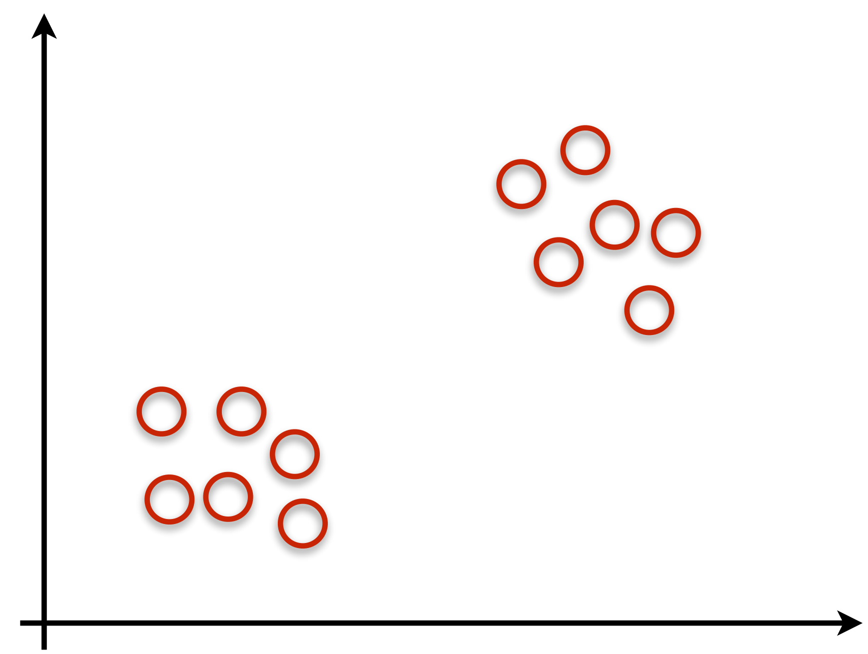
dynamical mean field theories

Machine Learning 101

Supervised learning

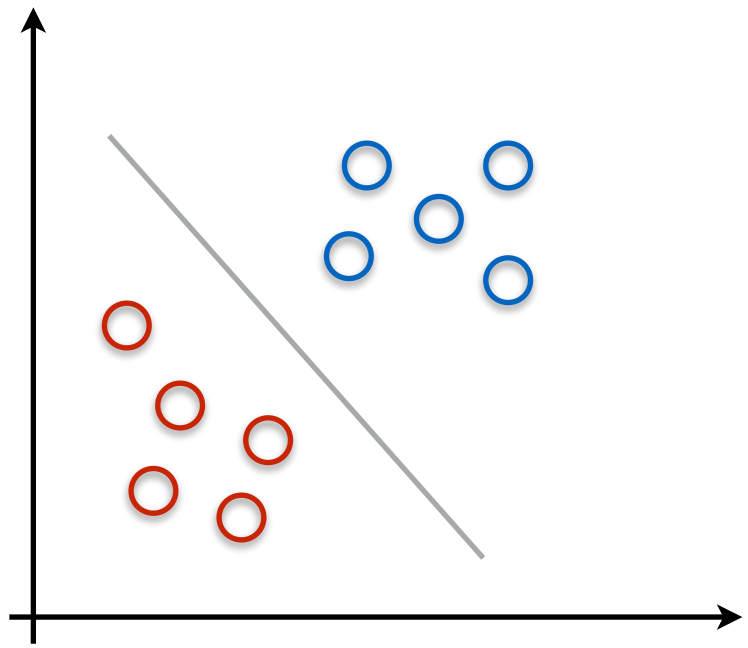


Unsupervised learning

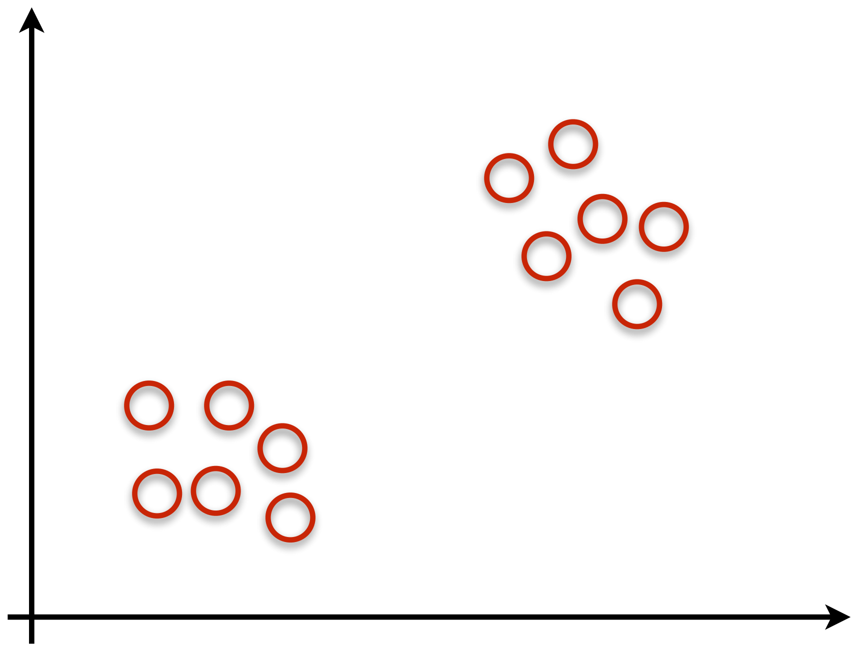


Machine Learning 101

Supervised learning

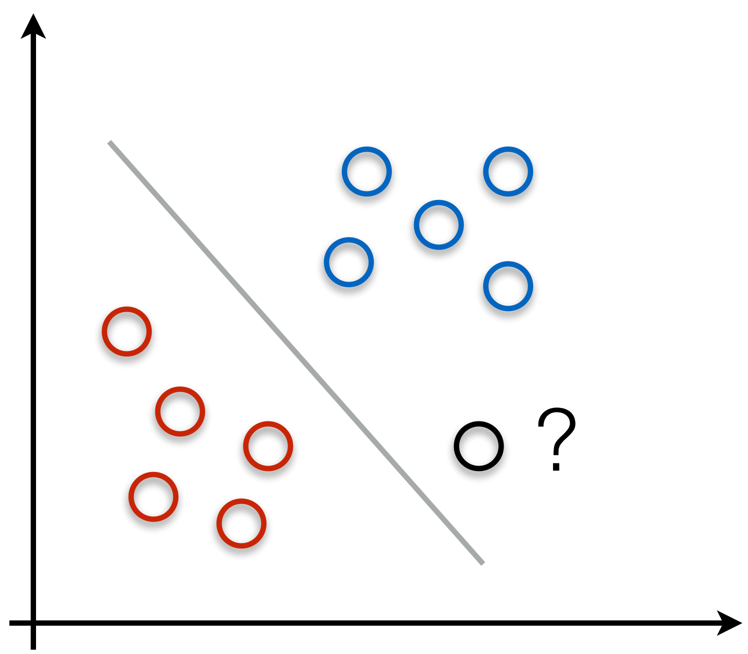


Unsupervised learning

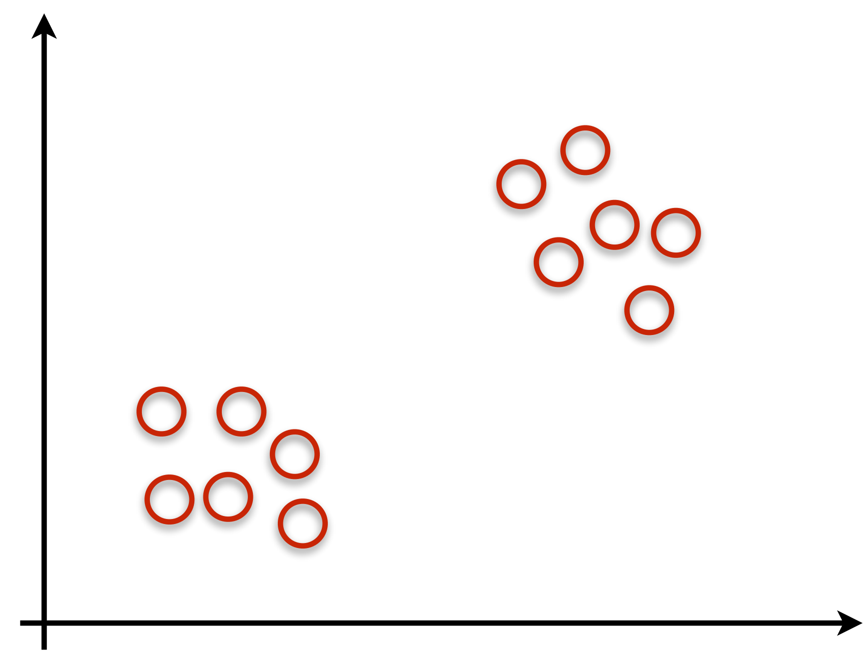


Machine Learning 101

Supervised learning

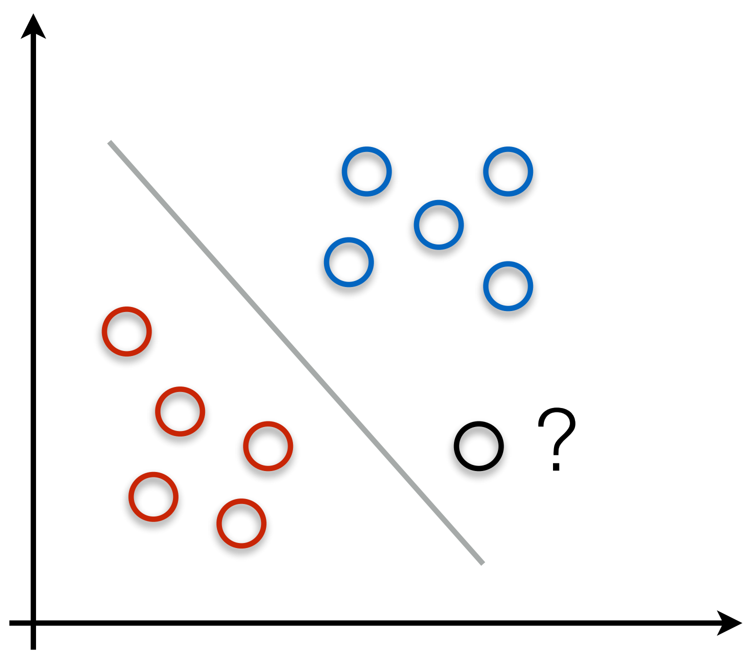


Unsupervised learning



Machine Learning 101

Supervised learning

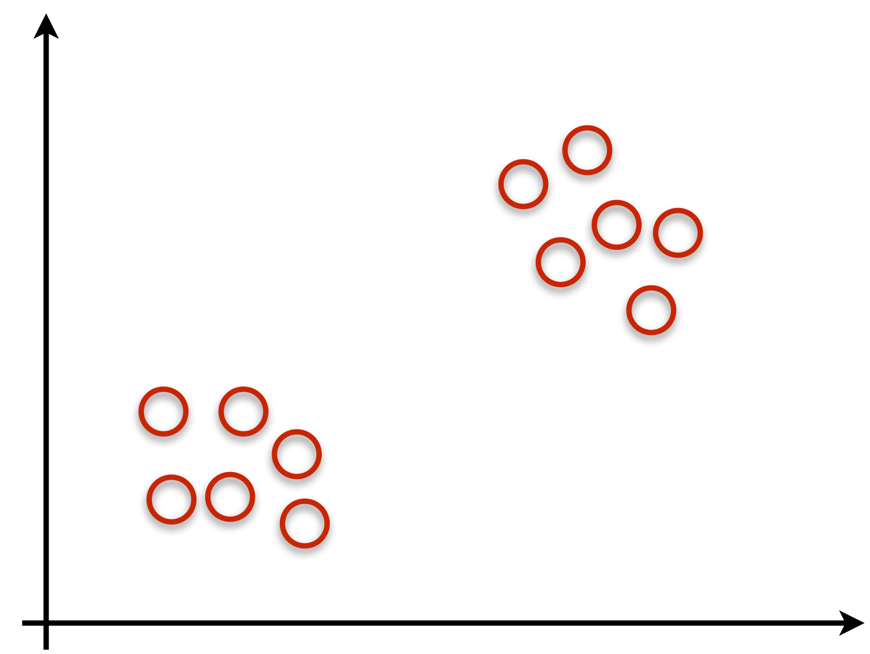


Classification

Spam detection

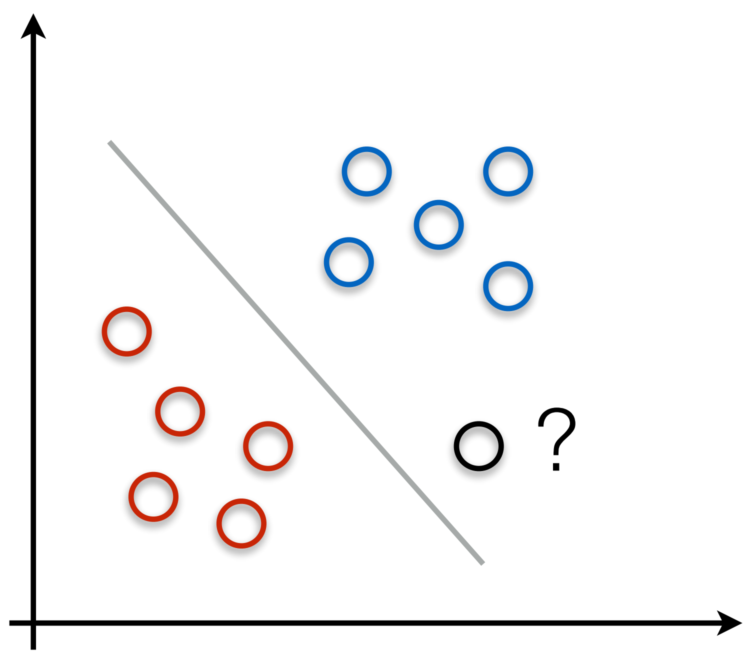
Image recognition

Unsupervised learning



Machine Learning 101

Supervised learning

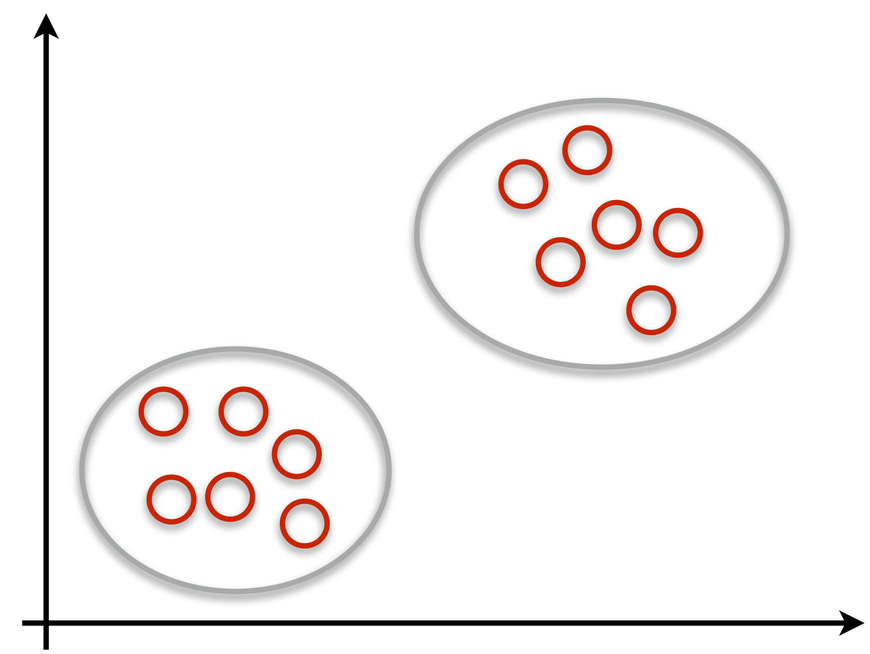


Classification

Spam detection

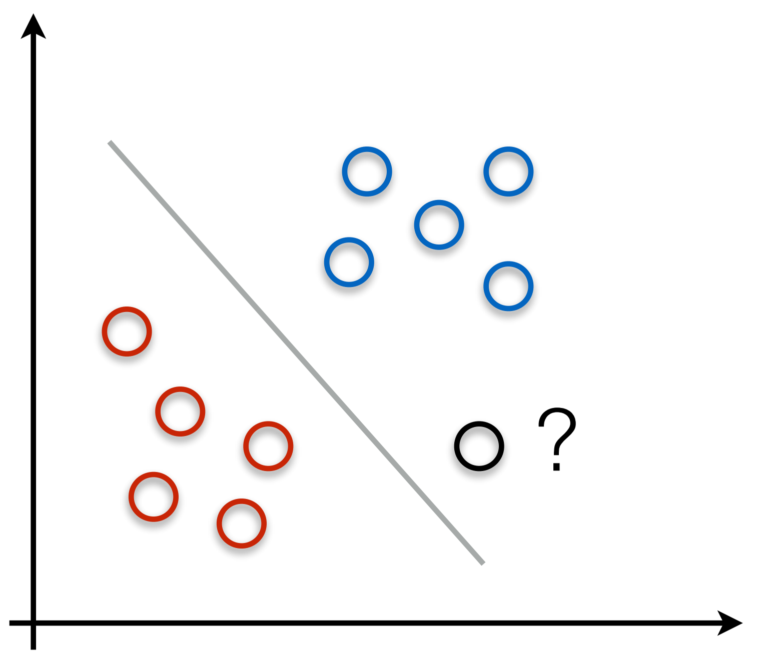
Image recognition

Unsupervised learning



Machine Learning 101

Supervised learning

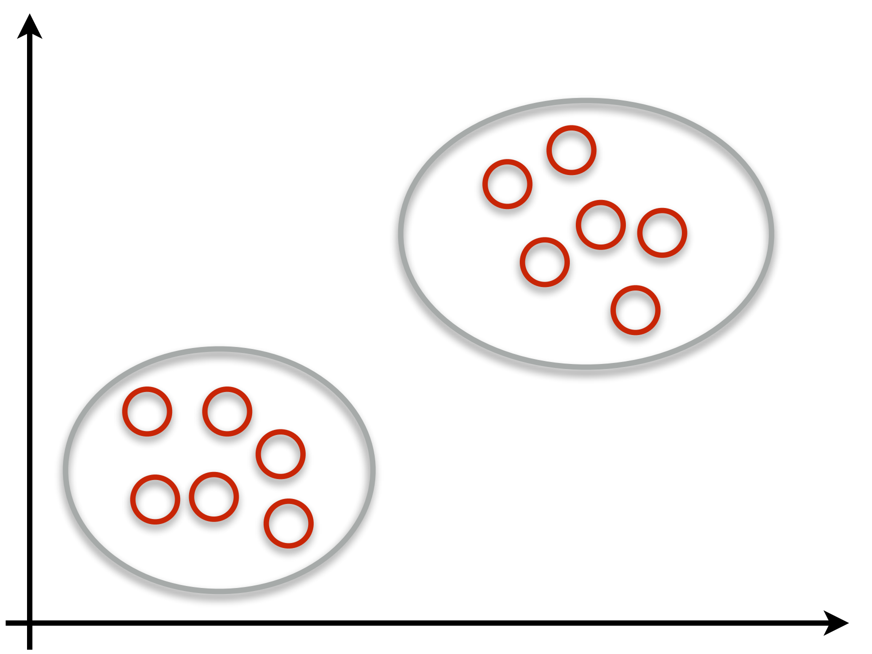


Classification

Spam detection

Image recognition

Unsupervised learning

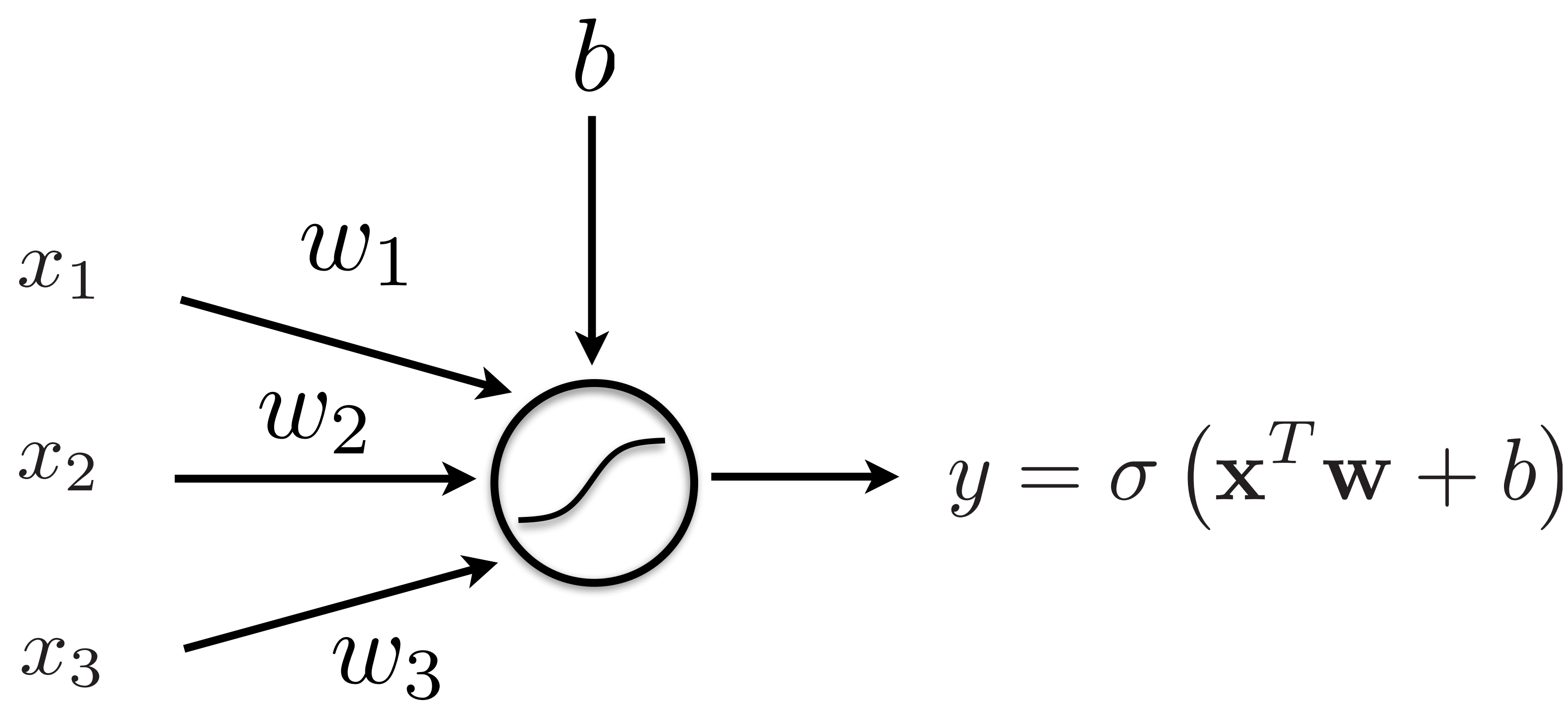


Clustering

Online advertising

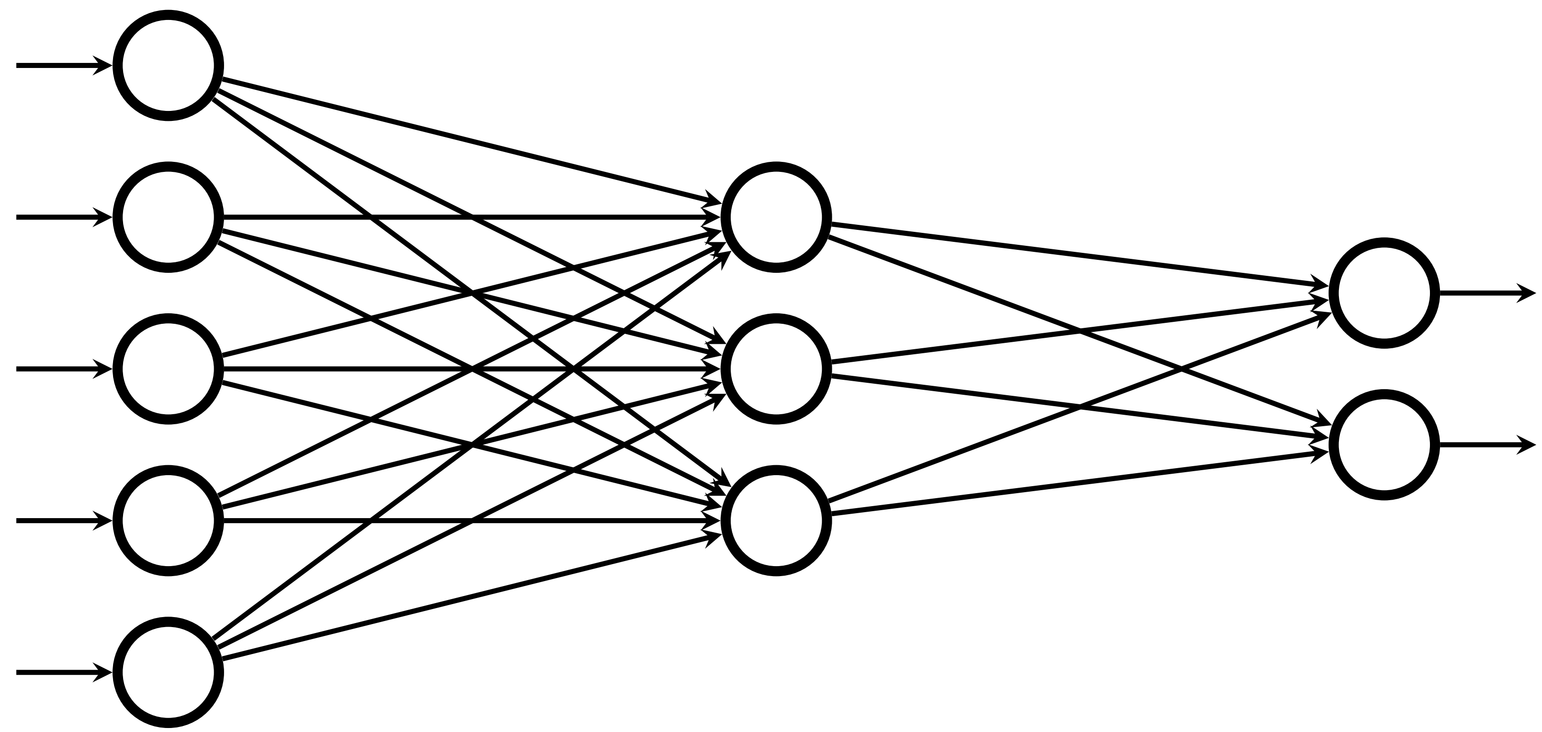
Anomaly detection

Artificial Neural Networks

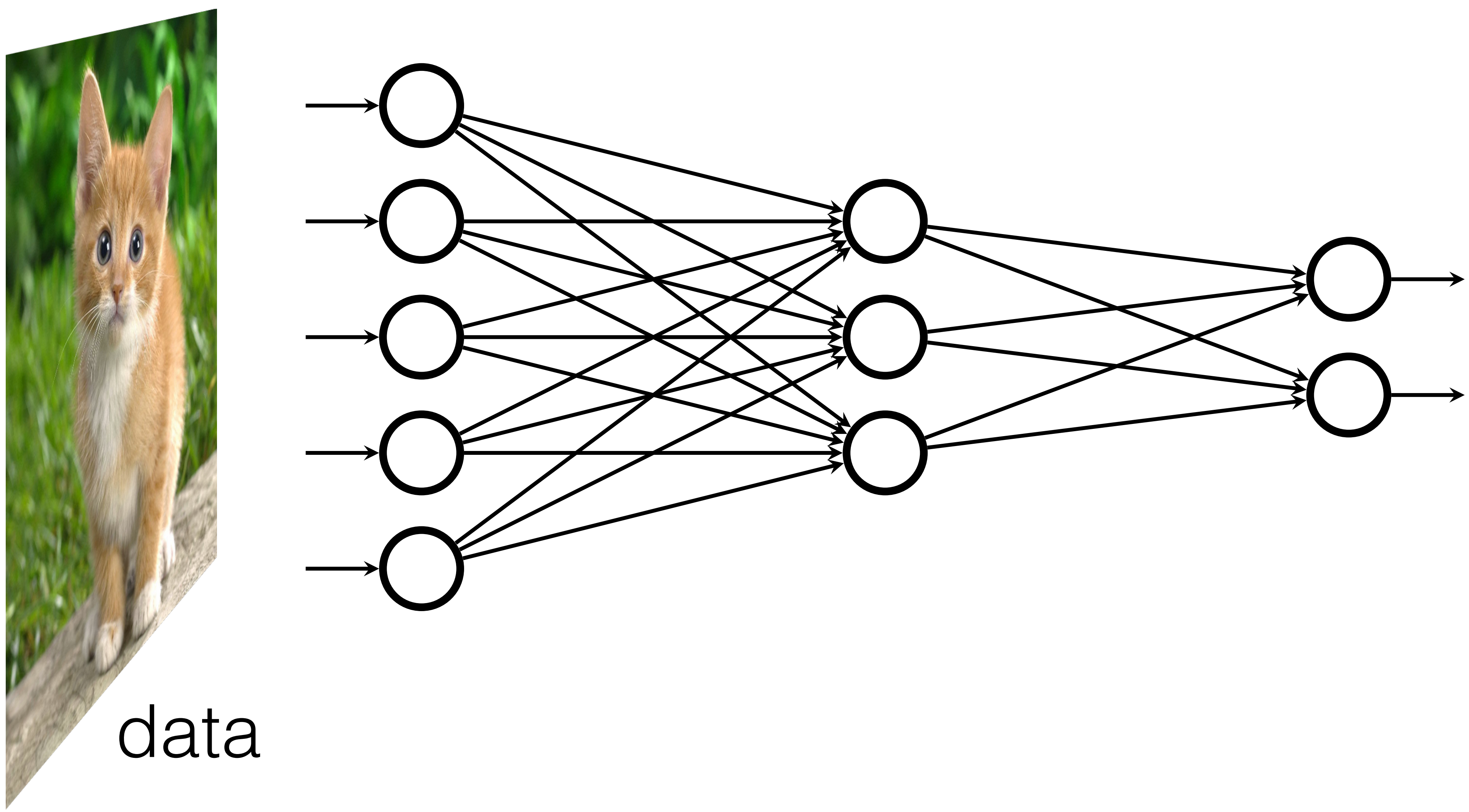


Computing unit: artificial neuron

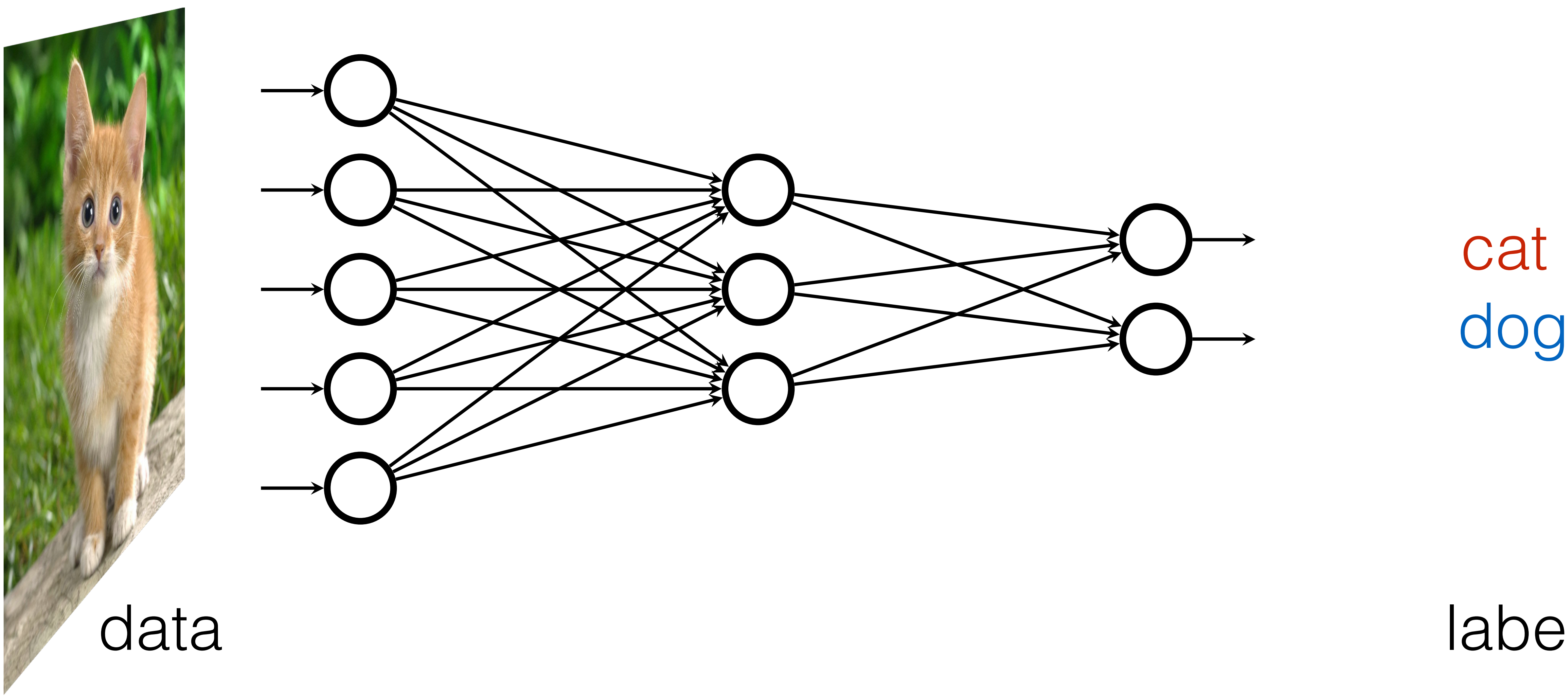
Artificial Neural Networks



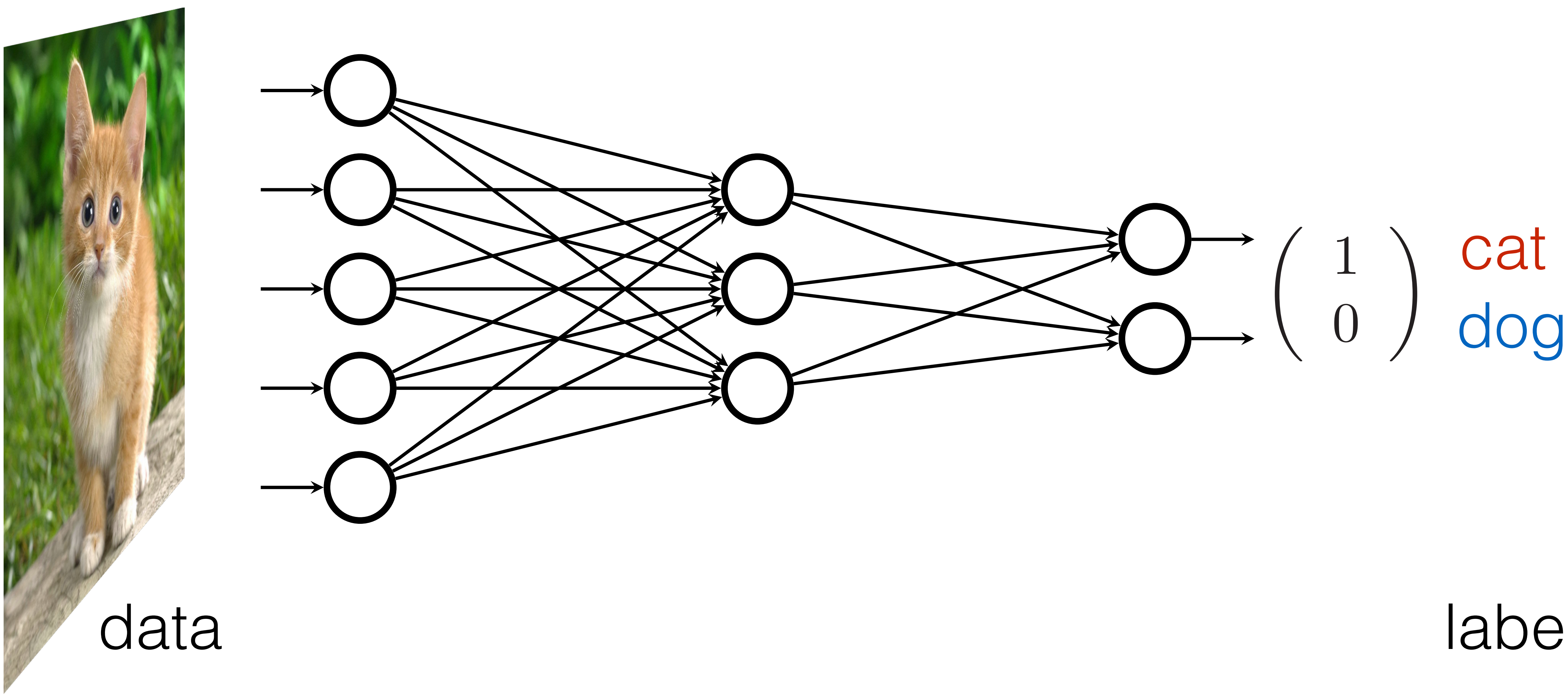
Artificial Neural Networks



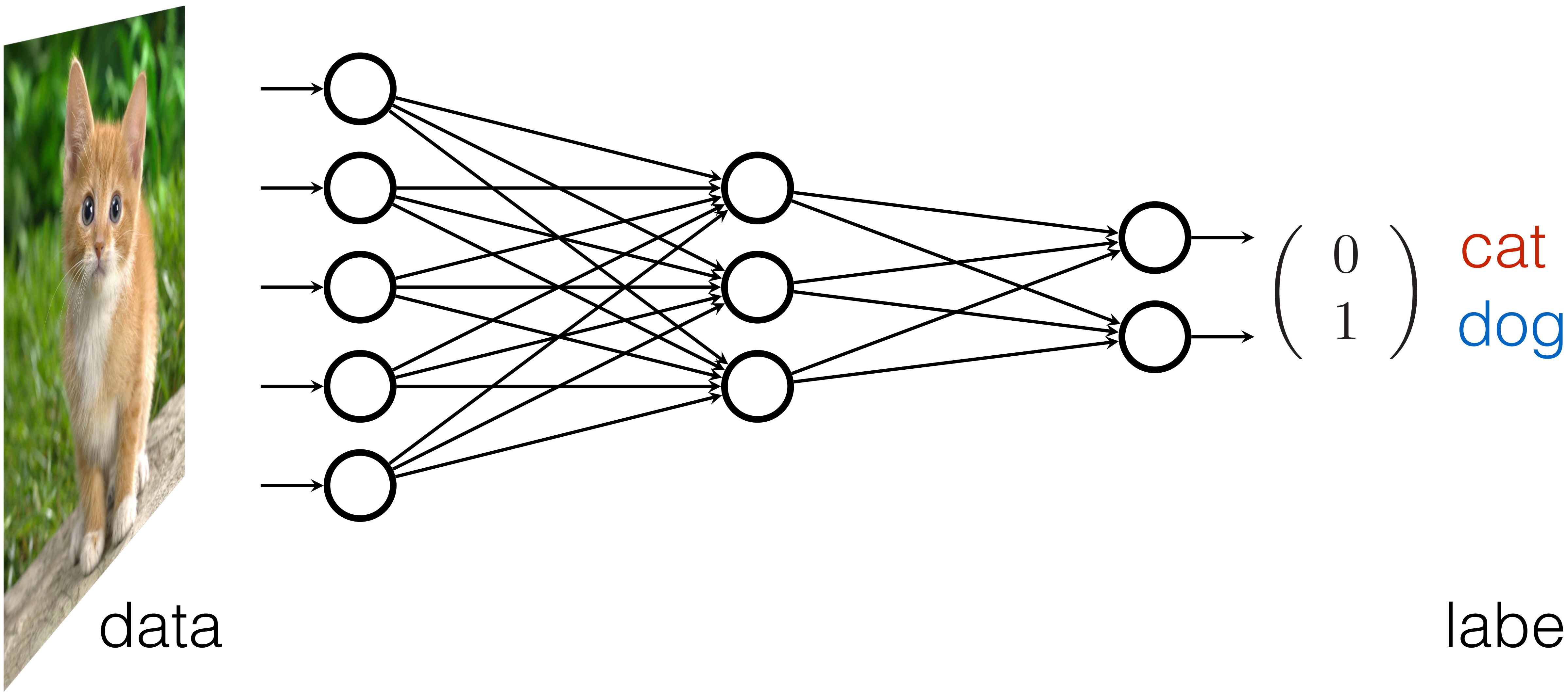
Artificial Neural Networks



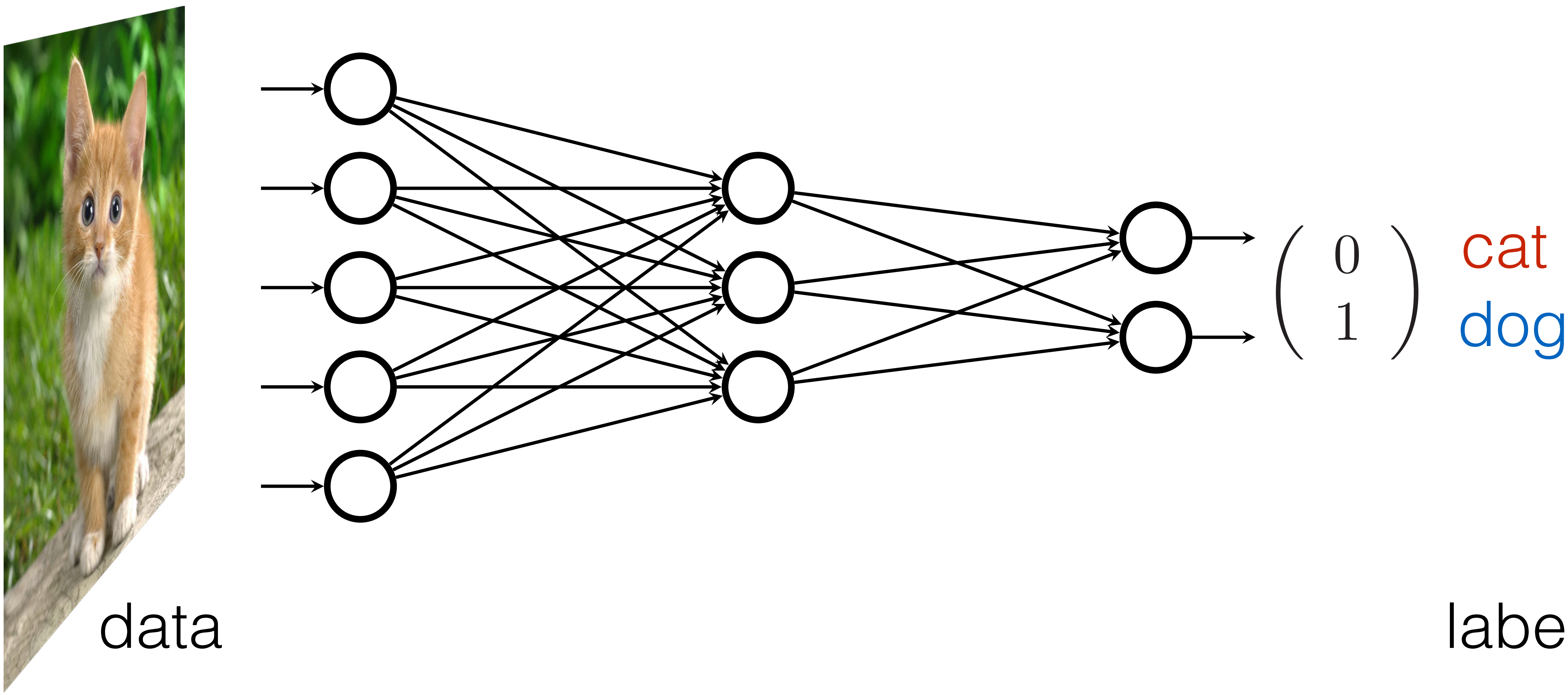
Artificial Neural Networks



Artificial Neural Networks



Artificial Neural Networks

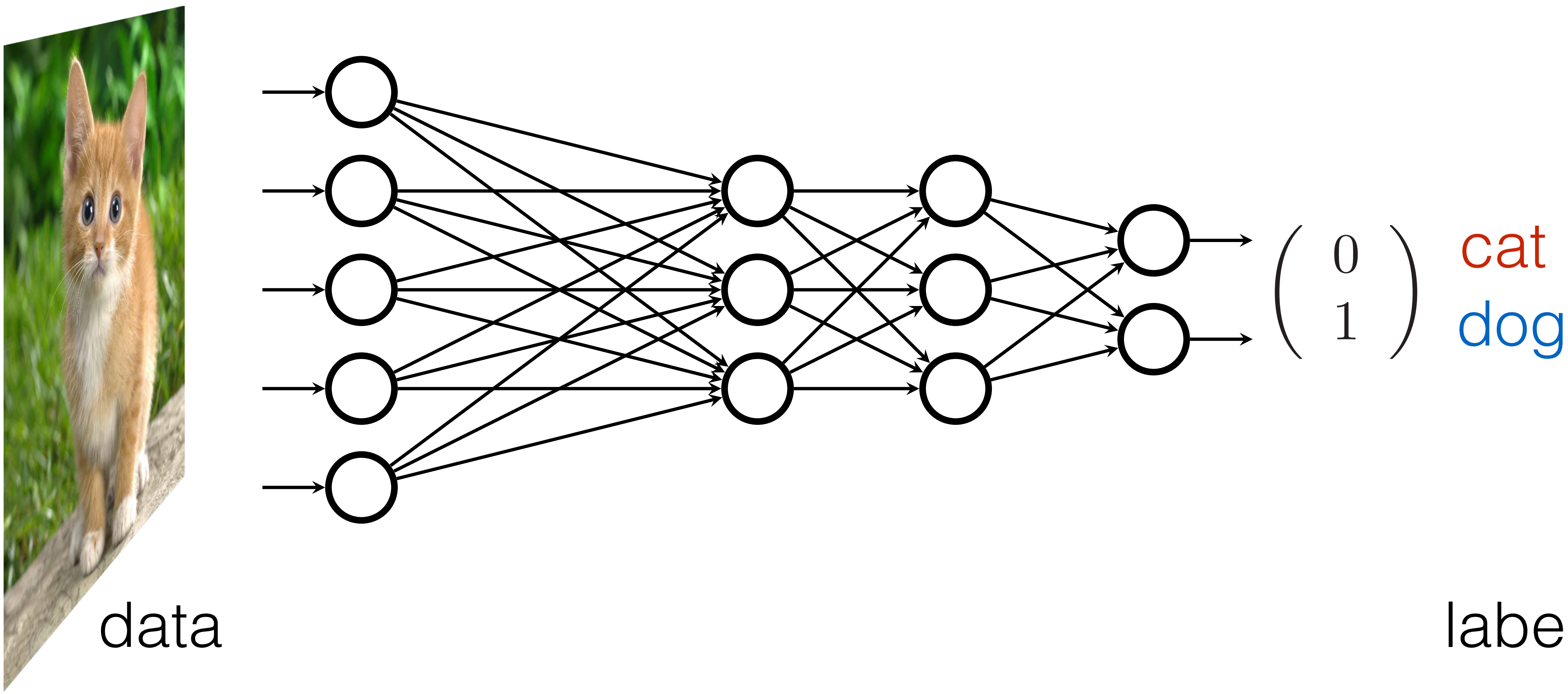


Universal Function Approximator

Cybenko 1989

Hornik, Stinchcombe, White 1989

Artificial Neural Networks

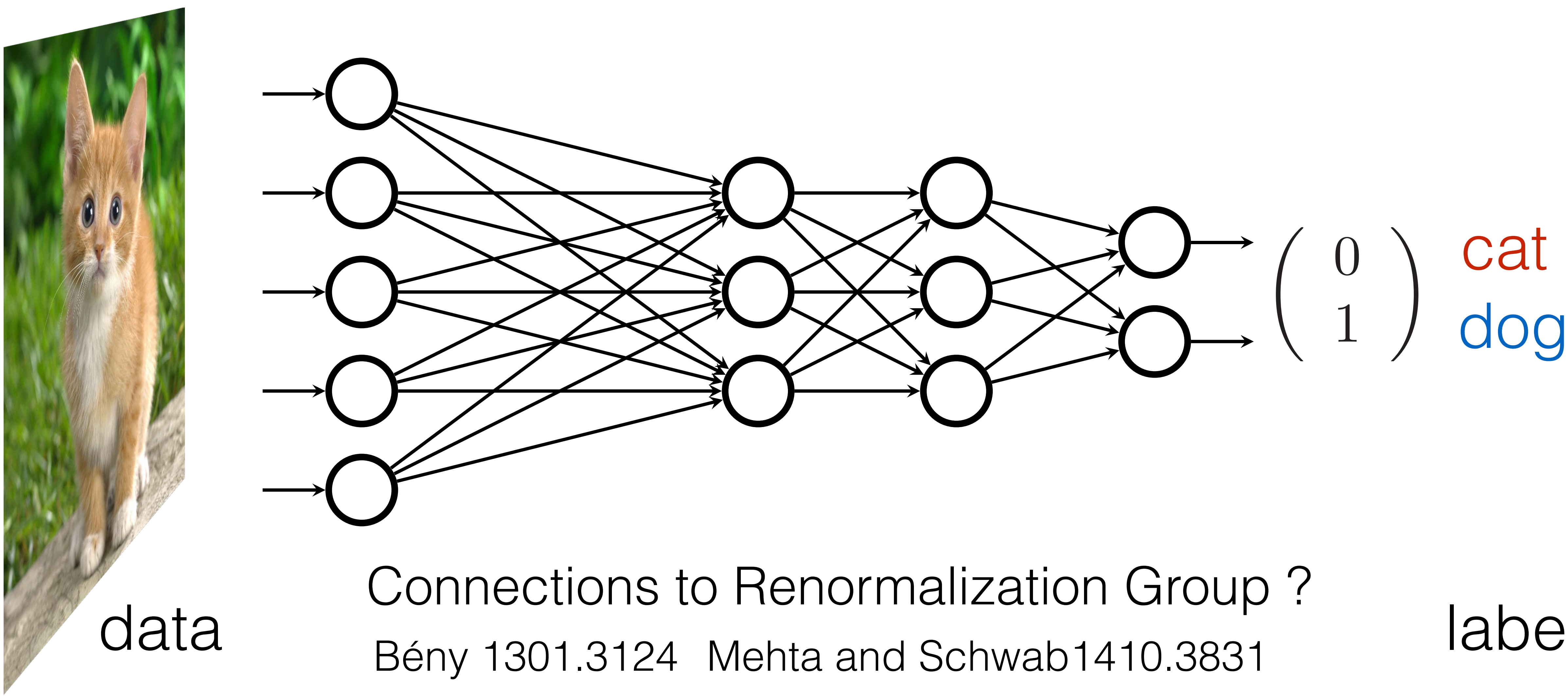


Universal Function Approximator

Cybenko 1989

Hornik, Stinchcombe, White 1989

Artificial Neural Networks



Connections to Renormalization Group ?

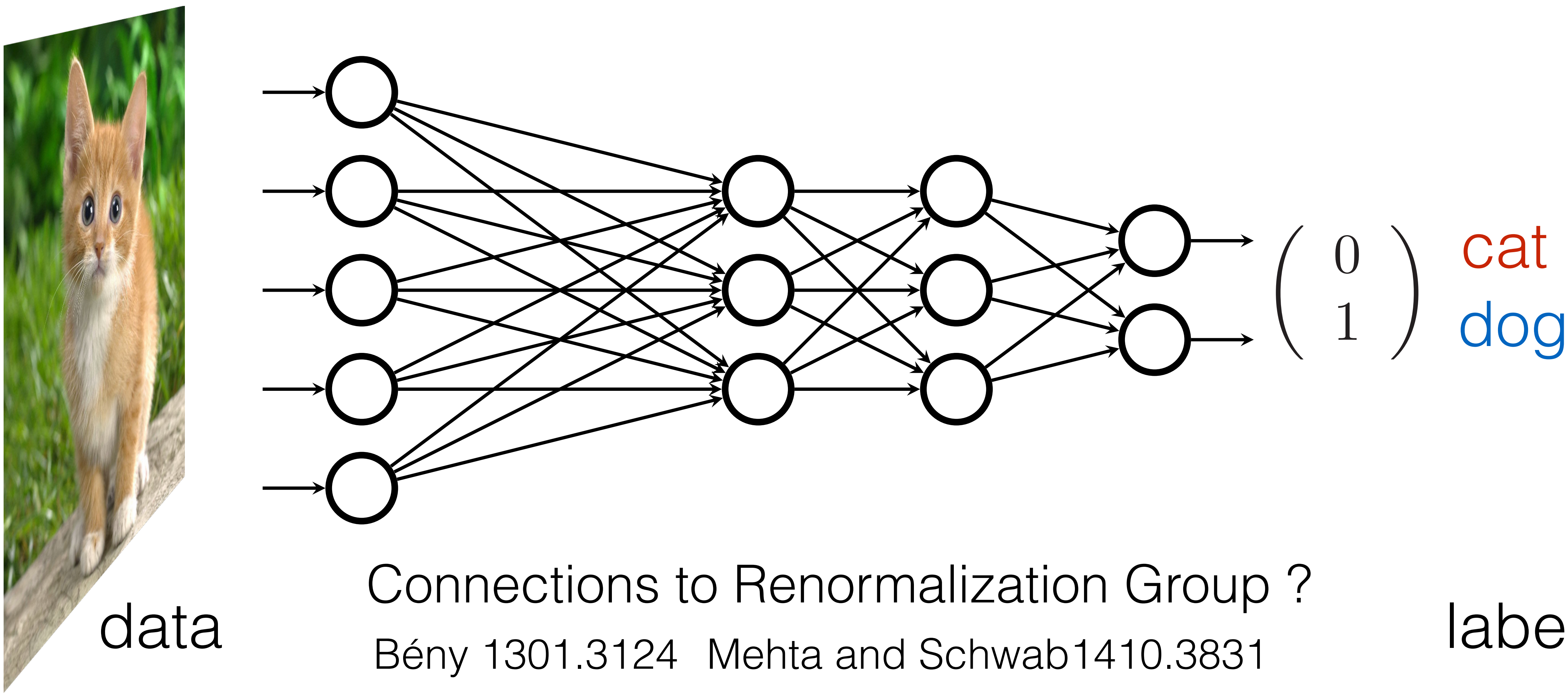
Bény 1301.3124 Mehta and Schwab 1410.3831

Universal Function Approximator

Cybenko 1989

Hornik, Stinchcombe, White 1989

Artificial Neural Networks



Connections to Renormalization Group ?

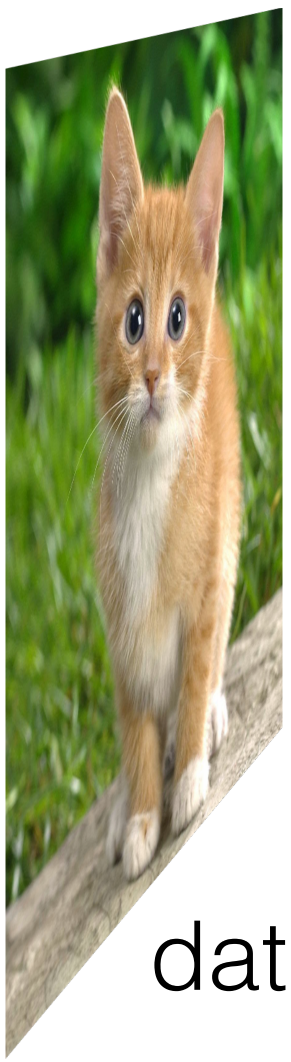
Bény 1301.3124   Mehta and Schwab 1410.3831

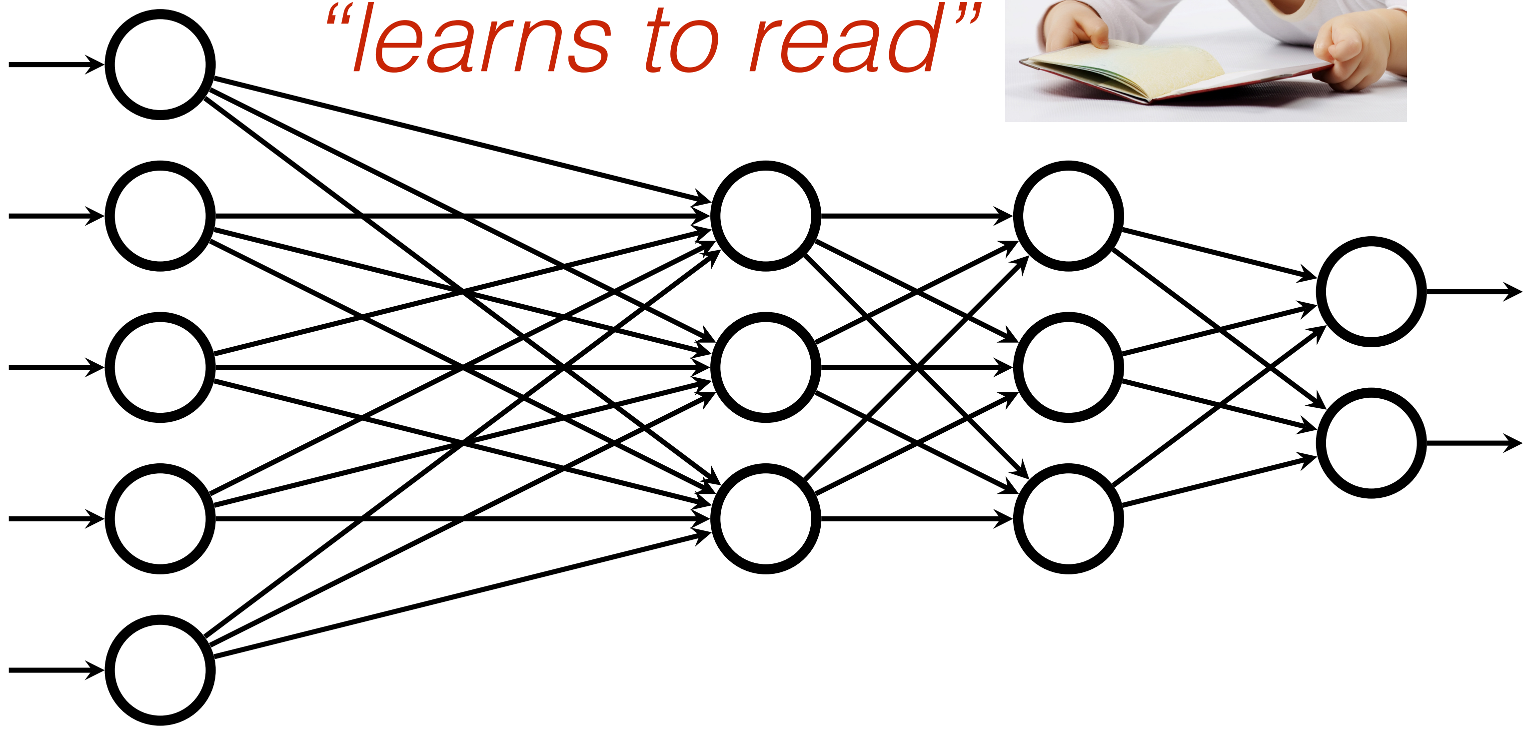
Why deep learning works? Not only a math problem, but also because of the law of physics: symmetry, locality, and compositionality

Lin and Tegmark, 1608.08225

Artificial Neural Networks

discriminative
learning:
"learns to read"





$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

cat
dog

data

Connections to Renormalization Group ?

Bény 1301.3124 Mehta and Schwab 1410.3831

label

Why deep learning works? Not only a math problem, but also because of the law of physics: symmetry, locality, and compositionality

Lin and Tegmark, 1608.08225

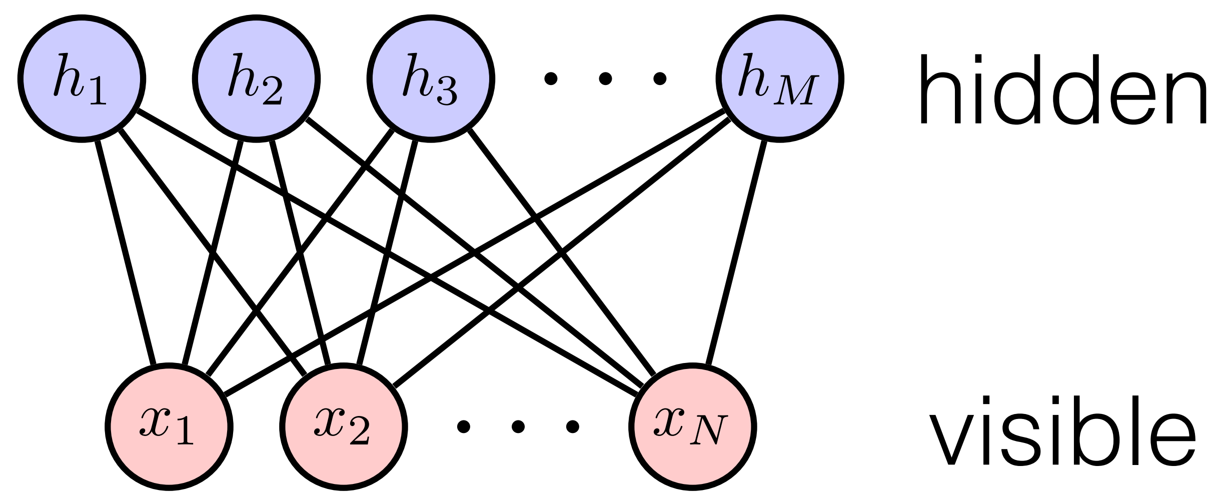
Restricted Boltzmann Machines

generative
learning:
"learns to write"

Restricted Boltzmann Machines

Smolensky 1986 Hinton and Sejnowski 1986

$$E(\mathbf{x}, \mathbf{h}) = -\sum_{i=1}^{N} a_i x_i - \sum_{j=1}^{M} b_j h_j - \sum_{i=1}^{N} \sum_{j=1}^{M} x_i W_{ij} h_j$$

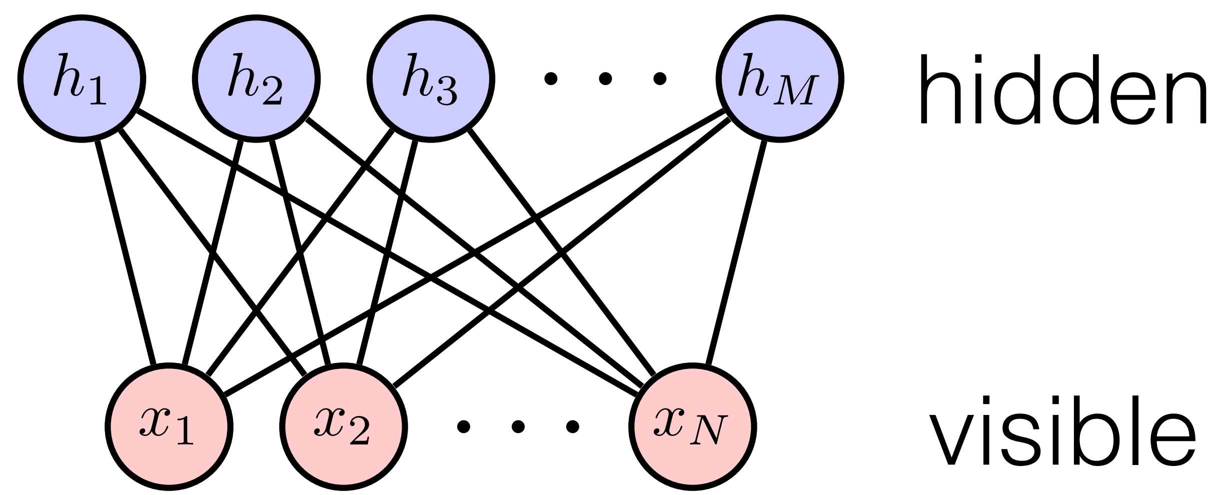


Restricted Boltzmann Machines

Smolensky 1986 Hinton and Sejnowski 1986

$$E(\mathbf{x}, \mathbf{h}) = -\sum_{i=1}^{N} a_i x_i - \sum_{j=1}^{M} b_j h_j - \sum_{i=1}^{N} \sum_{j=1}^{M} x_i W_{ij} h_j$$

Learn

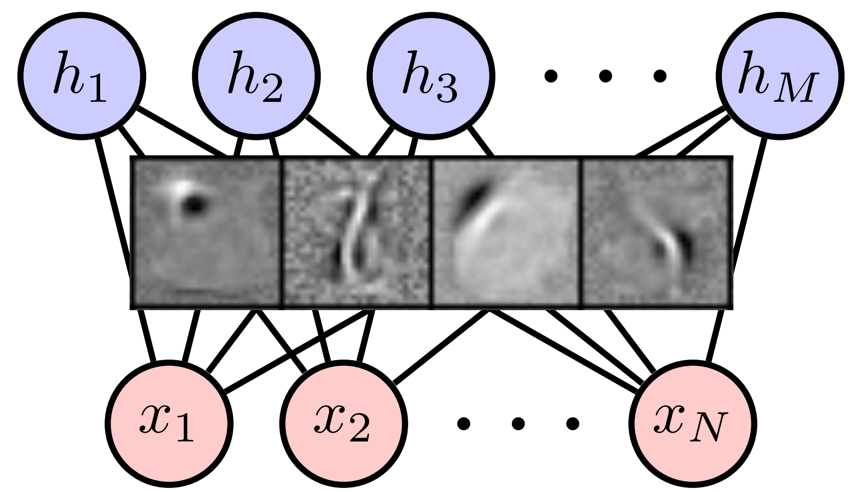


Restricted Boltzmann Machines

Smolensky 1986 Hinton and Sejnowski 1986

$$E\left(\mathbf{x},\mathbf{h}\right) = -\sum_{i=1}^{N} a_i x_i - \sum_{j=1}^{M} b_j h_j - \sum_{i=1}^{N}\sum_{j=1}^{M} x_i W_{ij} h_j$$

Learn



hidden

visible

Marginal Probability

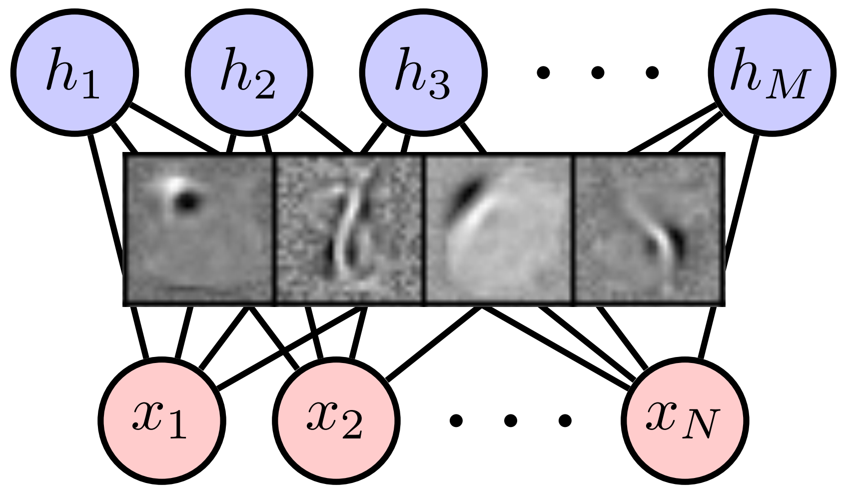
$$p(\mathbf{x}) = \sum_{\mathbf{h}} e^{-E(\mathbf{x},\mathbf{h})}$$

Restricted Boltzmann Machines

Smolensky 1986 Hinton and Sejnowski 1986

$$E\left(\mathbf{x},\mathbf{h}\right) = -\sum_{i=1}^{N} a_i x_i - \sum_{j=1}^{M} b_j h_j - \sum_{i=1}^{N}\sum_{j=1}^{M} x_i W_{ij} h_j$$

Learn



Generate

Marginal Probability

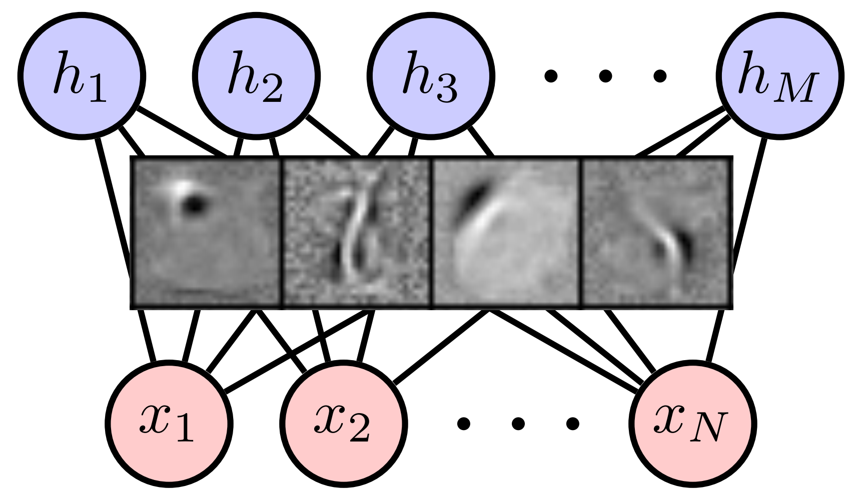
$$p(\mathbf{x}) = \sum_{\mathbf{h}} e^{-E(\mathbf{x},\mathbf{h})}$$

Restricted Boltzmann Machines

Smolensky 1986 Hinton and Sejnowski 1986

$$E\left(\mathbf{x},\mathbf{h}\right) = -\sum_{i=1}^{N} a_i x_i - \sum_{j=1}^{M} b_j h_j - \sum_{i=1}^{N}\sum_{j=1}^{M} x_i W_{ij} h_j$$

Learn



Generate

Marginal Probability

$$p(\mathbf{x}) = \sum_{\mathbf{h}} e^{-E(\mathbf{x},\mathbf{h})}$$

Universal approximator of probability distributions

. Freund and Haussler, 1989 Le Roux and Bengio, 2008

Why machine learning for many-body physics ?

- Conceptual connections: a new and natural way to think about (quantum) many-body systems

- Data driven approach: making scientific discovery based on big datasets

- Techniques: neural networks, kernel methods, pattern recognition, feature extraction, dimensional reduction, clustering analysis, probabilistic modeling, recommender systems, hardware acceleration, software frameworks…

Ideas

Ideas

A general way to do fittings

Solving inverse problems

Variational wave functions

Quantum state tomography/classifier/decoding

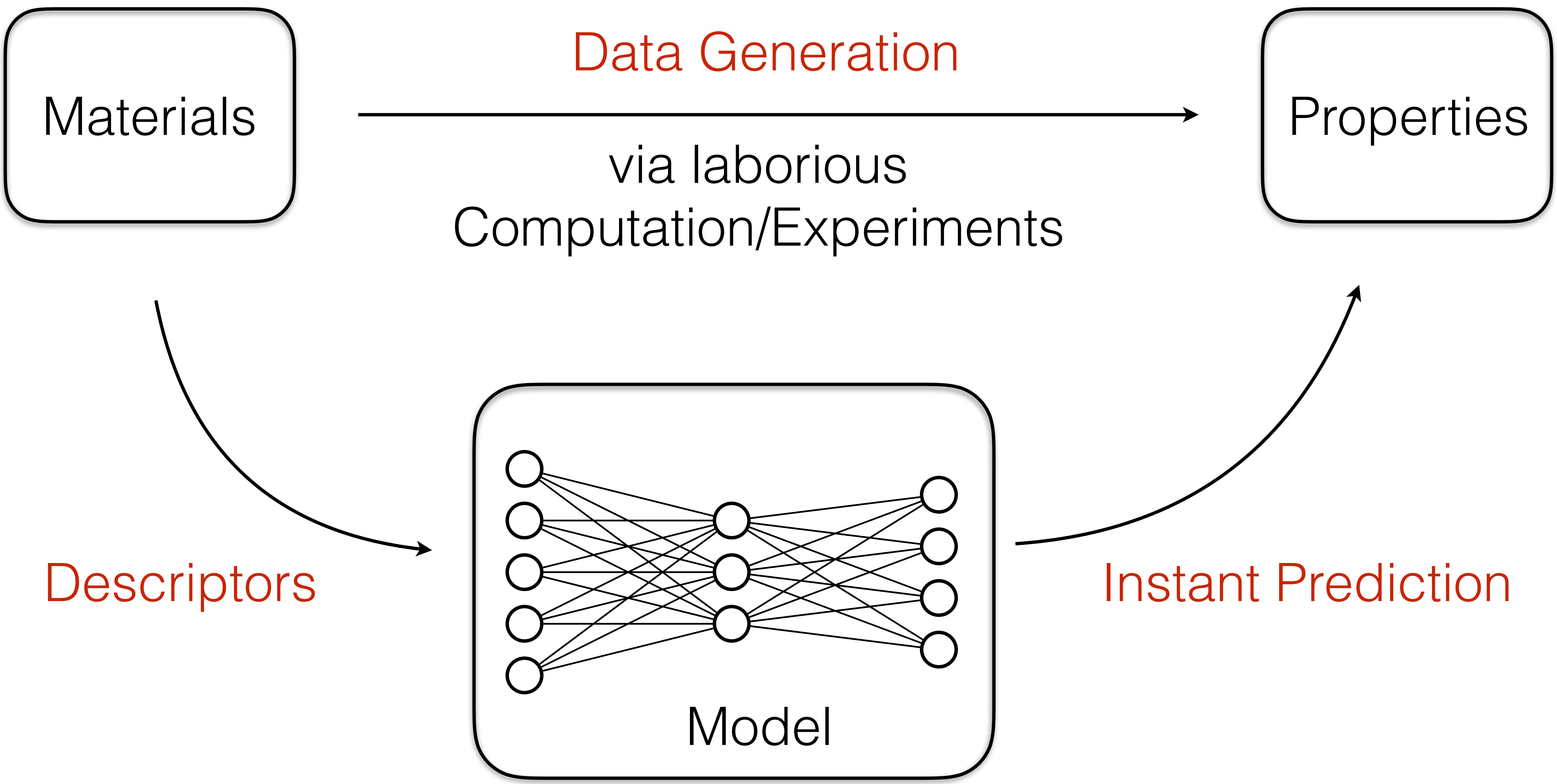
Classification/discovery phases of matter

Connection to tensor networks & RG

Recommender engines for QMC

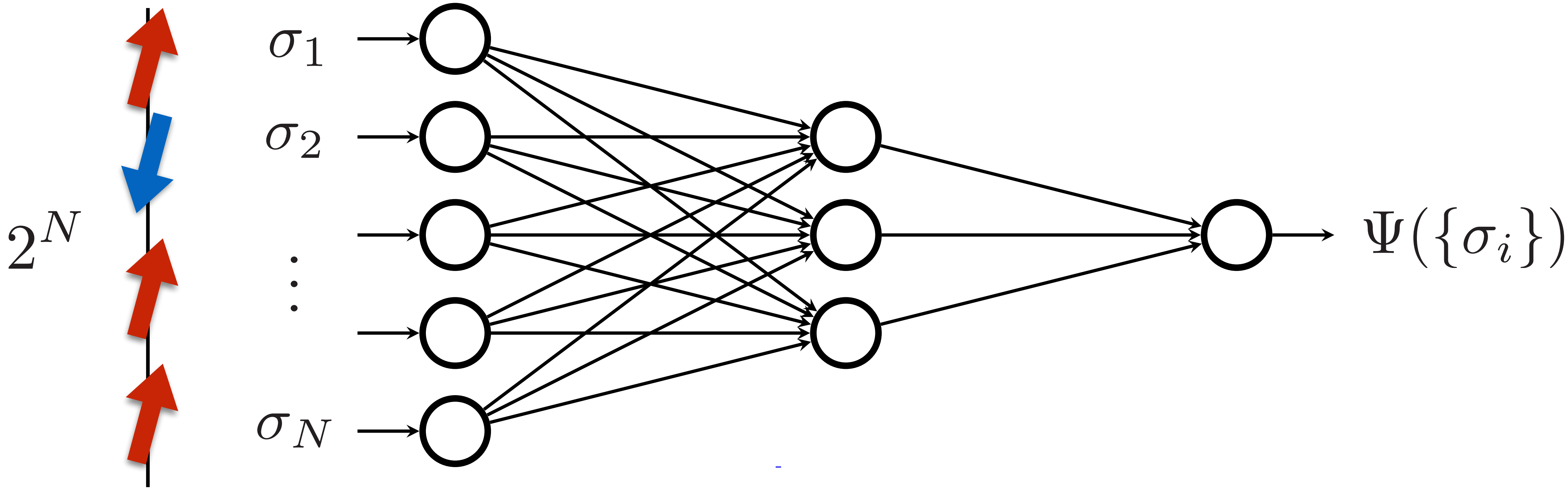
Function Approximation

Material Discovery



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

Variational wave functions

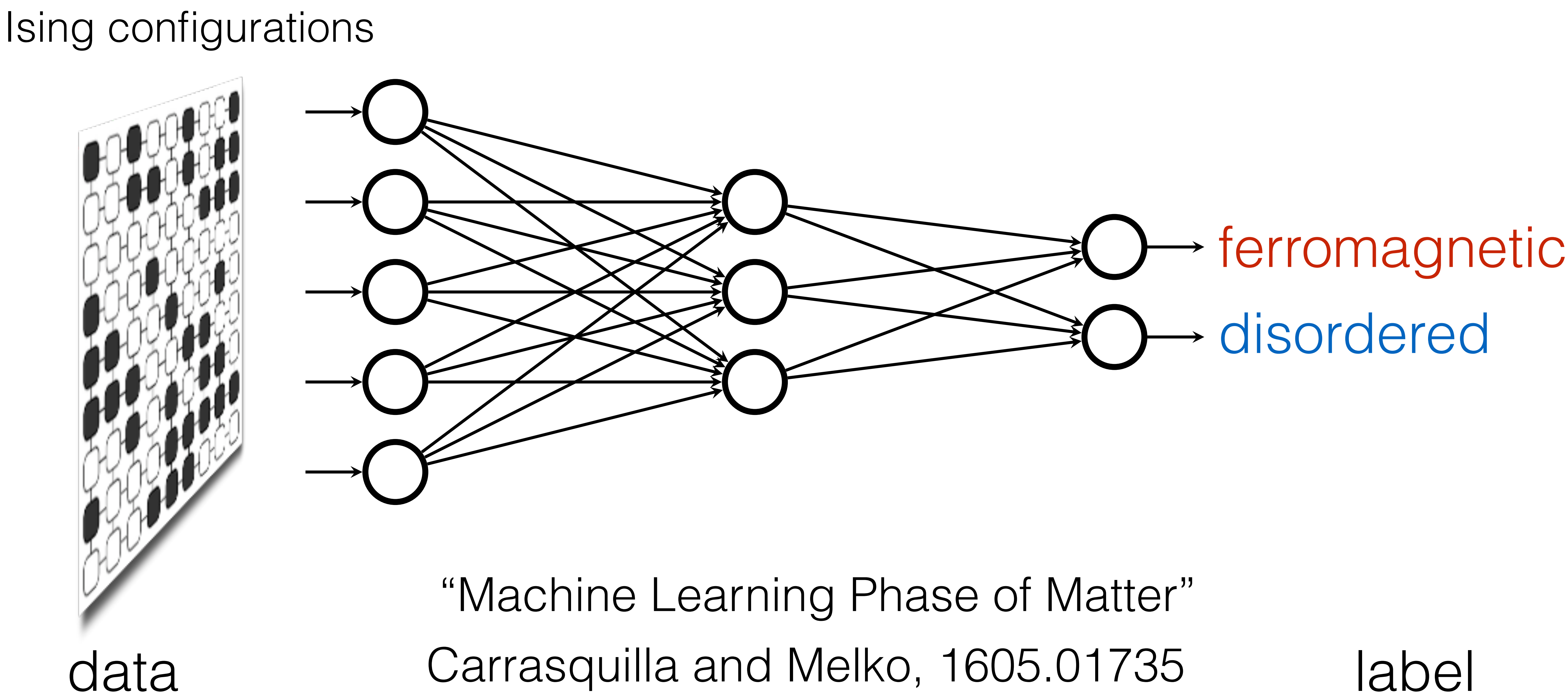


- Neural net as an efficient many-body wave function
- Universal function approximator
- Feature discovery and abstraction power from deep hierarchical structure

Carleo and Troyer, 1606.02318 Deng, Li, Das Sarma, 1609.09060 Zi Cai, 1704.05148

“Phase” Recognition

Supervised Approach



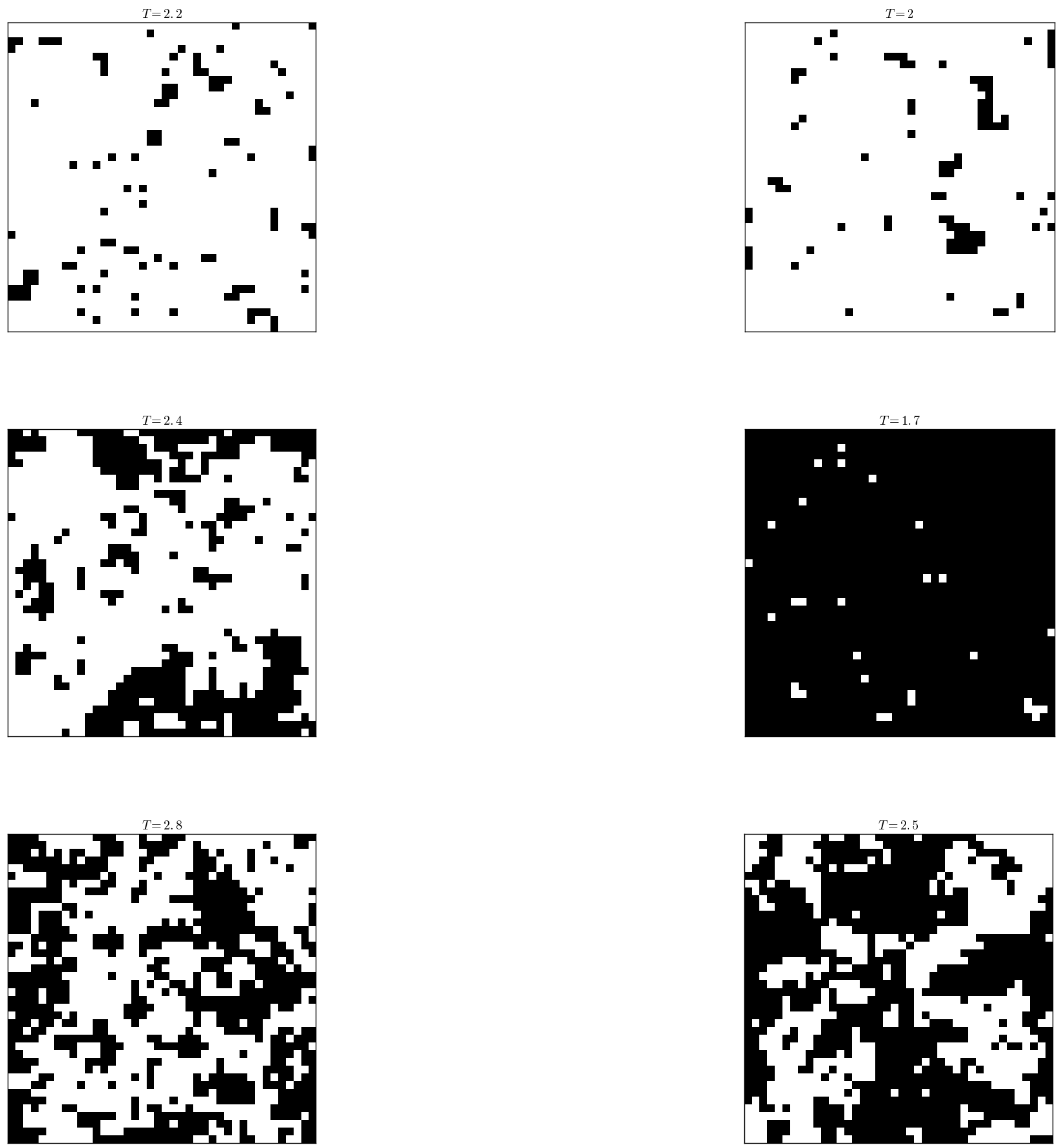
"Machine Learning Phase of Matter"

Carrasquilla and Melko, 1605.01735

Broecker, Carrasquilla, Melko, Trebst, 1608.07848
Ch'ng, Carrasquilla, Melko, Khatami, 1609.02552
Schindler, Regnault, Neupert,1704.01578
Ponte, Melko, 1704.05848

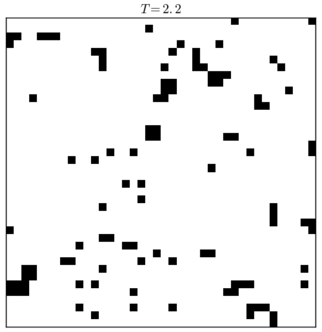
Tanaka, Tomiya 1609.09087
Ohtsuki, Ohtsuki,1610.00462
1612.04909
Zhang, Kim,1611.01518

Unsupervised Approach

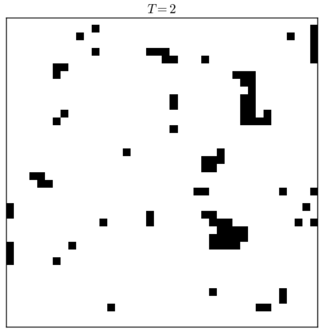


Unsupervised Approach

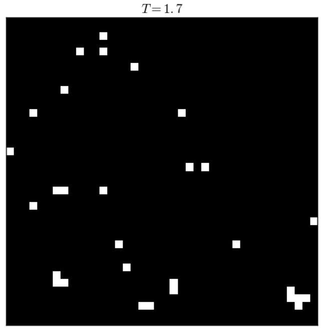
$T = 2.2$



$T = 2$

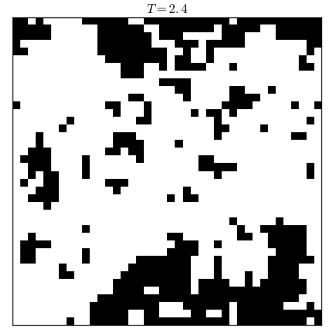


$T = 1.7$

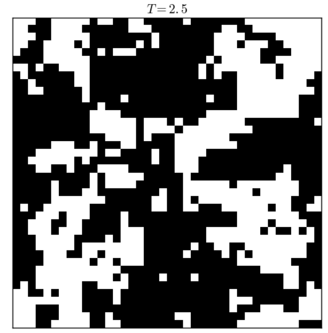


ferromagnetic

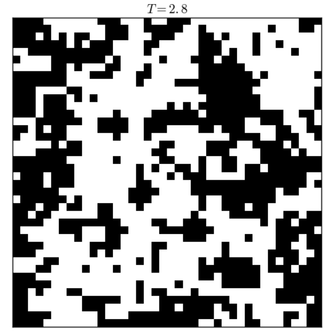
$T = 2.4$



$T = 2.5$

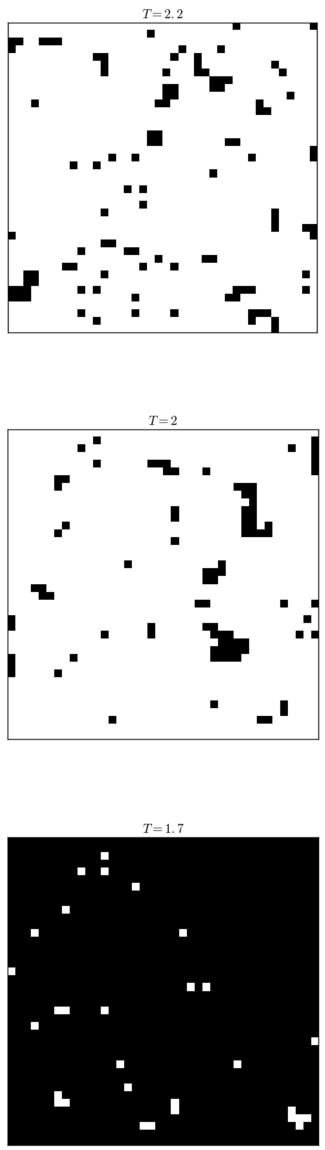


$T = 2.8$

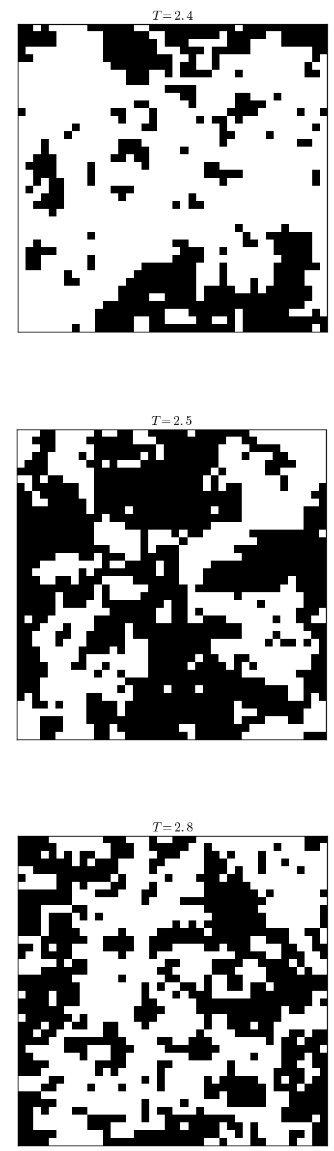


disordered

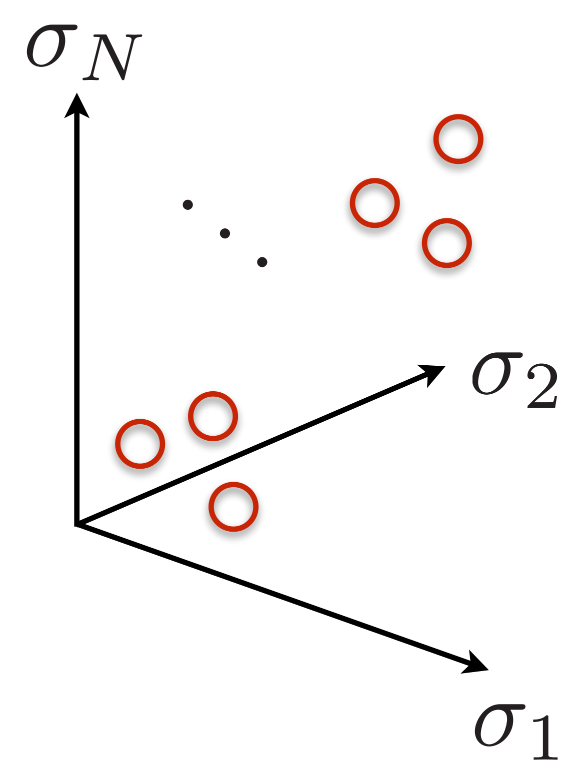
Unsupervised Approach



ferromagnetic

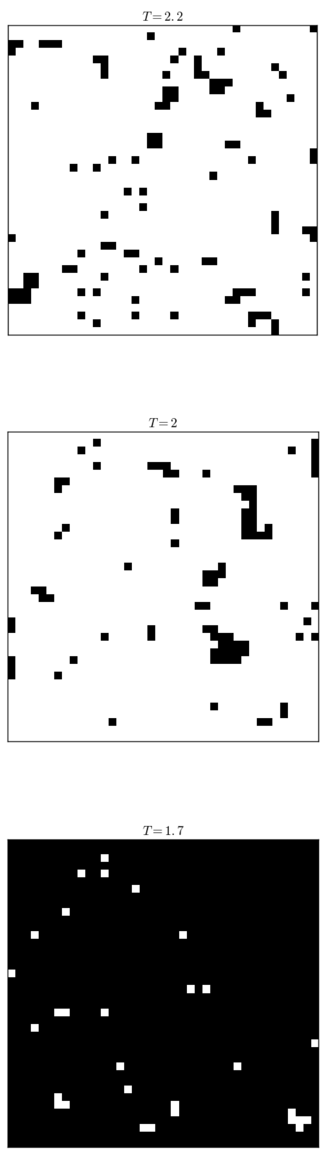


disordered

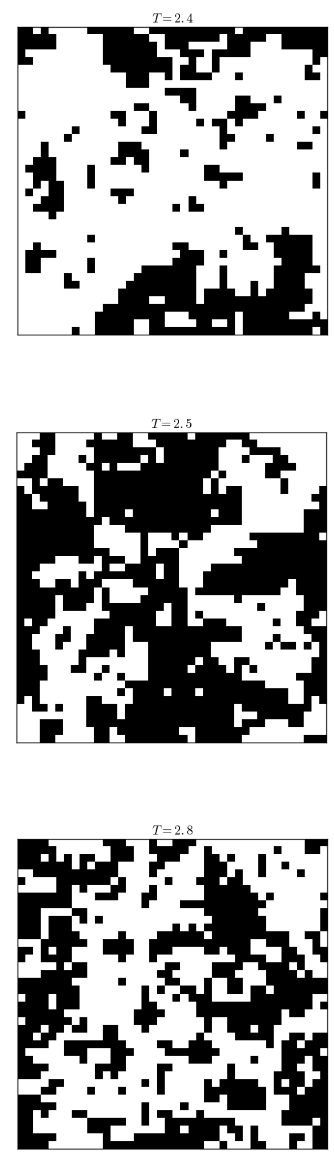


only data, no label

Unsupervised Approach



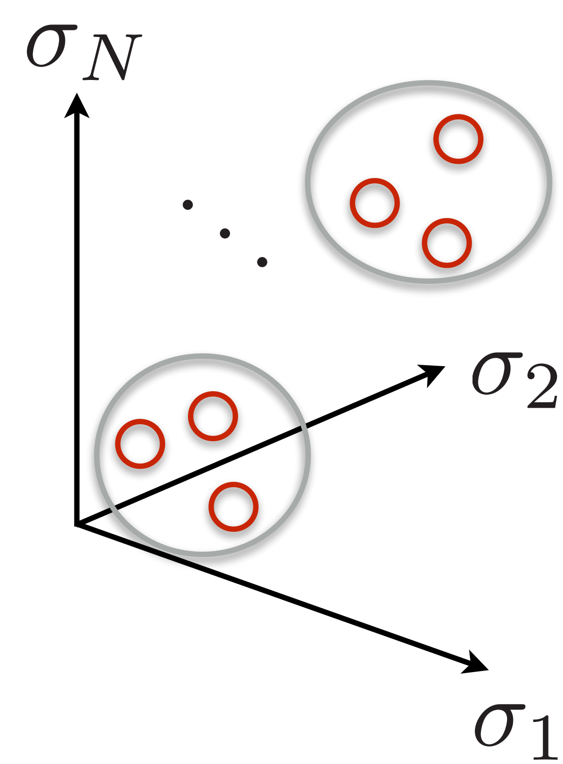
ferromagnetic



disordered

LW, 1606.00318

Discovering phase transition with dimensional reduction and clustering analysis



only data, no label

Nieuwenburg, Liu, Huber, 1610.02048
Liu, Nieuwenburg, 1706.08111
Broecker, Assaad, Trebst, 1707.00663

Wetzel, 1703.02435
Hu, Singh, Scalettar, 1704.00080
Wetzel, Scherzer, 1705.05582
Wang and Zhai, 1706.07977

Algorithmic Innovations

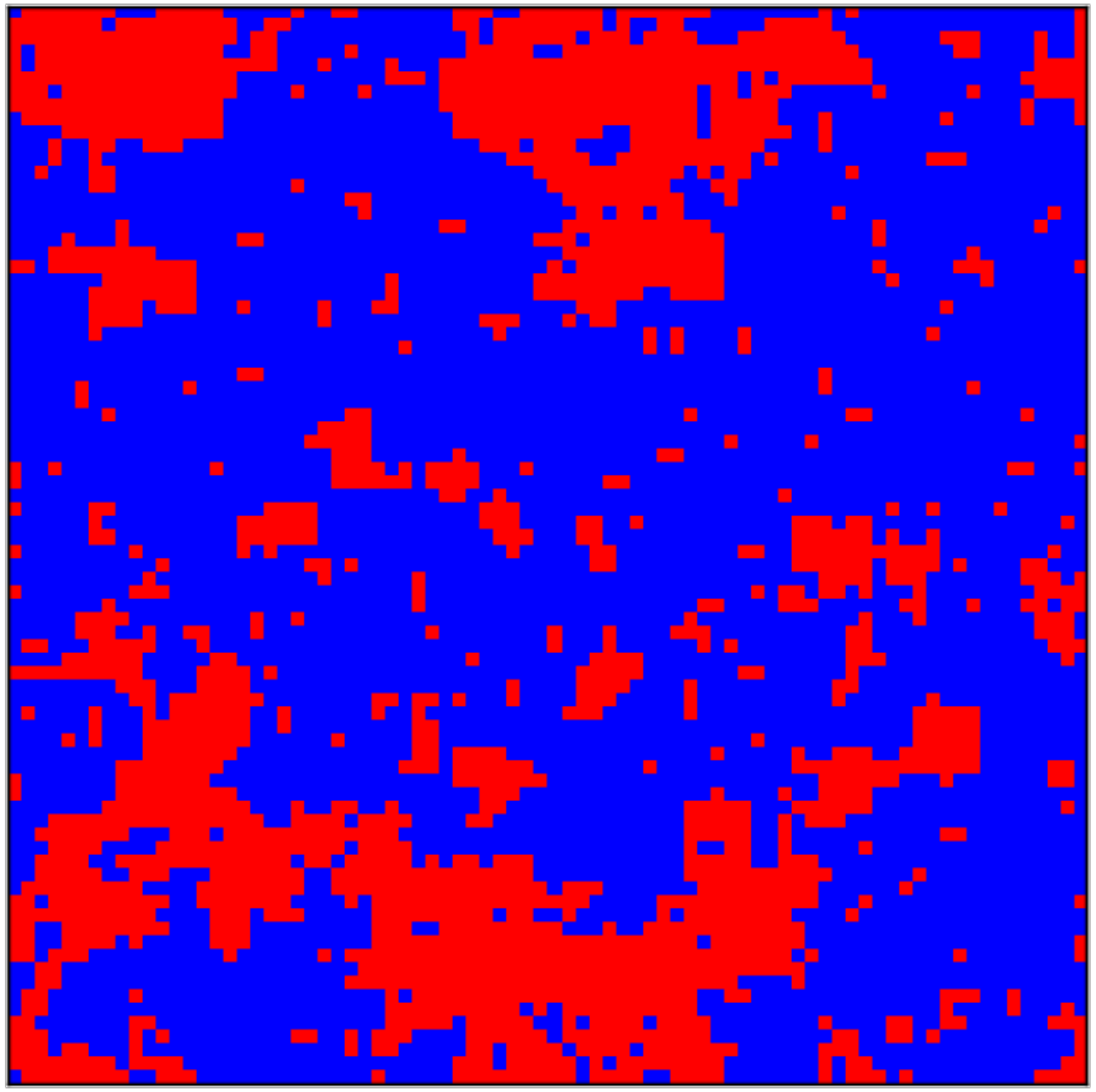
Liu, Qi, Meng, Fu,1610.03137
Liu, Shen, Qi, Meng, Fu,1611.09364
Xu, Qi, Liu, Fu, Meng,1612.03804
Nagai, Shen, Qi, Liu, Fu, 1705.06724

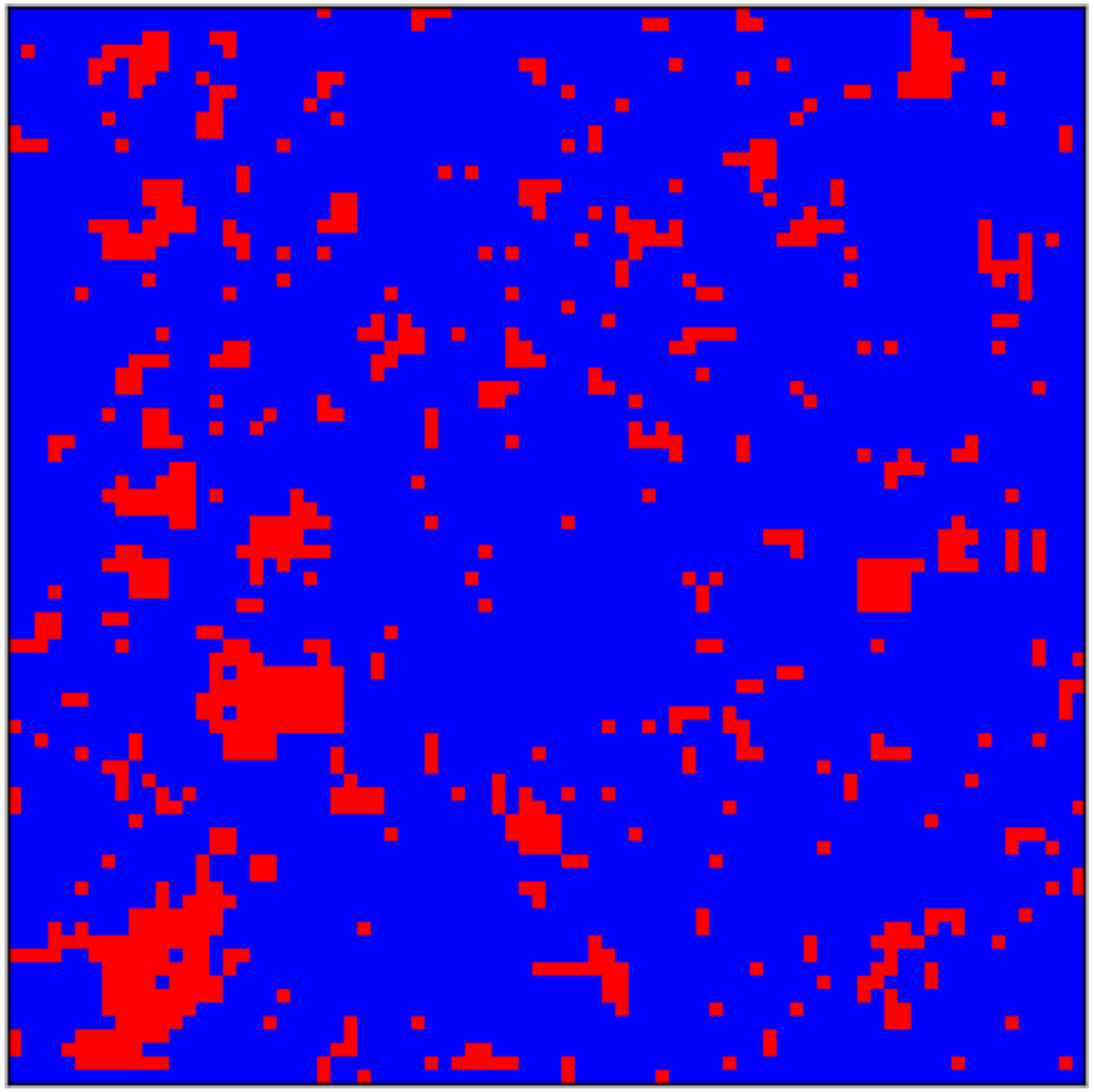
Li Huang and LW, 1610.02746
Li Huang, Yi-feng Yang and LW, 1612.01871
LW, 1702.08586

A Video from Google DeepMind

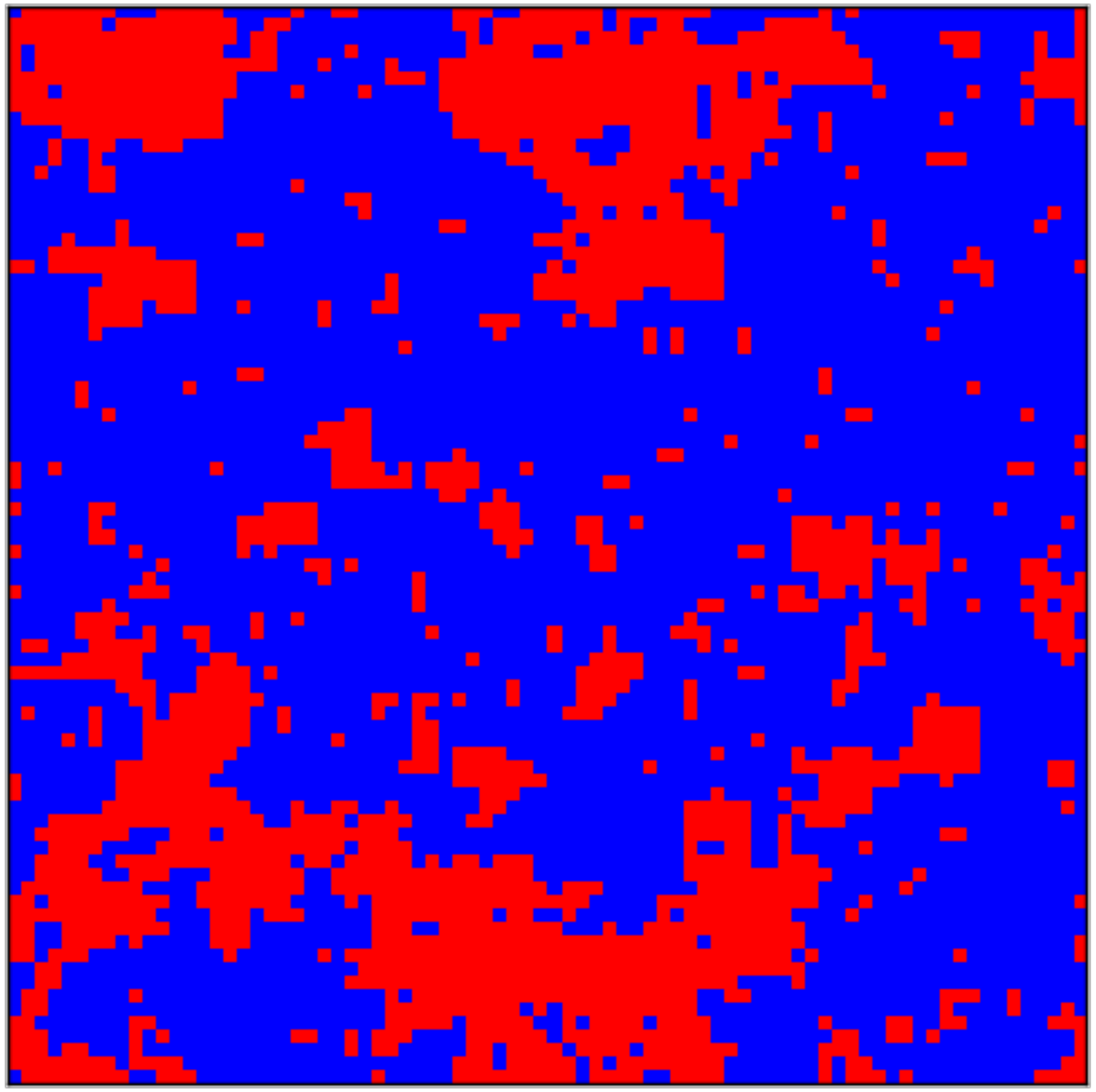
http://www.nature.com/nature/journal/v518/n7540/fig_tab/nature14236_SV2.html

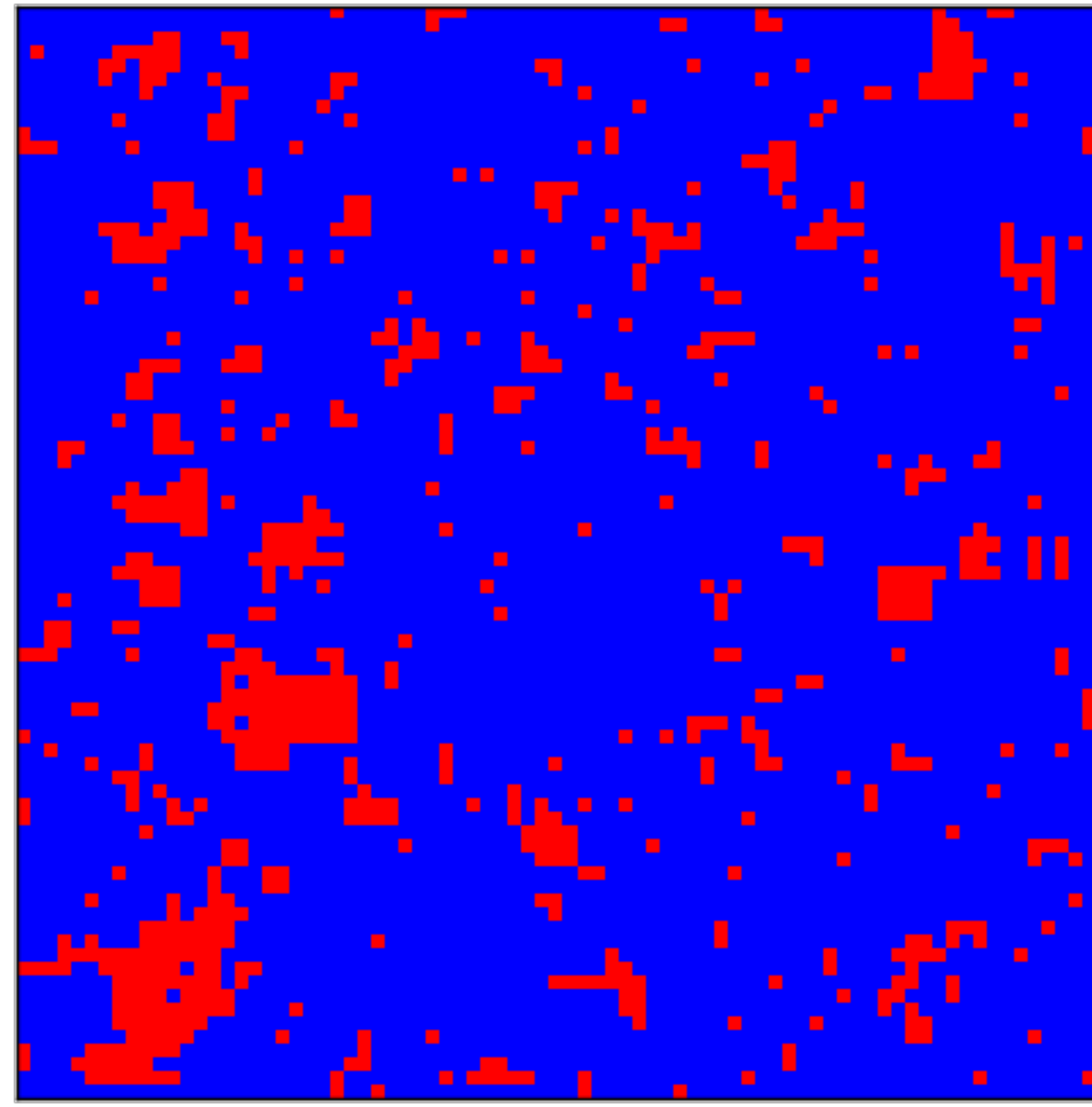
Local vs Cluster algorithms





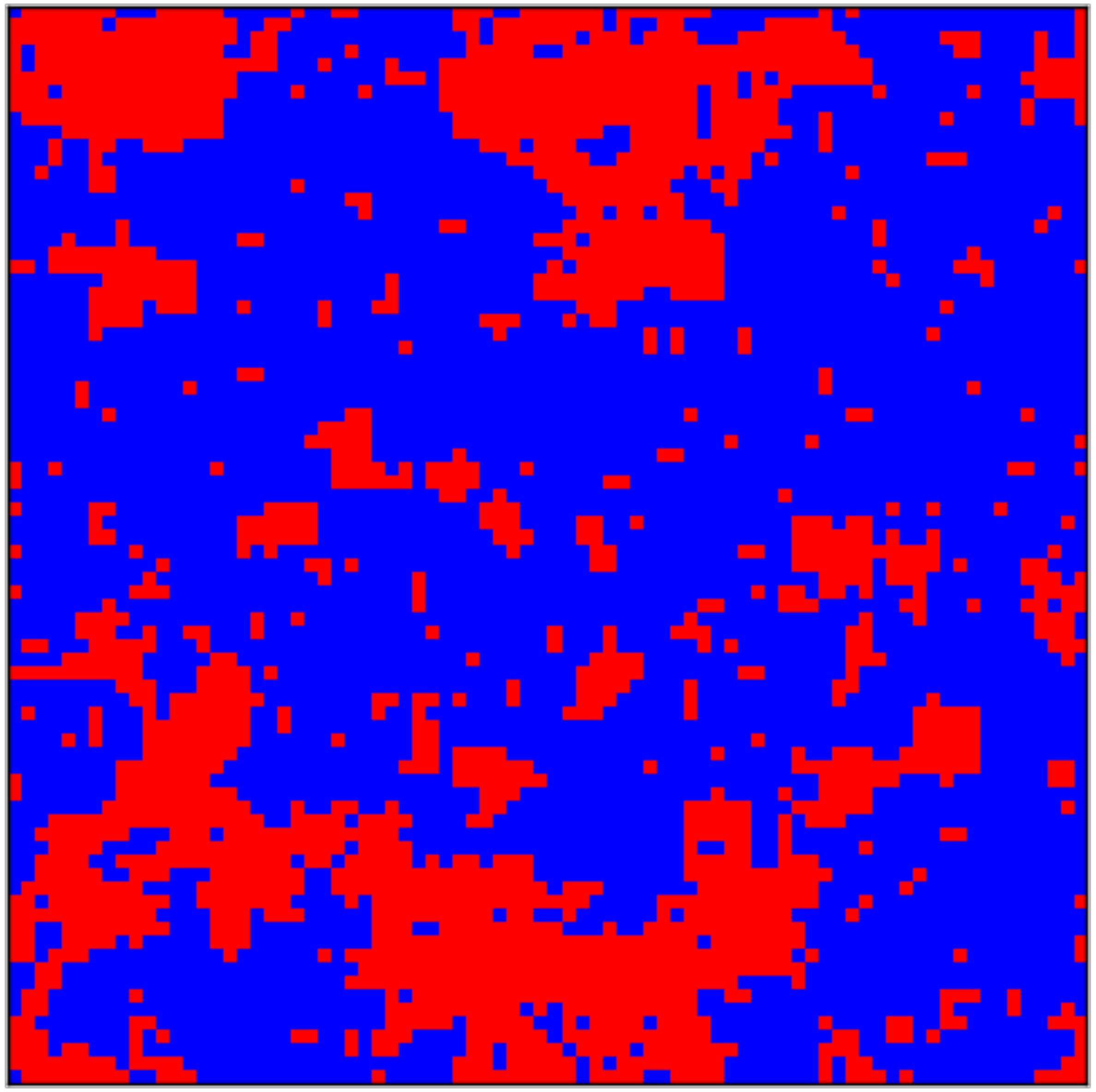
Local vs Cluster algorithms

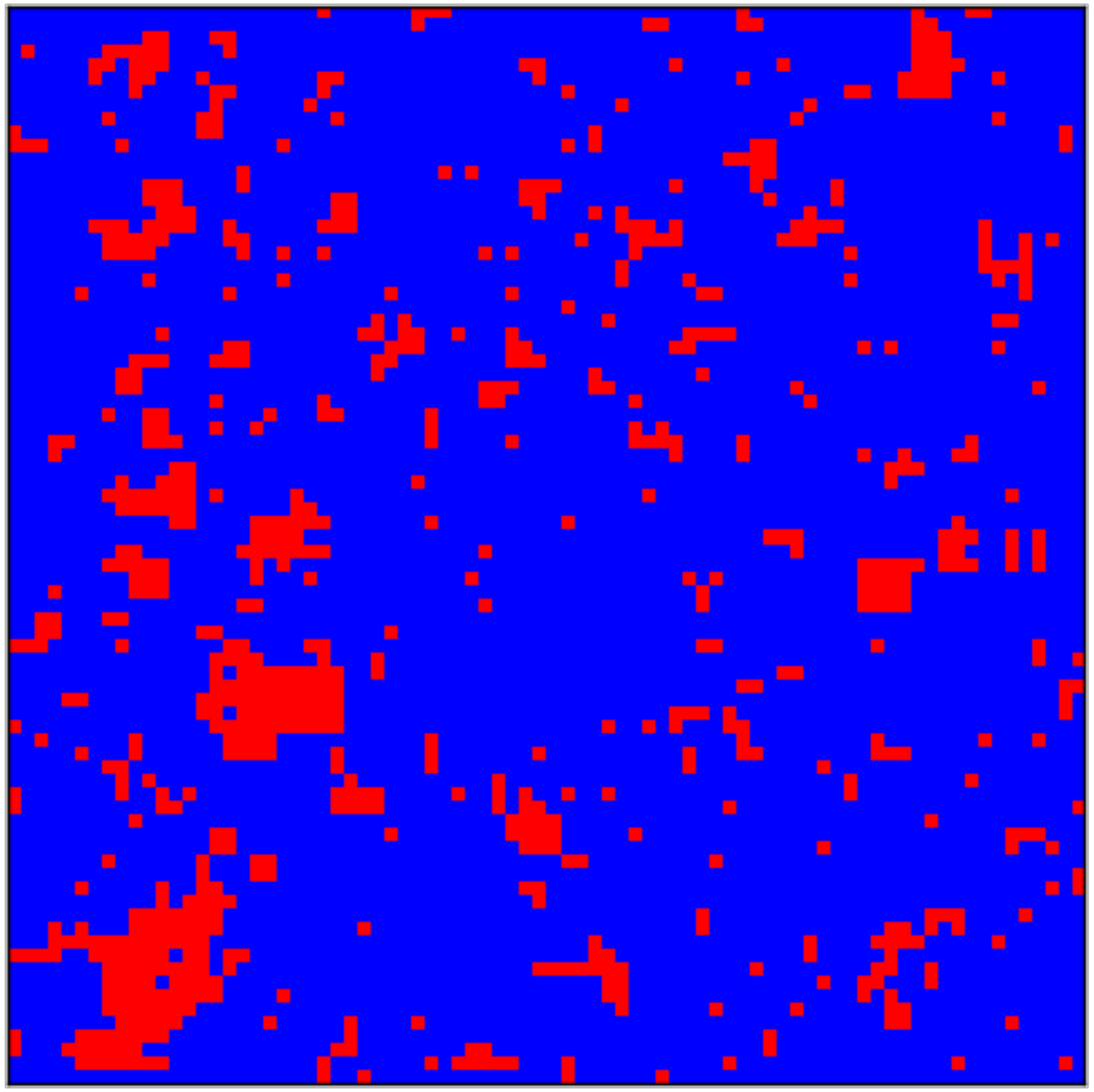




is slower than

Local vs Cluster algorithms





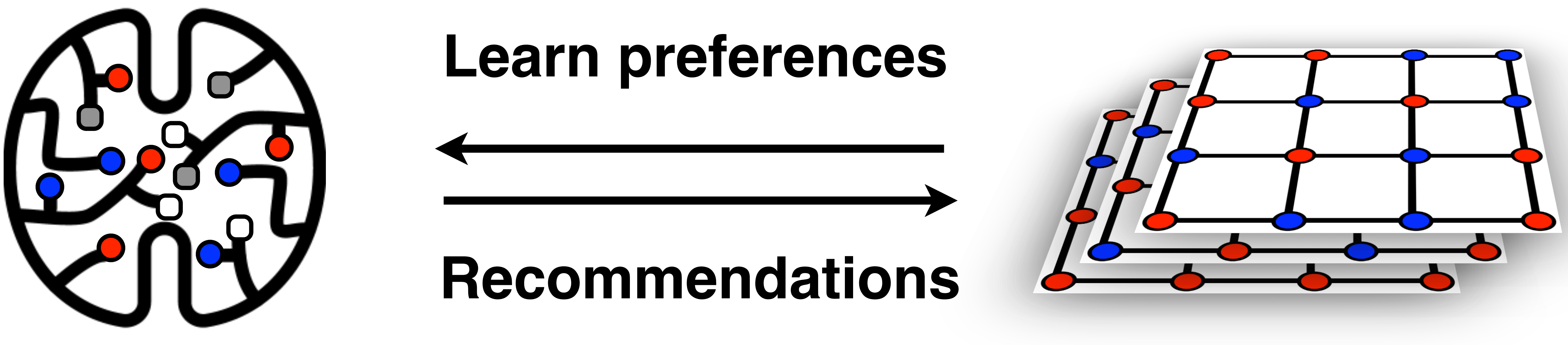
Algorithmic innovation outperforms Moore's law!

Discovering cluster updates with BM

Learn preferences

Recommendations

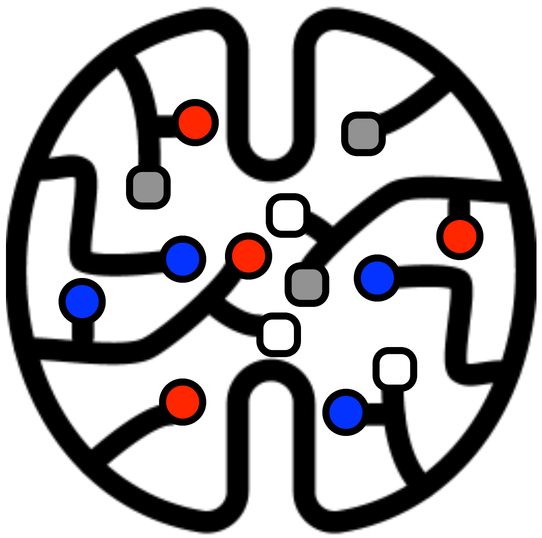
Discovering cluster updates with BM



- Use Boltzmann Machines as recommender systems for Monte Carlo simulation

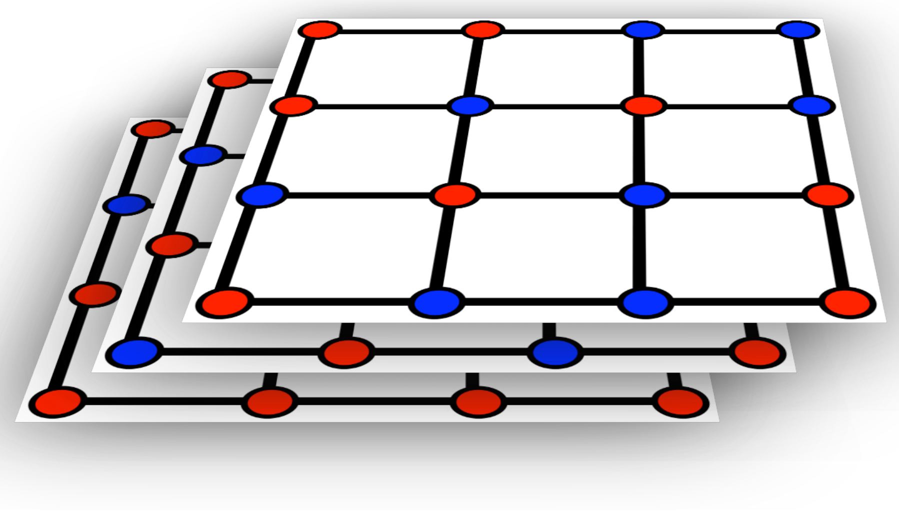
Li Huang and LW, 1610.02746

Discovering cluster updates with BM



Learn preferences

Recommendations



- Use Boltzmann Machines as recommender systems for Monte Carlo simulation

Li Huang and LW, 1610.02746

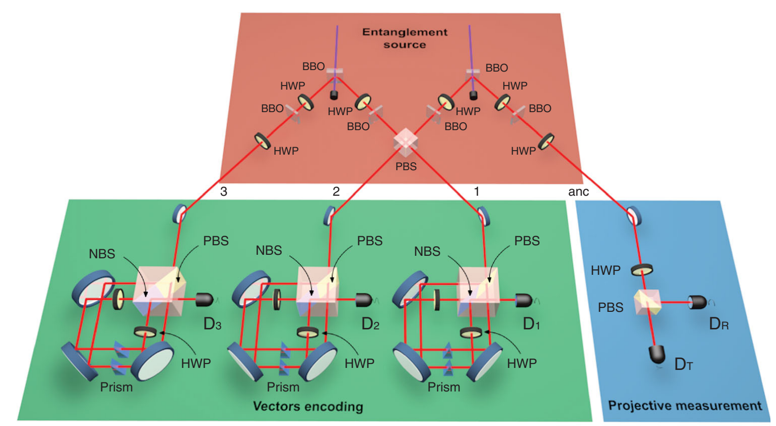
- Moreover, BM parametrizes Monte Carlo policies and explores novel algorithms!

LW, 1702.08586

Quantum Many-Body Physics for Machine Learning

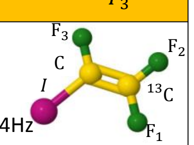
Quantum Machine Learning

- Use a quantum computer to speed up classical ML subroutines
  - Optimization
  - Linear algebra
  - Sampling
  - Clustering
  - Support vector machine
  - Principal component analysis



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)

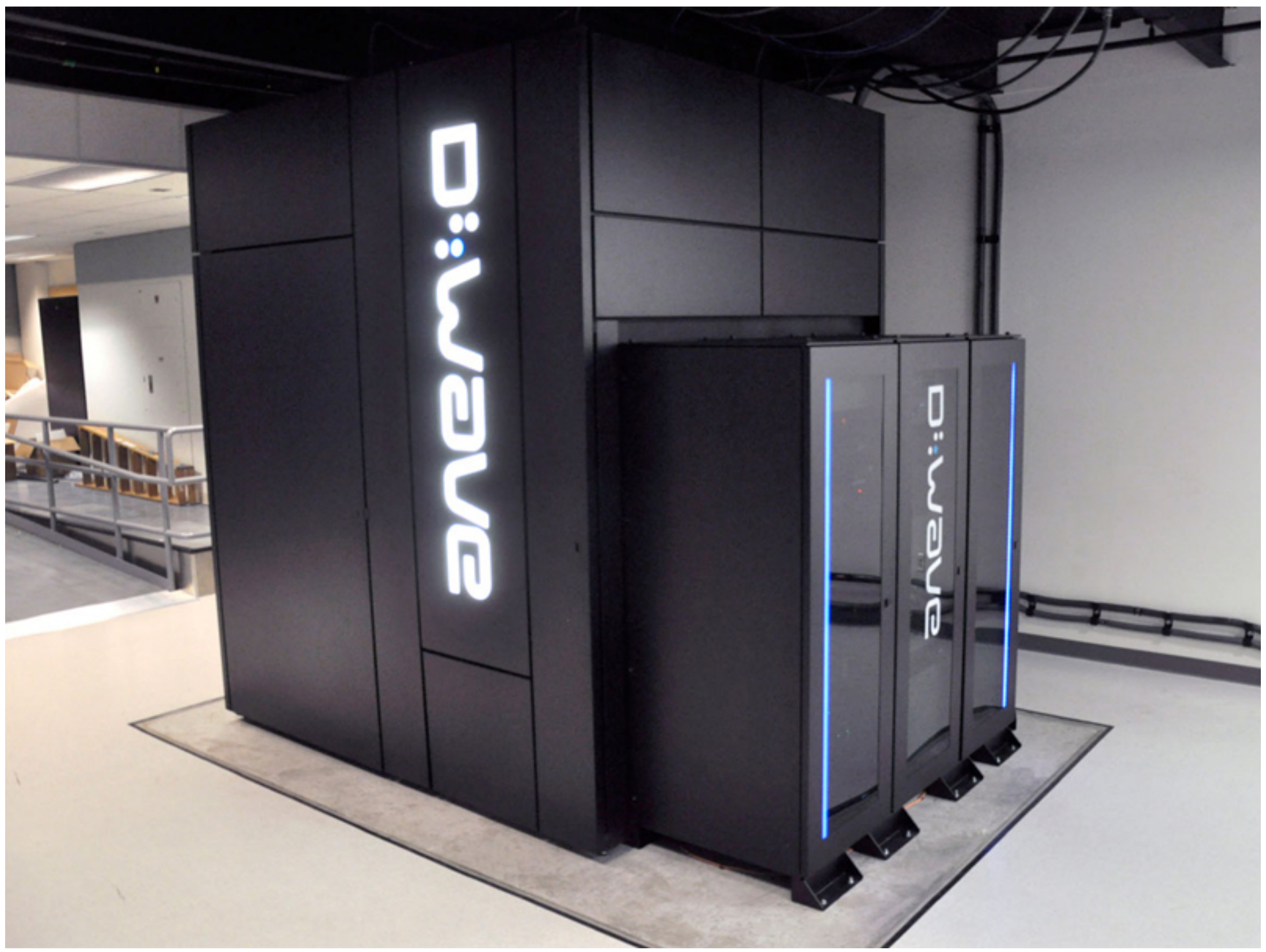
- Quantum data and quantum architecture

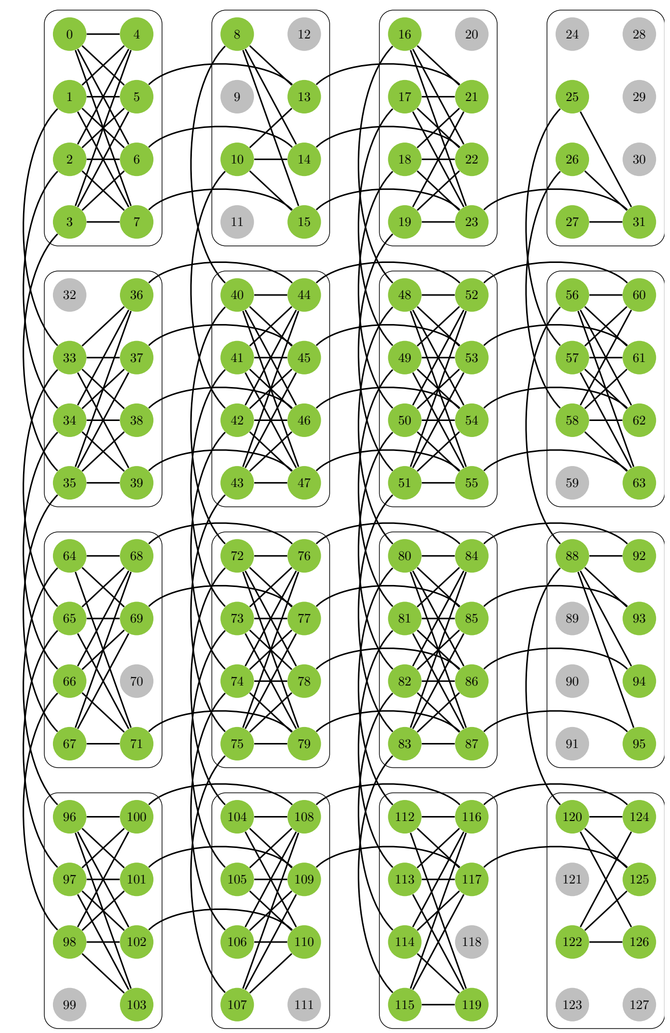
"Advances in quantum machine learning", Adcock et al, 1512.02900

"Quantum machine learning", Biamonte et al, 1611.09347

Quantum Boltzmann Machine

$15 million “quantum Ising simulator”





Is there any advantage of quantum architecture ?

Amin et al, 1601.02036

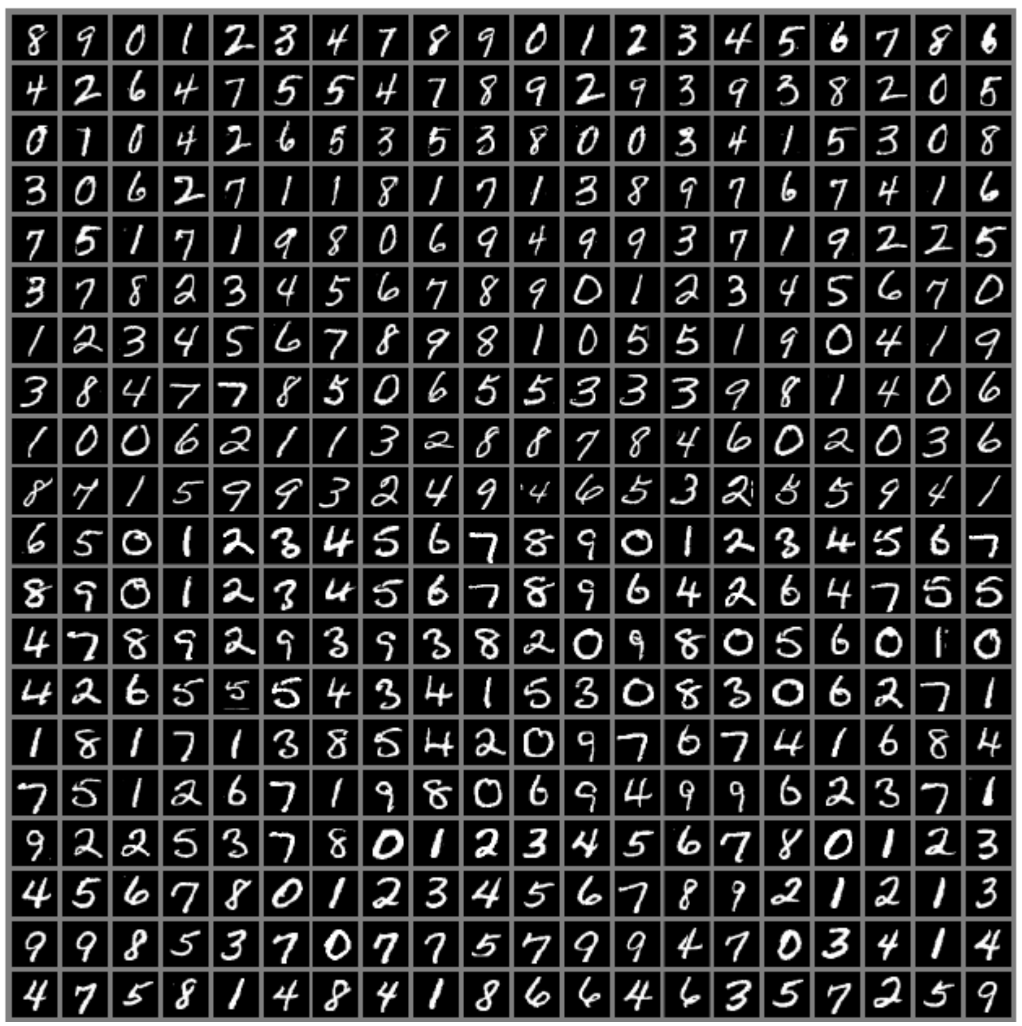
Quantum entanglement perspective on deep learning

Xun Gao, L.-M. Duan, 1701.05039

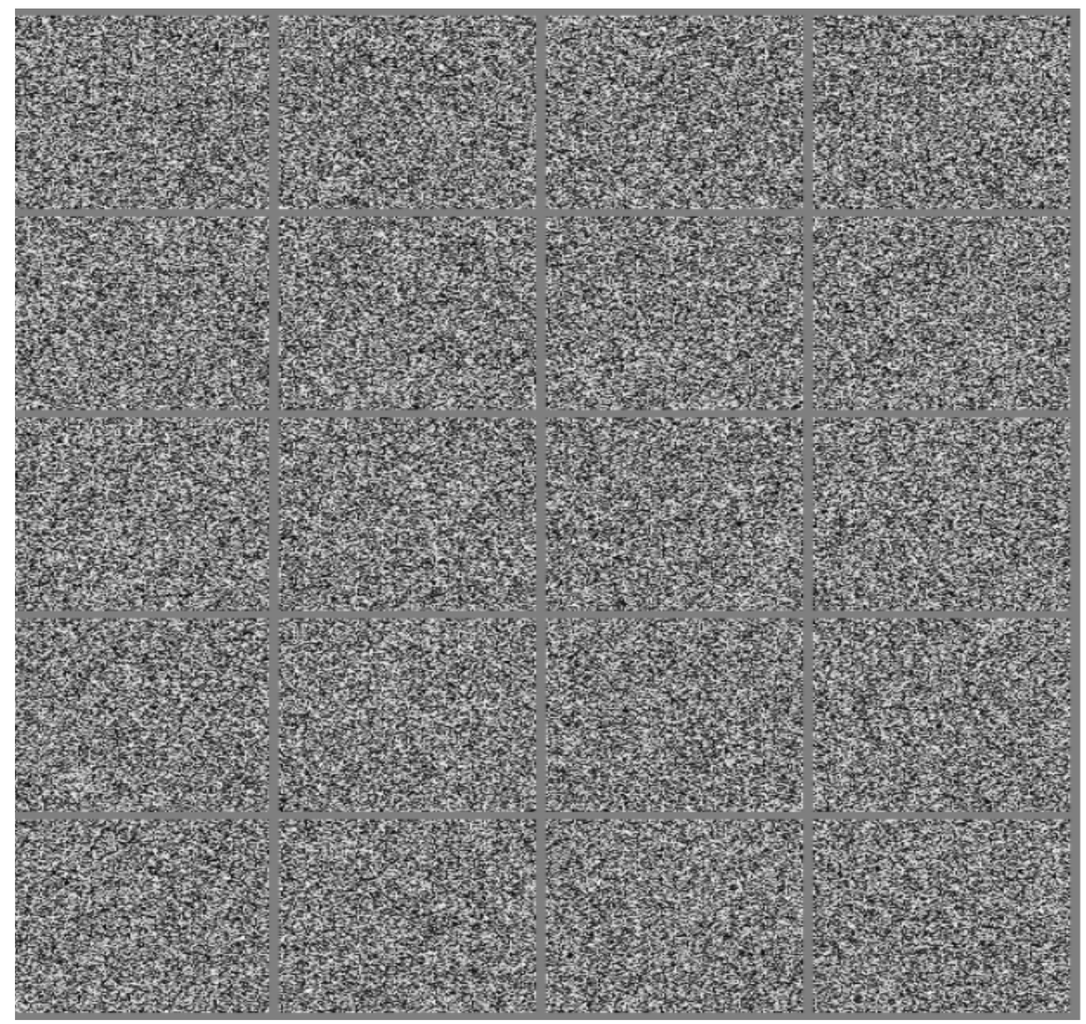
Yichen Huang and J. E. Moore, 1701.06246

Dong-Ling Deng, Xiaopeng Li and S. Das Sarma, 1701.04844

Jing Chen, Song Cheng, Haidong Xie, LW, and Tao Xiang, 1701.04831



MNIST database



random images

from the “Deep Learning” book by Goodfellow, Bengio, Courville https://www.deeplearningbook.org/

Deep Learning and Quantum Entanglement: Fundamental Connections with Implications to Network Design

Yoav Levine YOAVLEVINE@CS.HUJI.AC.IL
David Yakira DAVIDYAKIRA@CS.HUJI.AC.IL
Nadav Cohen COHENNADAV@CS.HUJI.AC.IL
Amnon Shashua SHASHUA@CS.HUJI.AC.IL
The Hebrew University of Jerusalem

arXiv:1704.01552v2 [cs.LG] 10 Apr 2017

Abstract

Deep convolutional networks have witnessed unprecedented success in various machine learning applications. Formal understanding on what makes these networks so successful is gradually unfolding, but for the most part there are still significant mysteries to unravel. The inductive bias, which reflects prior knowledge embedded in the network architecture, is one of them. In this work, we establish a fundamental connection between the fields of quantum physics and deep learning. We use this connection for asserting novel theoretical observations regarding the role that the number of channels in each layer of the convolutional network fulfills in the overall inductive bias. Specifically, we show an equivalence between the function realized by a deep convolutional arithmetic circuit (ConvAC) and a quantum many-body wave function, which relies on their common underlying tensorial structure. This facilitates the use of quantum entanglement measures as well-defined quantifiers of a deep network's expressive ability to model intricate correlation structures of its inputs. Most importantly, the construction of a deep convolutional arithmetic circuit in terms of a Tensor Network is made available. This description enables us to carry a graph-theoretic analysis of a convolutional network, tying its expressiveness to a min-cut in the graph which characterizes it. Thus, we demonstrate a direct control over the inductive bias of the designed deep convolutional network via its channel numbers, which we show to be related to the min-cut in the underlying graph. This result is relevant to any practitioner designing a convolutional network for a specific task. We theoretically analyze convolutional arithmetic circuits, and empirically validate our findings on more common convolutional networks which involve ReLU activations and max pooling. Beyond the results described above, the description of a deep convolutional network in well-defined graph-theoretic tools and the formal structural connection to quantum entanglement, are two interdisciplinary bridges that are brought forth by this work.

Thank you!

量子纠缠：从量子物质态到深度学习

程 嵩[1,2] 陈 靖[1,2] 王 磊[1,†]

(1 中国科学院物理研究所 北京 100190)

(2 中国科学院大学 北京 100049)

Quantum entanglement: from quantum states of matter to deep learning

CHENG Song[1,2] CHEN Jing[1,2] WANG Lei[1,†]

(1 Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China)

(2 University of Chinese Academy of Sciences, Beijing 100049, China)

2017－06－05收到

† email: wanglei@iphy.ac.cn

DOI: 10.7693/wl20170702

摘 要 量子纠缠在量子物质态的研究中扮演着日趋重要的角色，它可以标记传统范式难以区分的新奇量子态和量子相变，并指导设计高效的数值算法来精确地研究量子多体问题。最近，随着一些深度学习技术在量子物理问题中的应用，人们惊奇地发现：从量子纠缠的视角审视深度学习，或许有助于反过来理解和解决一些深度学习中的问题。量子纠缠定量化地刻画了现实数据集的复杂度，并指导相应的人工神经网络结构设计。沿着这个思路，物理学家们对于量子多体问题所形成的种种洞察和理论可以以一种意想不到的方式应用在现实世界中。

关键词 量子纠缠，张量网络，人工神经网络，深度学习

《物理》杂志

2017年7月刊