

Micro-workshop on AI applications to data analysis

H. Niewodniczański Institute of Nuclear Physics, Polish Academy of Sciences

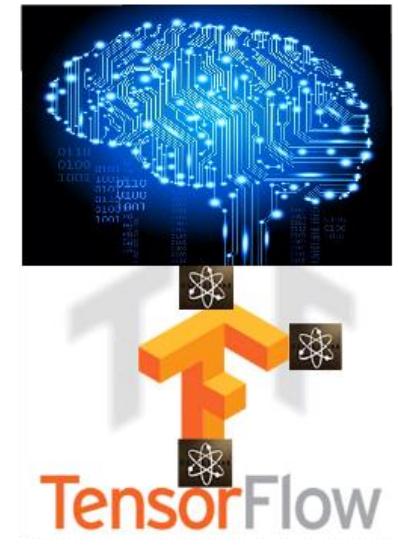
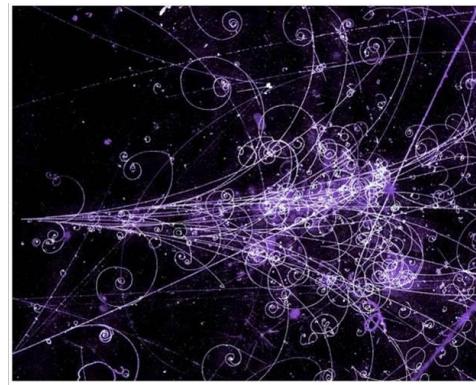
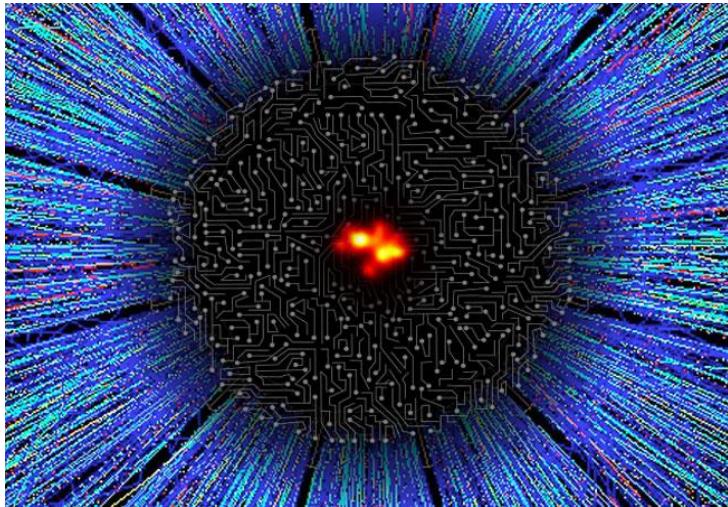
Modern AI-Methods for physics

Olena Linnyk

30.01.2019



Deep Learning bridges theory and experiment

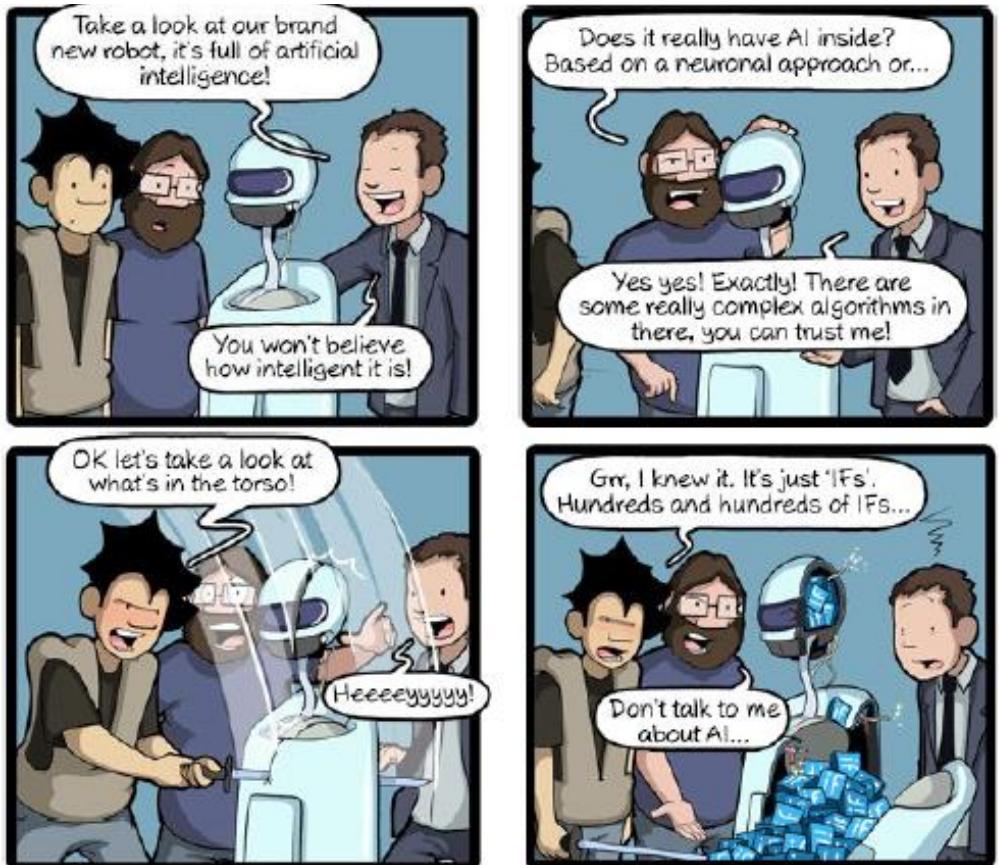


Our tools – tensorflow, keras libraries.

One can also use scikit-learn, R,
or write custom.

‘Old’ AI

- Science of
automate intellectual tasks normally performed by humans
- born in 1950s – in game of chess
 - symbolic AI, hard-code **handcraft rules** – until 1980s
- Narrow AI (currently) - can do just one or few defined things
- General AI



CommitStrip.com

,New' AI

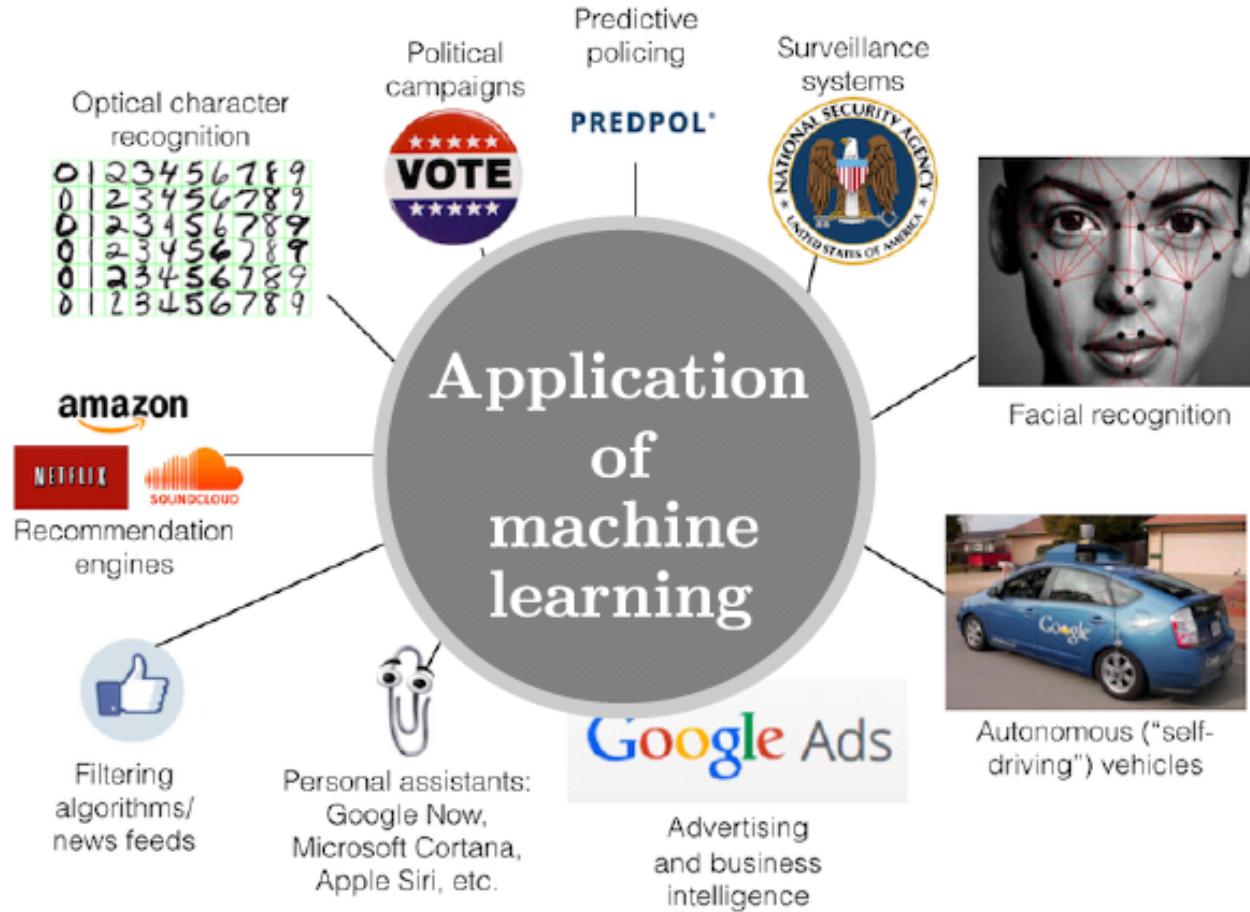
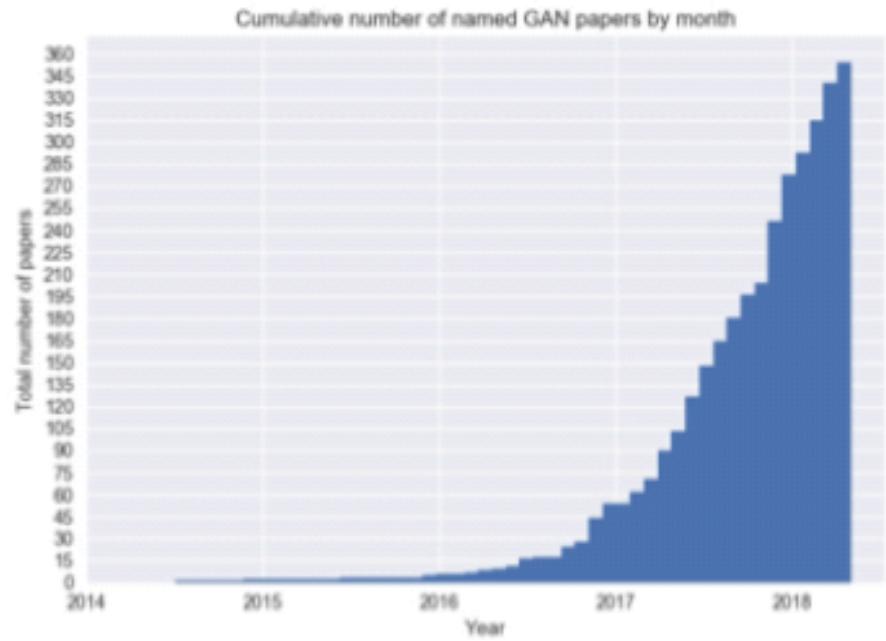
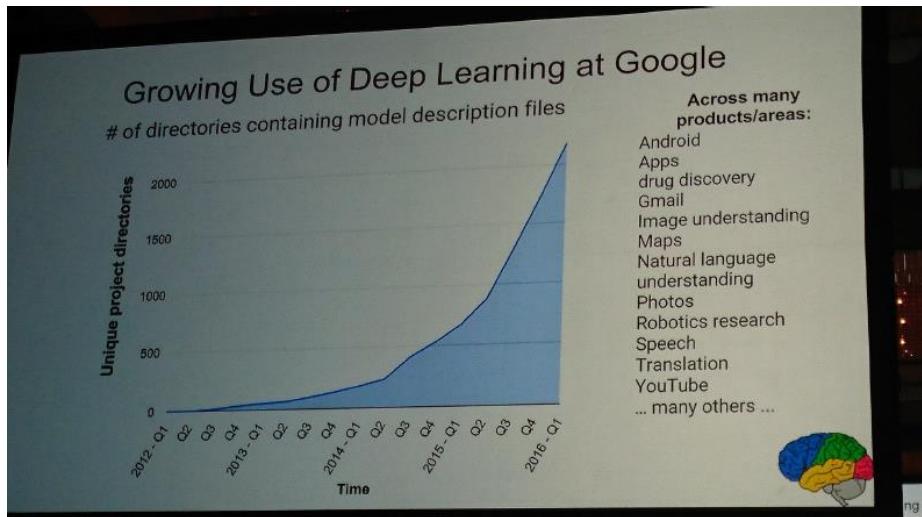


Image source: <https://redshiftzero.github.io/2015/08/29/Manipulation-and-Machine-Learning/>

Deep Learning revolution



Superhuman? ...



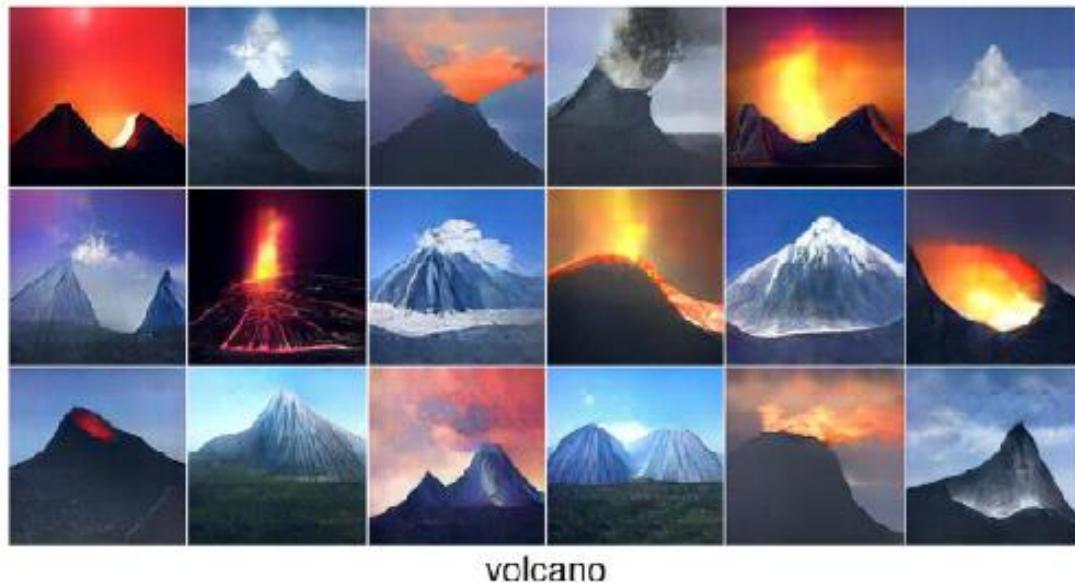
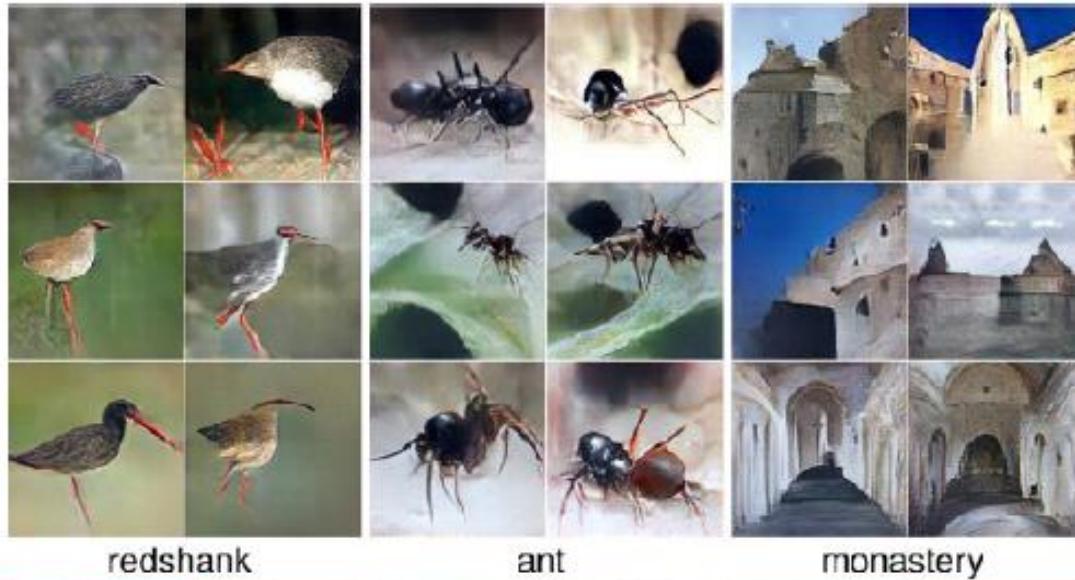
AlphaGo 4 : Lee Sedol 1
Seoul, March 2016



AlphaGo Master vs Ke Jie
Wuzhen, May 2017

Google DeepMind, London
Nature 529, 484–489 (2016)

GENERATIVE MODEL FOR IMAGES



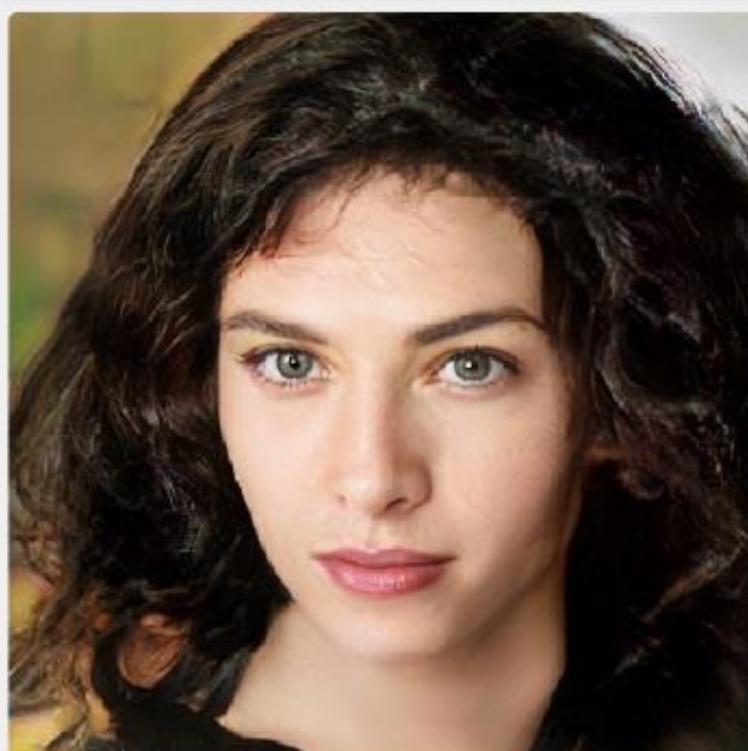
GENERATIVE MODEL FOR IMAGES

How an A.I. ‘Cat-and-Mouse Game’ Generates Believable Fake Photos

By CADE METZ and KEITH COLLINS JAN. 2, 2018



This one is computer-generated

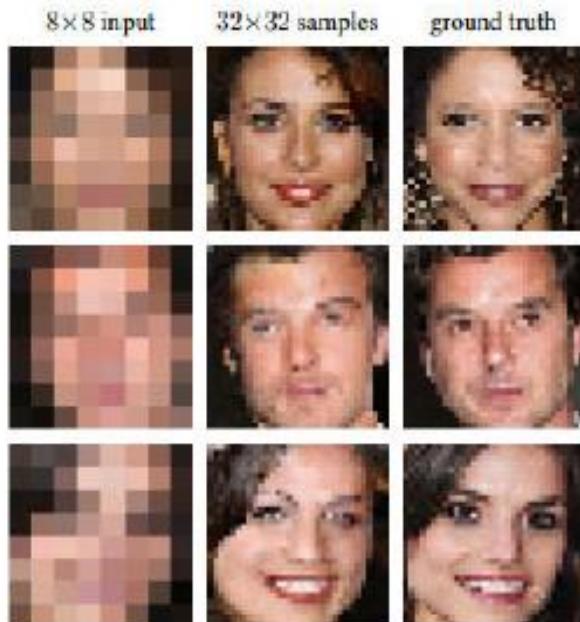


This one is also computer-generated

Generation

Image Super Resolution

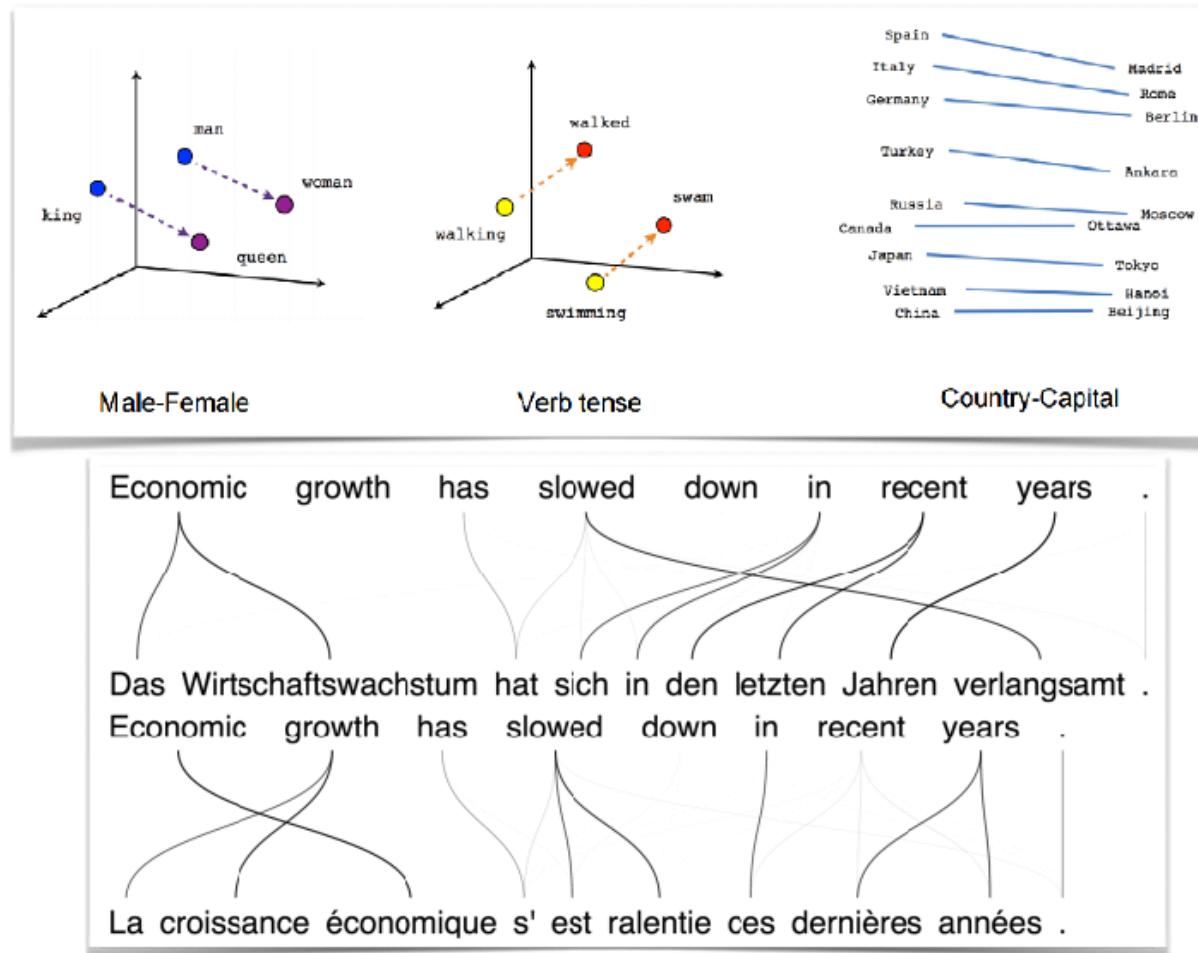
Conditional generative model $P(\text{high res image} \mid \text{low res image})$



Ledig et al., 2017

Words: two, too, to...

WORD EMBEDDINGS & TRANSLATION



Generation

Chinese Poetry Generation with Planning based Neural Network

Zhe Wang[†], Wei He[‡], Hua Wu[†], Haiyang Wu[†], Wei Li[‡], Haifeng Wang[‡], Enhong Chen[†]

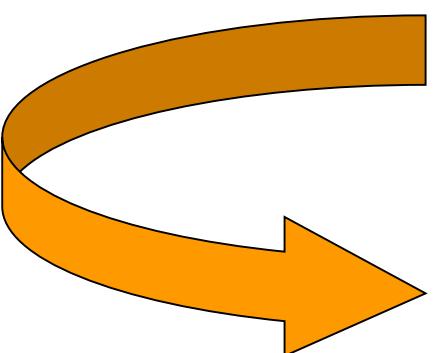
[†]University of Science and Technology of China, Hefei, China

[‡]Baidu Inc., Beijing, China

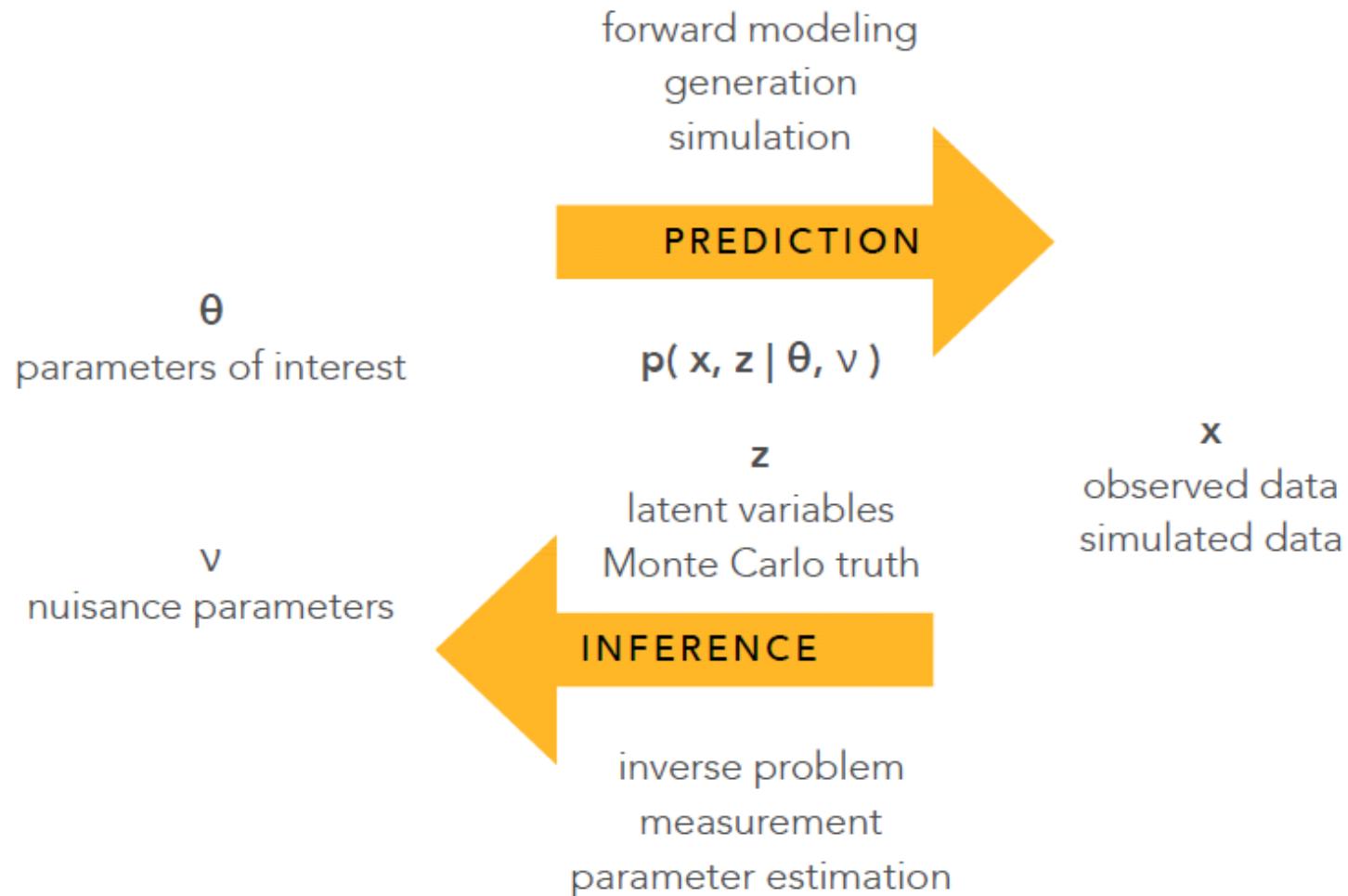
<p>秋夕湖上</p> <p>By a Lake at Autumn Sunset</p> <p>一夜秋凉雨湿衣，</p> <p>A cold autumn rain wetted my clothes last night, 西窗独坐对夕晖。</p> <p>And I sit alone by the window and enjoy the sunset. 湖波荡漾千山色，</p> <p>With mountain scenery mirrored on the rippling lake, 山鸟徘徊万籁微。</p> <p>A silence prevails over all except the hovering birds.</p>	<p>秋夕湖上</p> <p>By a Lake at Autumn Sunset</p> <p>获花风里桂花浮，</p> <p>The wind blows reeds with osmanthus flying, 恨竹生云翠欲流。</p> <p>And the bamboos under clouds are so green as if to flow down. 谁拂半湖新镜面，</p> <p>The misty rain ripples the smooth surface of lake, 飞来烟雨暮天愁。</p> <p>And I feel blue at sunset .</p>
---	---

Table 6: A pair of poems selected from the blind test. The left one is a machine-generated poem, and the right one is written by Shaoti Ge, a poet lived in the Song Dynasty.

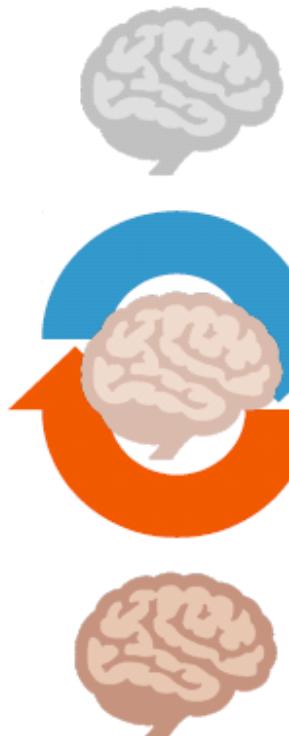
What can AI for physics?

1. Classification, Visualisation
 2. Optimisation, Regression
 3. Generation
 4. Dimensionality reduction
- 
1. Fitting, Prediction
 2. Interpolation, Extrapolation
 3. Simulation (Emulation)
 4. Huge-Dimensional Distributions

What can AI?



Fitting



1. Pick an appropriate brain (curve-fitting function)



while

2a. Teach the brain under supervision

2b. Testing



3. Brain learned the best-fit model when error is minimized

NN = A HIGHLY FLEXIBLE FAMILY OF FUNCTIONS

In calculus of variations, the optimization is over all functions: $\hat{s} = \operatorname{argmin}_s L[s]$ —

- In applied calculus of variations, we consider a highly flexible family of functions s_ϕ and optimize: i.e. $\hat{\phi} = \operatorname{argmin}_\phi L[s_\phi]$ and $\hat{s} \approx s_{\hat{\phi}}$
- Think of neural networks as a highly flexible family of functions
- Machine learning also includes non-convex optimization algorithms that are effective even with millions of parameters!

Shallow neural network

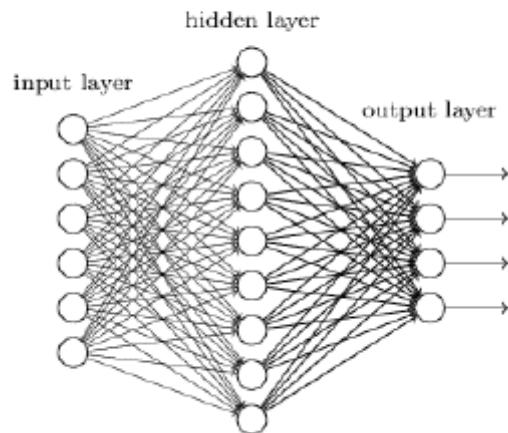
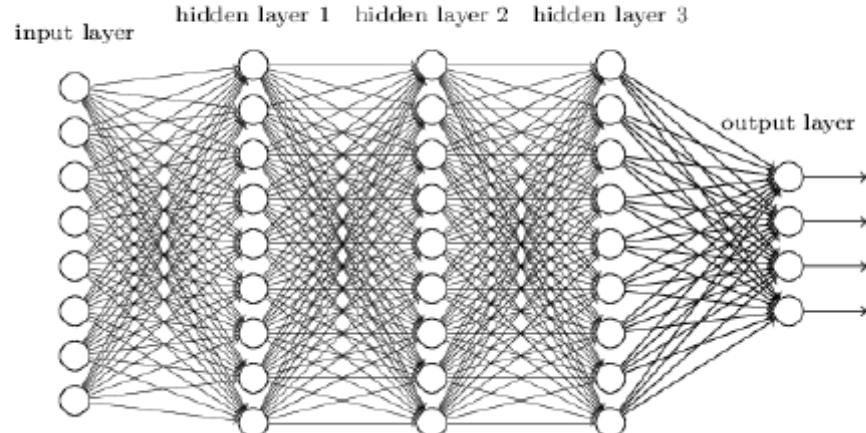
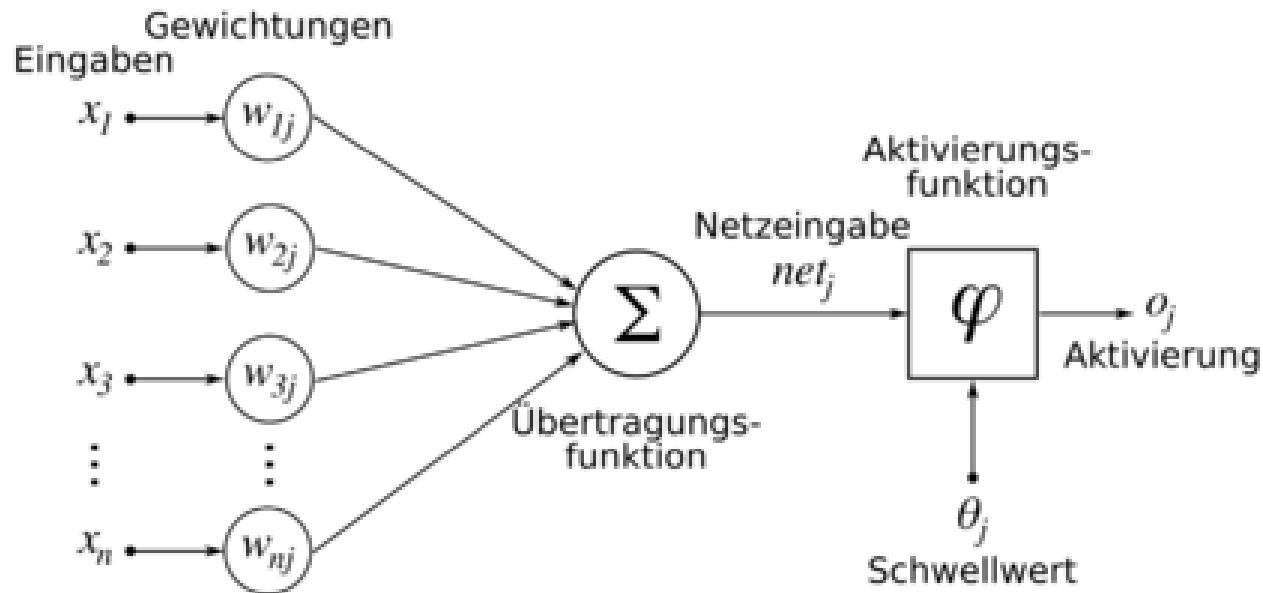
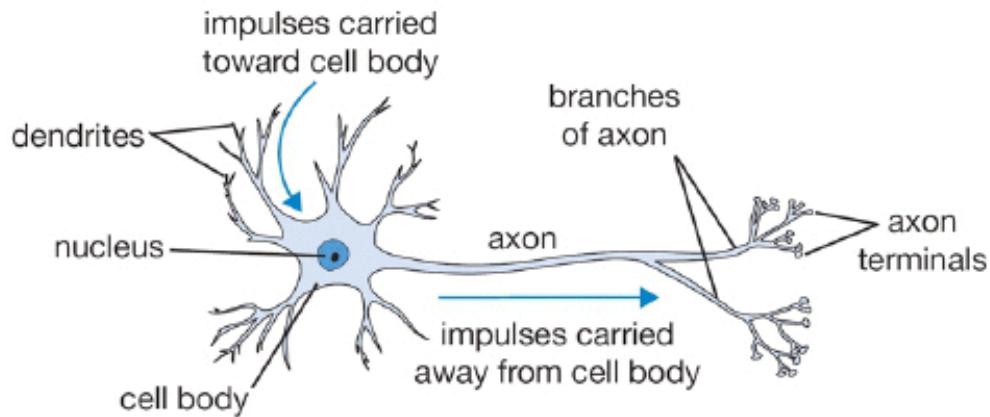


image credit: Michael Nielsen

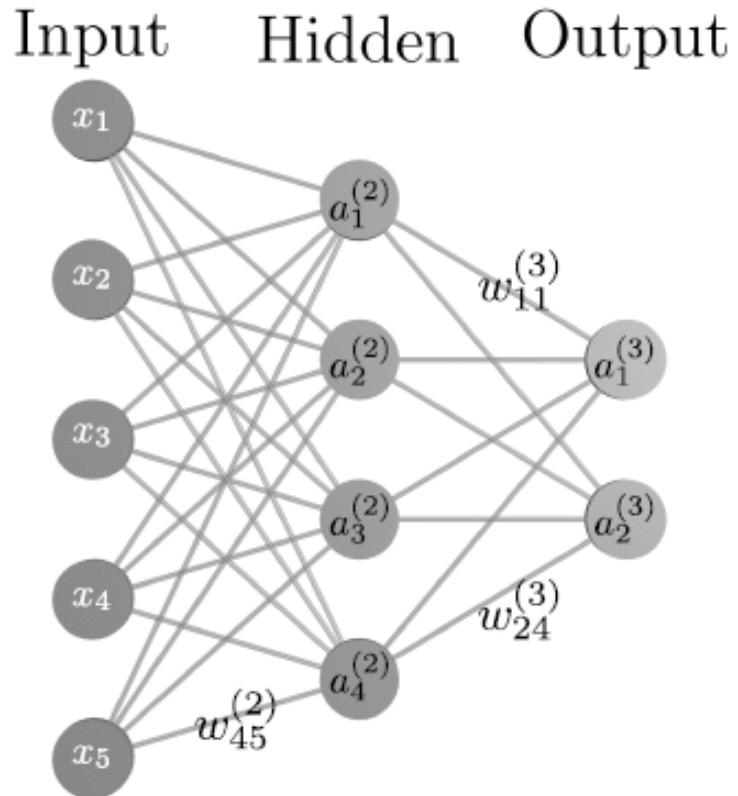
Deep neural network



Neuron



Layers, Back-propagation



$w_{jk}^{(l)}$ is the weight from the k^{th} neuron in the $(l - 1)^{th}$ layer to the j^{th} neuron in the l^{th} layer

Perform SGD to get ΔC and backpropagate errors from all previous layers.

IMAGE CLASSIFICATION

INPUT CONV POOL CONV POOL FC OUTPUT



Dog: 94%

Cat: 31%

Bird: 2%

Boat: 0%



Dog: 37%

Cat: 91%

Bird: 21%

Boat: 1%

2012

ImageNet Classification with Deep Convolutional Neural Networks

Alex Krizhevsky
University of Toronto
kriz@cs.utoronto.ca

Ilya Sutskever
University of Toronto
ilya@cs.utoronto.ca

Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca

Abstract

We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 error rates of 37.5% and 17.0% which is considerably better than the previous state-of-the-art. The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation. To reduce overfitting in the fully-connected layers we employed a recently-developed regularization method called “dropout” that proved to be very effective. We also entered a variant of this model in the ILSVRC-2012 competition and achieved a winning top-5 test error rate of 15.3%, compared to 26.2% achieved by the second-best entry.

2015

News & Analysis

Microsoft, Google Beat Humans at Image Recognition

Deep learning algorithms compete at ImageNet challenge

R. Colin Johnson
2/18/2015 03:15 AM EST
14 comments

1 saves
LOGIN TO RATE

Classification

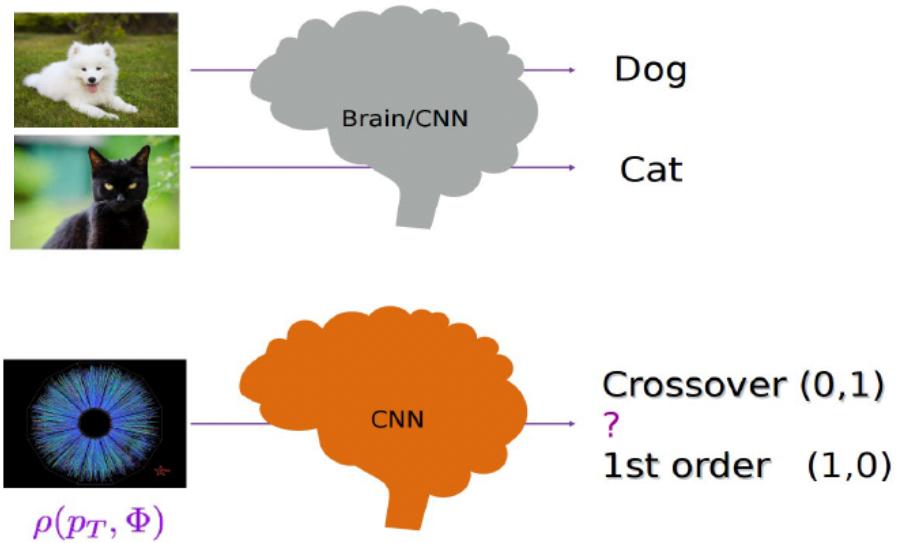


Article | OPEN | Published: 15 January 2018

An equation-of-state-meter of quantum chromodynamics transition from deep learning

Long-Gang Pang , Kai Zhou , Nan Su , Hannah Petersen, Horst Stöcker & Xin-Nian Wang

Nature Communications 9, Article number: 210 (2018) | Download Citation



Nature Communication 9, 210(2018)

AlphaStar: Mastering the Real-Time Strategy Game StarCraft II



► AlphaStar: The inside story



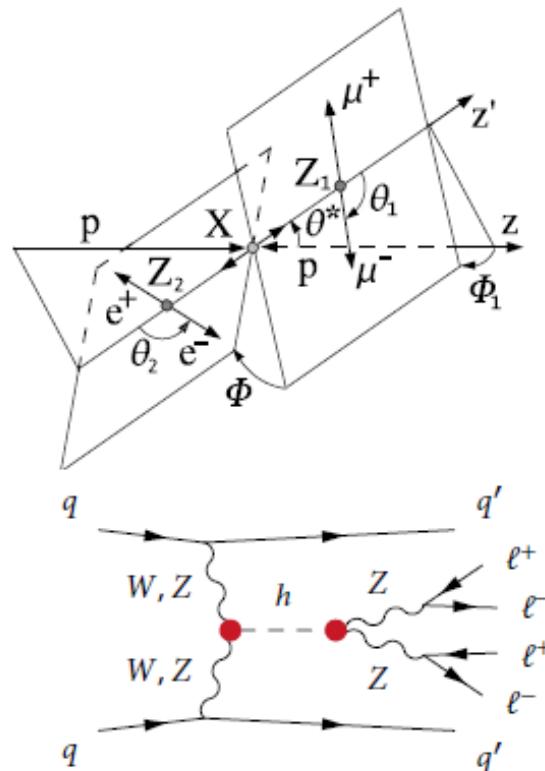
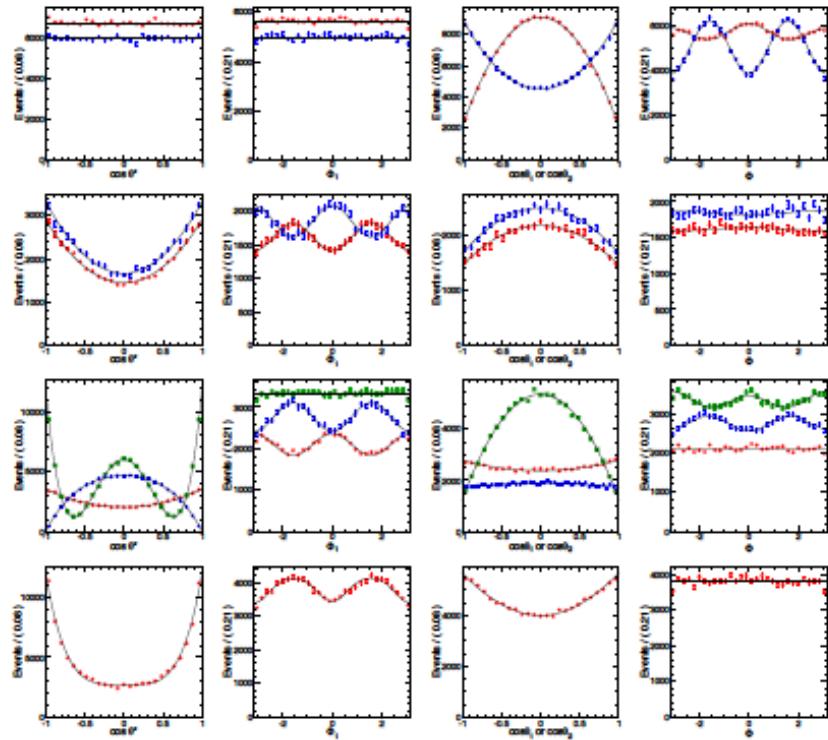
Dimentionality curse ->
dimentionality blessing



HIGH DIMENSIONAL EXAMPLE

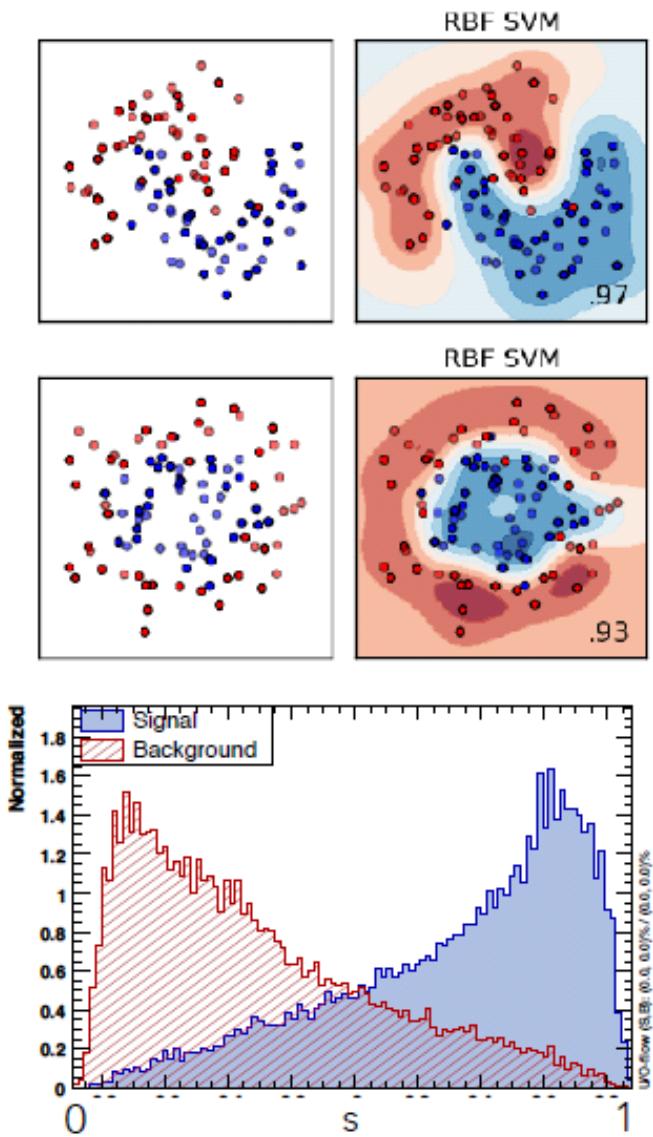
For instance, when looking for deviations from the standard model Higgs, we would like to look at all sorts of kinematic correlations

- thus each observation \mathbf{x} is high-dimensional



@KyleCranmer
New York University
Department of Physics
Center for Data Science
CILVR Lab

MACHINE LEARNING: CLASSIFIERS

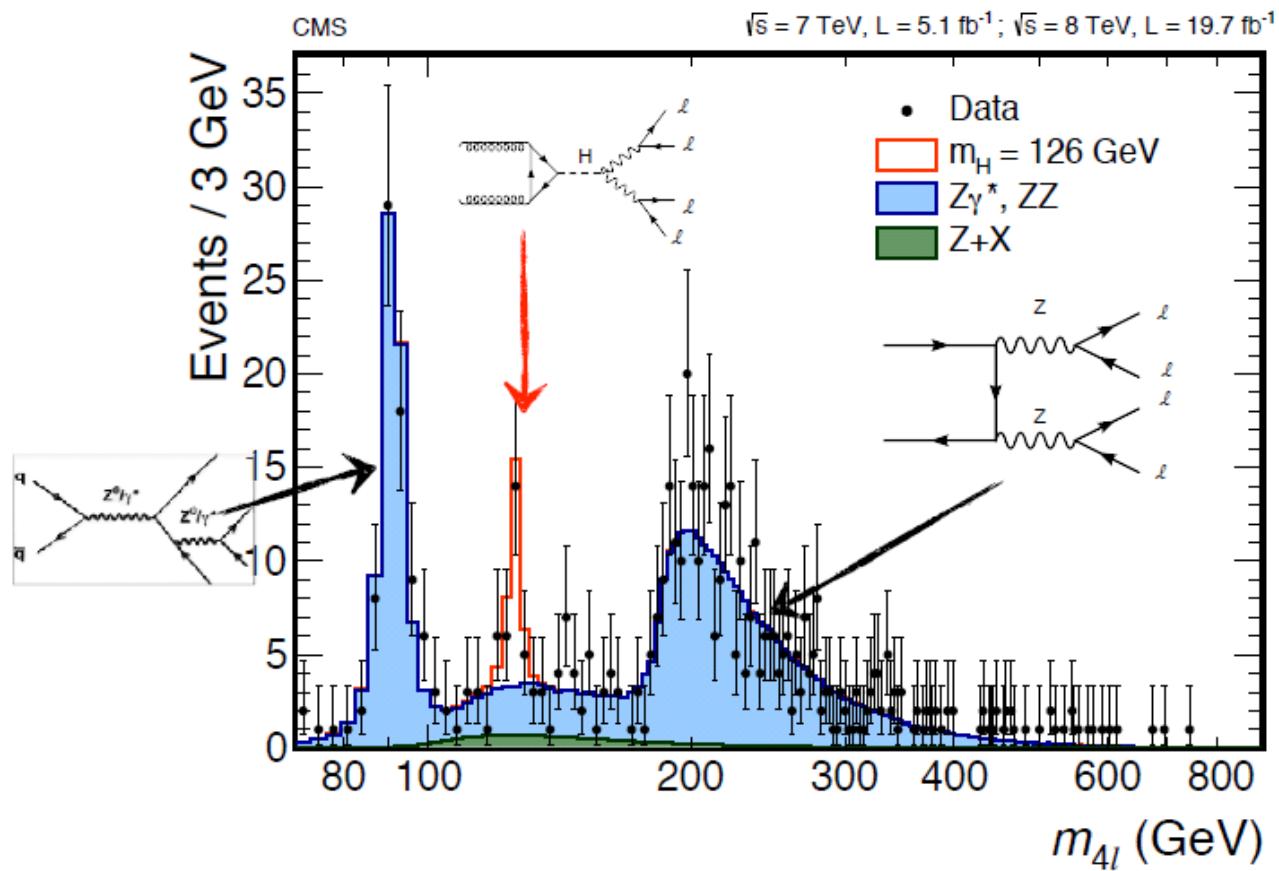


Common to use machine learning classifiers to separate signal (H_1) vs. background (H_0)

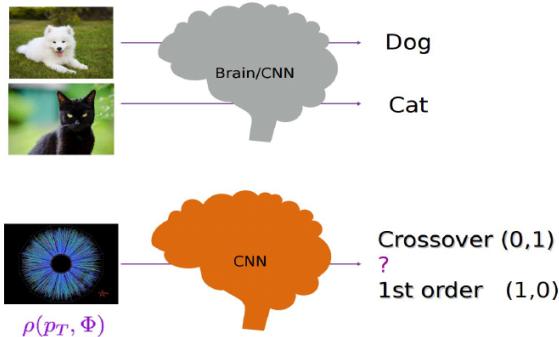
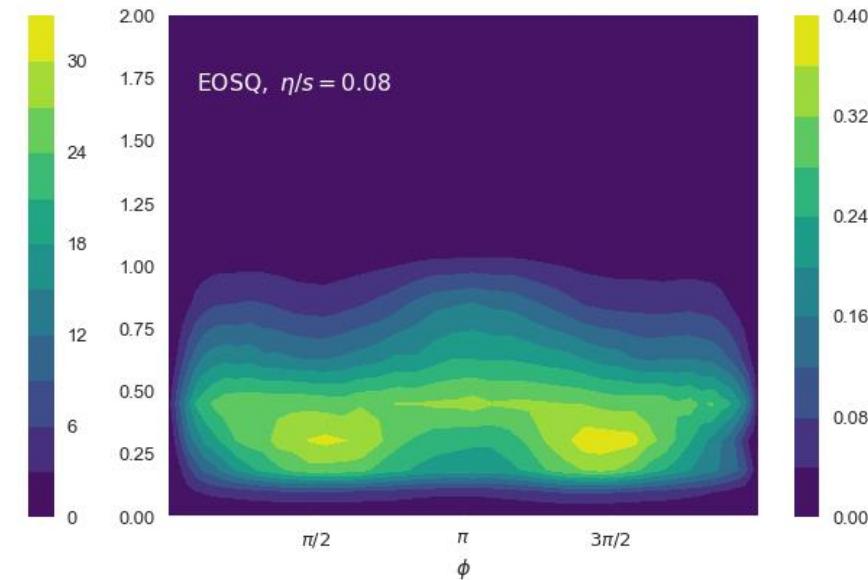
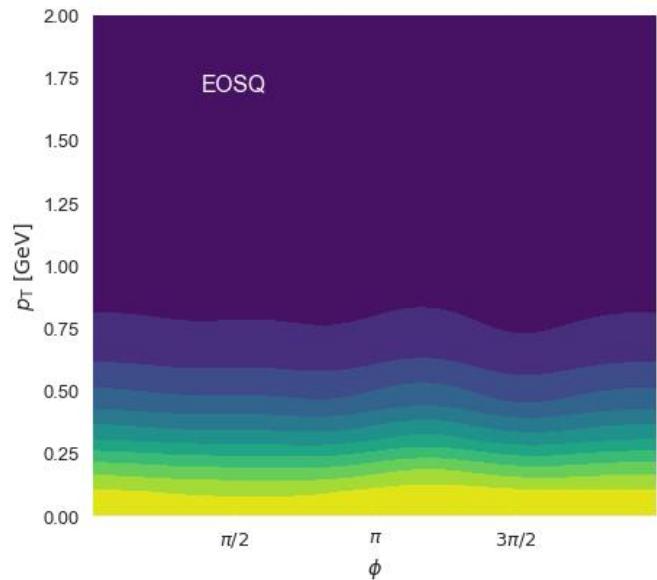
- want a function $s: X \rightarrow Y$ that maps signal to $y=1$ and background to $y=0$
- **calculus of variations:** find function $s(x)$ that minimizes **loss**:

$$L[s] = \int p(x|H_0) (0 - s(x))^2 dx + \int p(x|H_1) (1 - s(x))^2 dx$$

Example: Signal vs Background



Classification, importance map



nature
COMMUNICATIONS

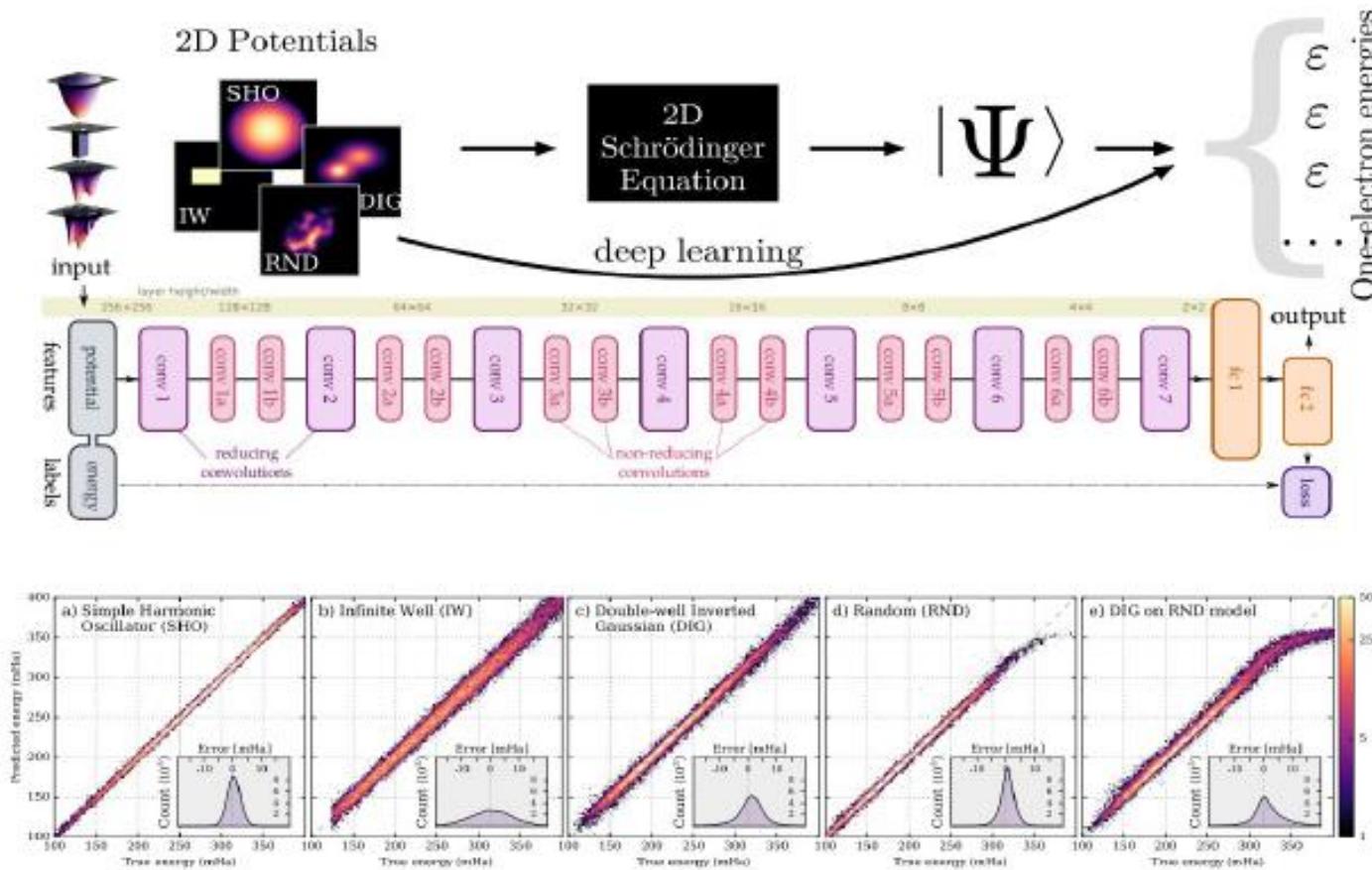
Article | OPEN | Published: 15 January 2018

An equation-of-state-meter of quantum chromodynamics transition from deep learning

Long-Gang Pang , Kai Zhou , Nan Su , Hannah Petersen, Horst Stöcker & Xin-Nian Wang

Nature Communications 9, Article number: 210 (2018) | Download Citation

Emulation

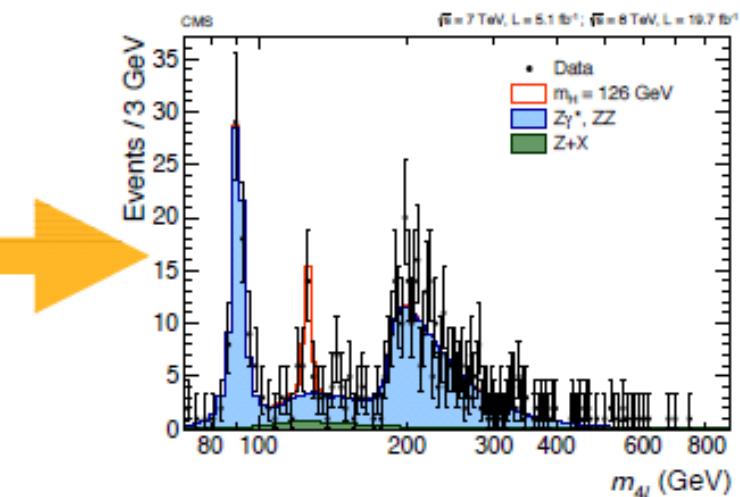
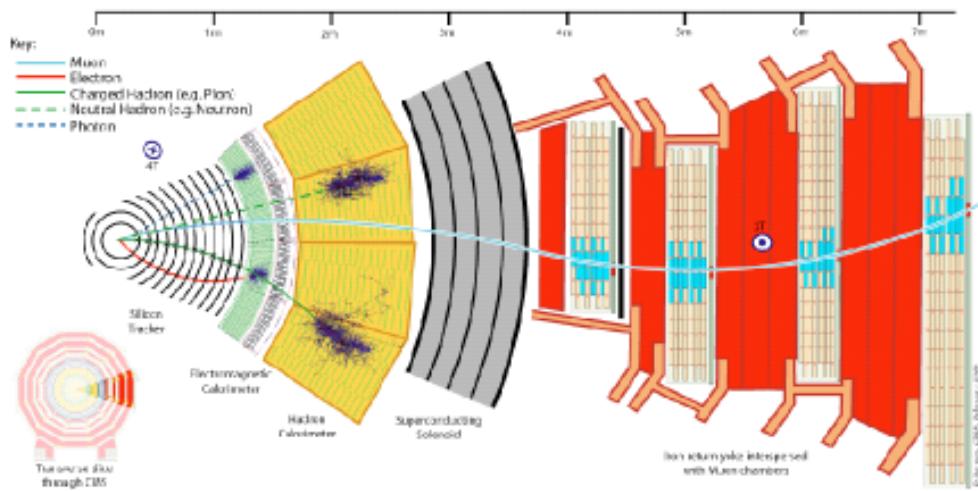


[Deep learning and the Schrodinger equation](#), by K. Mills, M. Spanner, Tamblyn (February 7, 2017)

10^8 SENSORS \rightarrow 1 REAL-VALUED QUANTITY

Most measurements and searches for new particles at the LHC are based on the distribution of a single variable / feature / summary statistic

- choosing a good variable (feature engineering) is a task for a skilled physicist and tailored to the goal of measurement or new particle search
- likelihood $p(x|\theta)$ approximated using histograms (univariate density estimation)



THE FORWARD MODEL

$$\mathcal{L}_{SM} = \frac{1}{4} \mathbf{W}_{\mu\nu} \cdot \mathbf{W}^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}$$

kinetic energies and self-interactions of the gauge bosons

$$+ \underbrace{L \gamma^\mu (\partial_\mu - \frac{1}{2} g \tau \cdot \mathbf{W}_\mu - \frac{1}{2} g' Y B_\mu) L + R \gamma^\mu (\partial_\mu - \frac{1}{2} g' Y B_\mu) R}_{\text{kinetic energies and framework interactions of fermions}}$$

kinetic energies and framework interactions of fermions

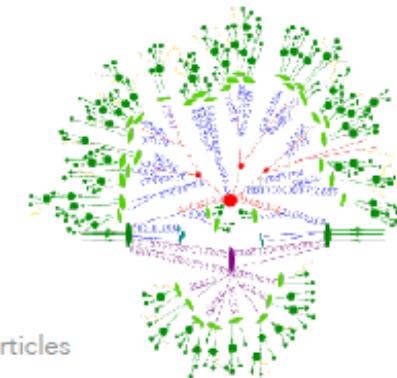
$$+ \underbrace{\frac{1}{2} |(\partial_\mu - \frac{1}{2} g \tau \cdot \mathbf{W}_\mu - \frac{1}{2} g' Y B_\mu) \phi|^2 - V(\phi)}_{W^\pm, Z, \gamma \text{ and Higgs masses and couplings}}$$

W[±], Z, γ and Higgs masses and couplings

$$+ \underbrace{g'' (\bar{q} \gamma^\mu T_d q) C_\mu^a}_{\text{interactions between quarks and gluons}} + \underbrace{(G_1 L \phi R + G_2 L \phi_c R + h.c.)}_{\text{fermion masses and couplings to Higgs}}$$

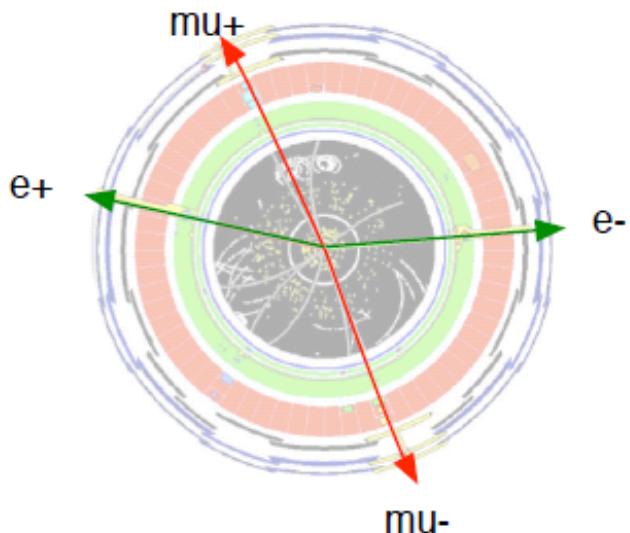
interactions between quarks and gluons fermion masses and couplings to Higgs

1) We begin with Quantum Field Theory



2) Theory gives detailed prediction for high-energy collisions

hierarchical: $2 \rightarrow O(10) \rightarrow O(100)$ particles

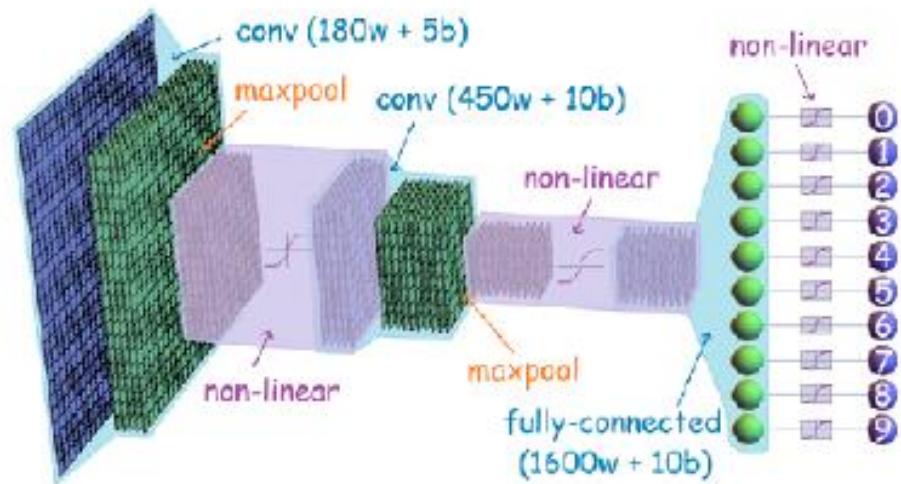
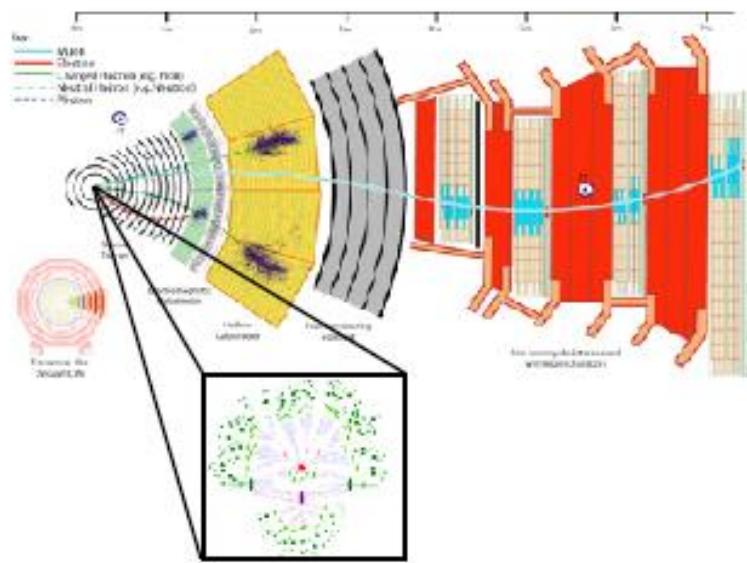


3) The interaction of outgoing particles with the detector is simulated.

>100 million sensors

4) Finally, we run particle identification and feature extraction algorithms on the simulated data as if they were from real collisions.

~10-30 features describe interesting part



GANS FOR PHYSICS

CaloGAN: Simulating 3D High Energy Particle Showers in Multi-Layer Electromagnetic Calorimeters with Generative Adversarial Networks

Michela Paganini^{a,b}, Luke de Oliveira^a, and Benjamin Nachman^a

^a*Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA, 94720, USA*

^b*Department of Physics, Yale University, New Haven, CT 06520, USA*

E-mail: michela.paganini@yale.edu, lukedoliveira@lbl.gov, bnachman@cern.ch

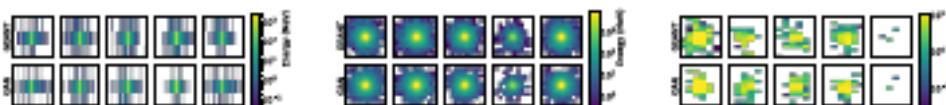


Figure 9: Five randomly selected e^+ showers per calorimeter layer from the training set (top) and the five nearest neighbors (by euclidean distance) from a set of CALOGAN candidates.

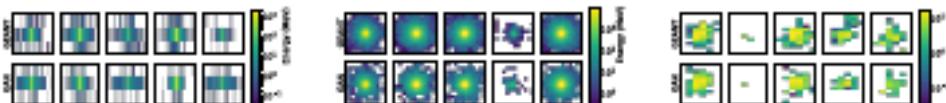


Figure 10: Five randomly selected γ showers per calorimeter layer from the training set (top) and the five nearest neighbors (by euclidean distance) from a set of CALOGAN candidates.

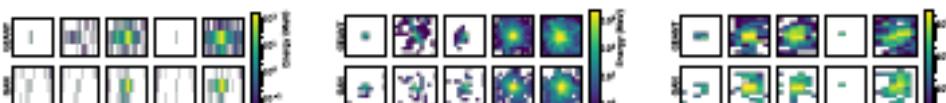
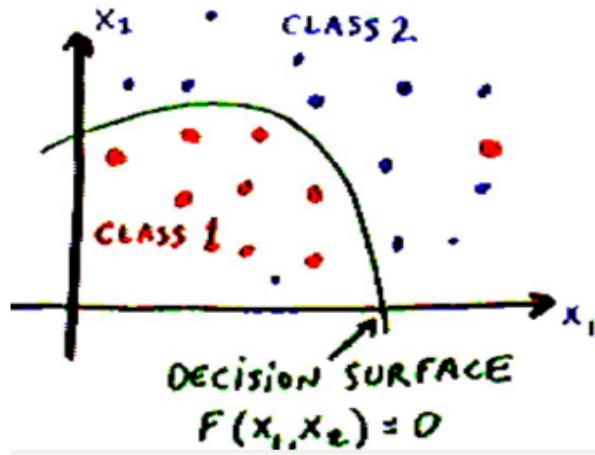


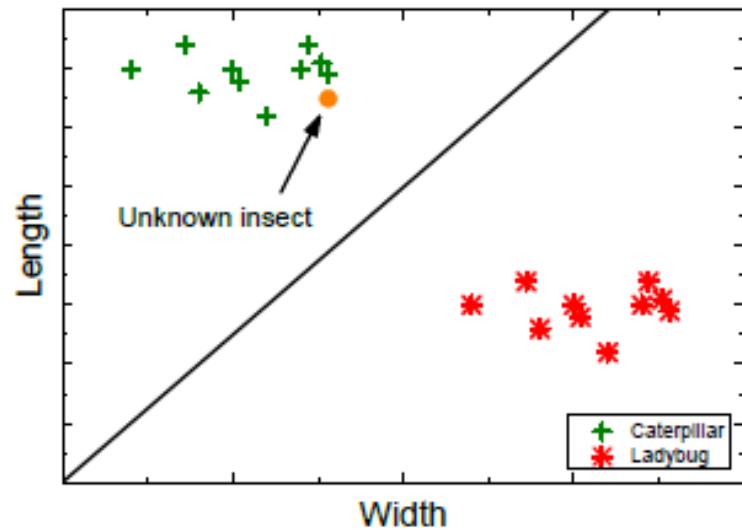
Figure 11: Five randomly selected π^+ showers per calorimeter layer from the training set (top) and the five nearest neighbors (by euclidean distance) from a set of CALOGAN candidates.

Classification supervised

Classification (discrete output/label - **classes**)



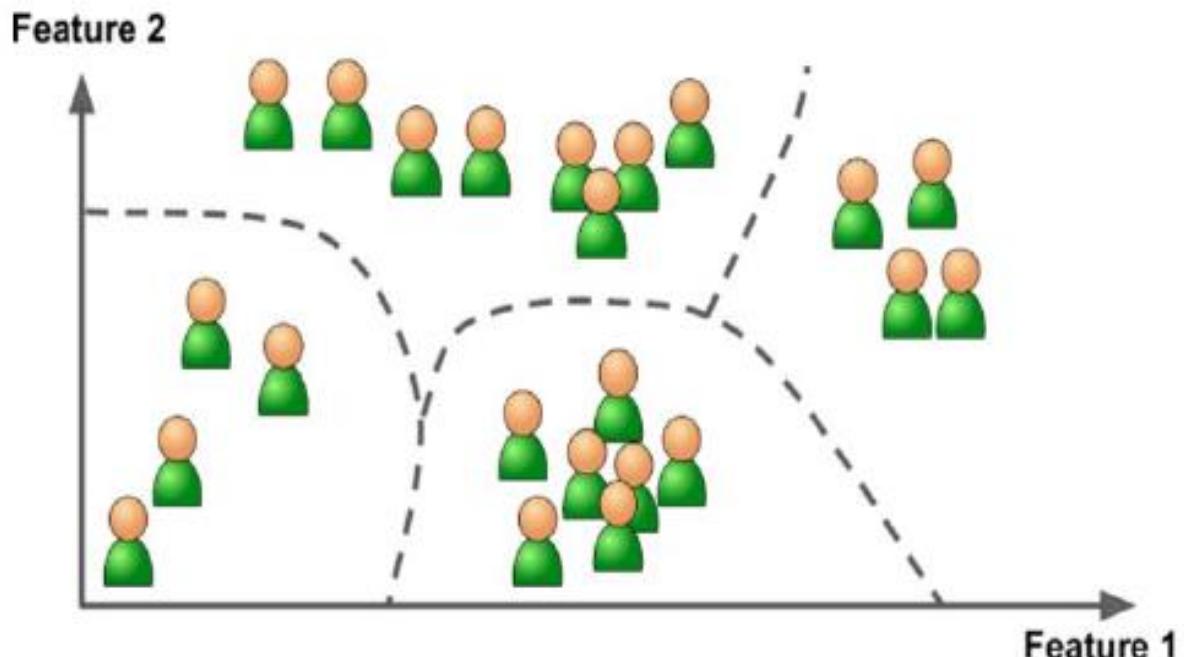
$$f : \mathbb{R}^n \rightarrow \{1, \dots, k\}$$



Clustering, unsupervised

Clustering (group dataset into clusters in **similar manner**)

- K-Means
- HCA
- Gaussian Mixture



Classification, semi-supervised

Some photo-hosting services, such as Google Photos, are good examples of this. Once you upload all your family photos to the service, it automatically recognizes that the same person A shows up in photos 1, 5, and 11, while another person B shows up in photos 2, 5, and 7. This is the unsupervised part of the algorithm (clustering). Now all the system needs is for you to tell it who these people are. Just one label per person,⁴ and it is able to name everyone in every photo, which is useful for searching photos.

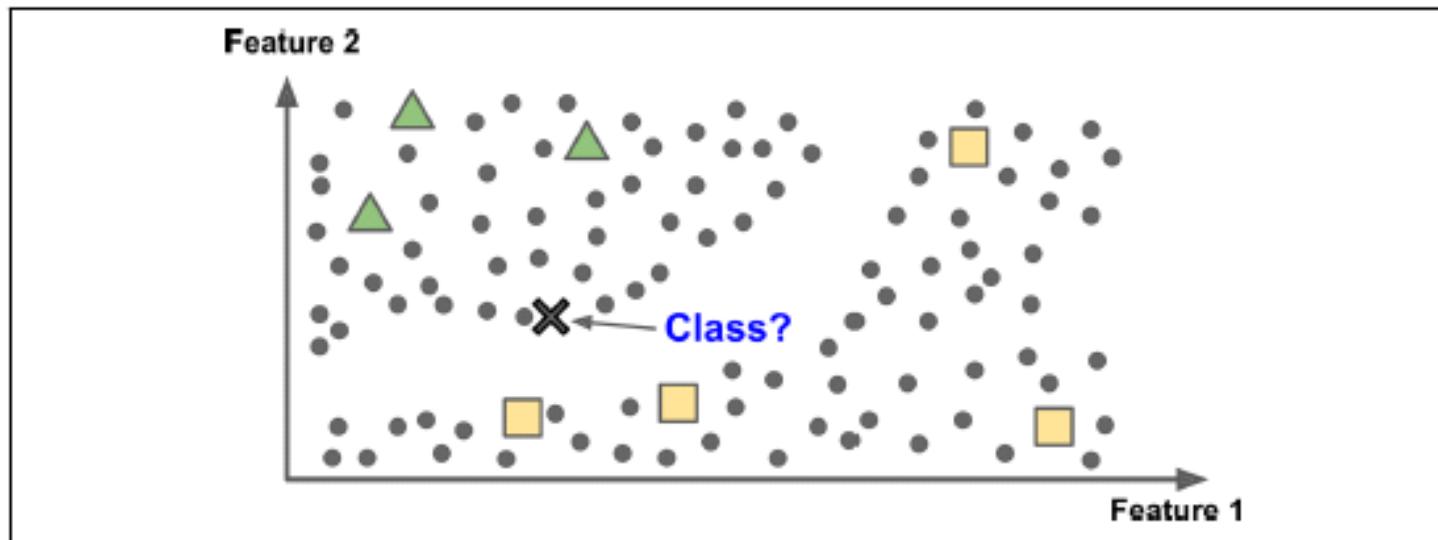
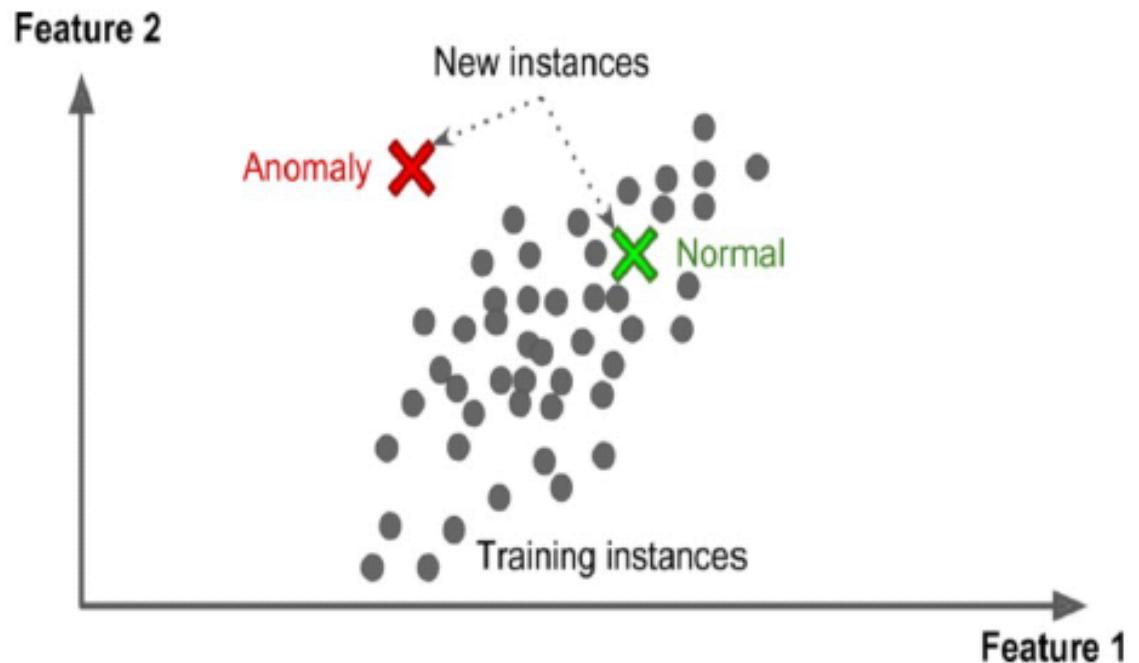


Figure 1-11. Semisupervised learning

Analyse von Daten

Anomaly detection (find out **outliers**)

- Autoencoder
- Density estimation
- KNN, k-means



Grain of salt

[Back](#) labradoodle or fried chicken [Select](#)



[Albums](#) chihuahua or muffin [Select](#)



Grain of salt

[Back](#) labradoodle or fried chicken [Select](#)



[Albums](#) chihuahua or muffin [Select](#)



Hausaufgaben

- <https://www.python.org/downloads/>

Linux:

- pip3 install --upgrade pip
 - pip3 install jupyter
 - pip3 install tensorflow, keras
- virtualenv?

Windows:

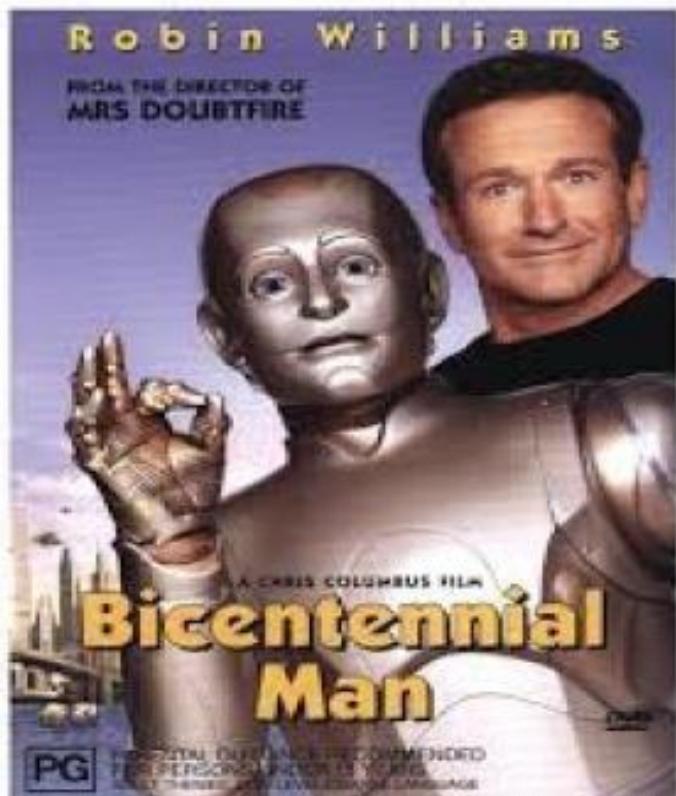
- <https://www.anaconda.com/download/>
- in the visual interface: tensorflow, keras

Codes:

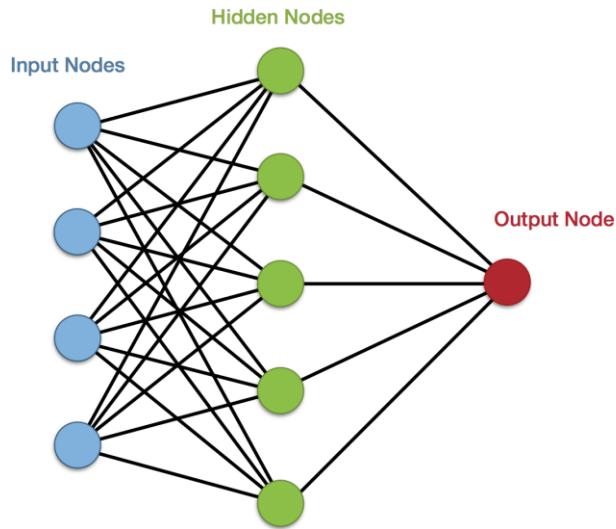
- <https://github.com/ageron/handson-ml>

Darüber ist das nicht

Ethische Fragestellungen, Super-Intelligence



Example:



Dimentionality reduction

Mask R-CNN Kaiming He, Georgia Gkioxari, Piotr Dollár, Ross Girshick, 2017.04.05



Visualisation

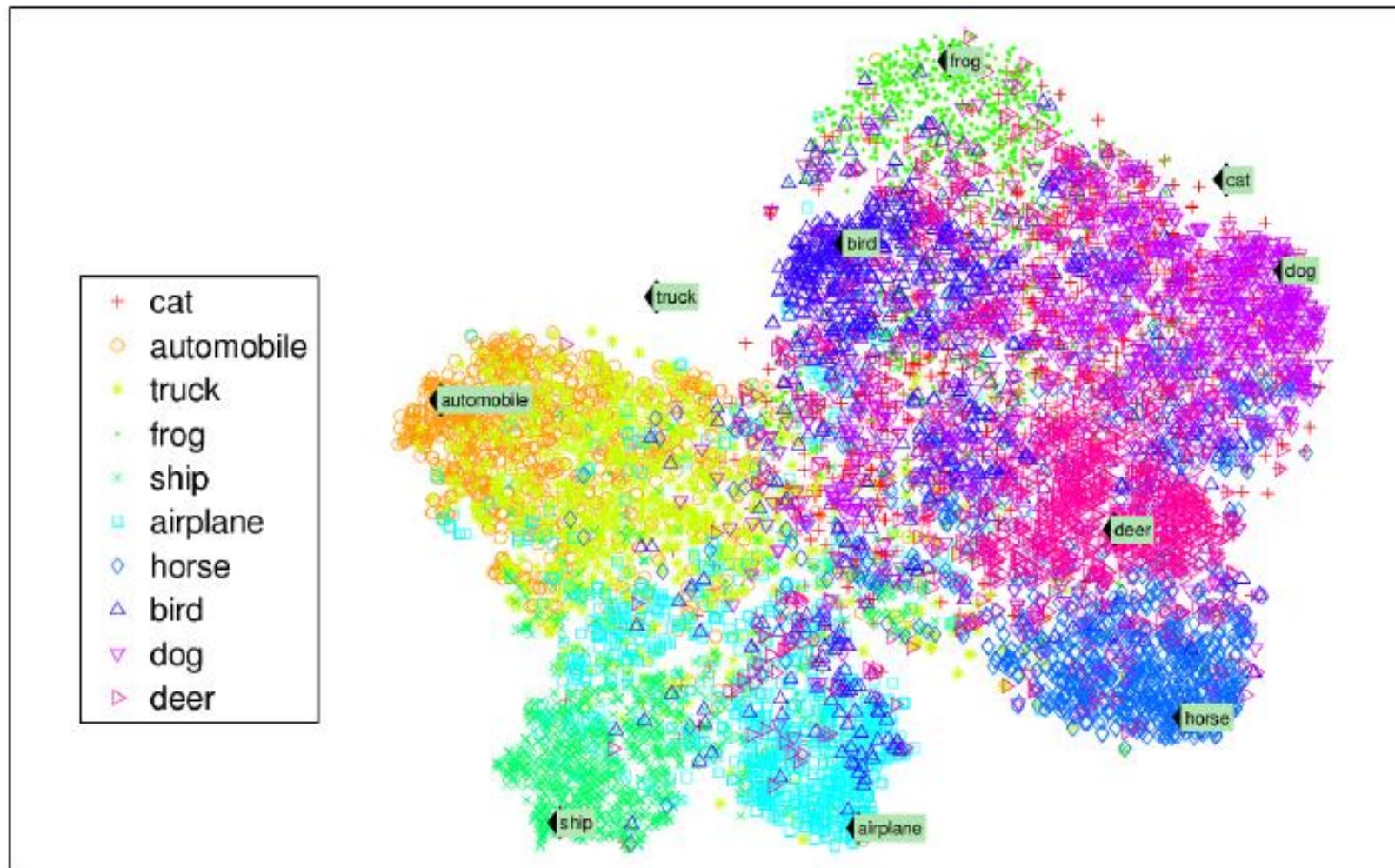


Figure 1-9. Example of a t-SNE visualization highlighting semantic clusters³

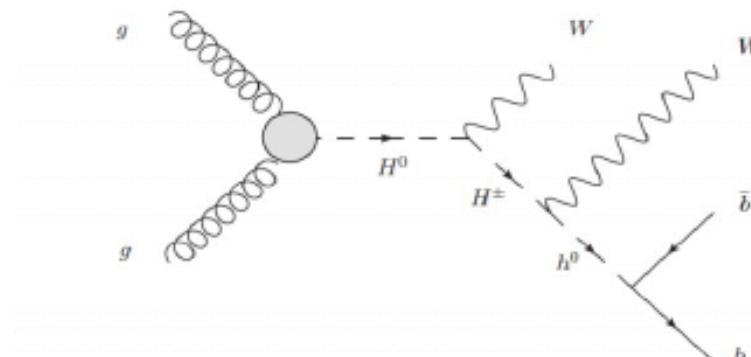
Generation

Searching for Exotic Particles in High-Energy Physics with Deep Learning

P. Baldi, P. Sadowski, and D. Whiteson

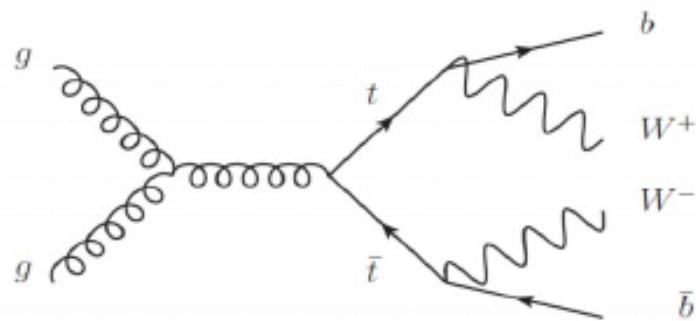
Nature Commun. 5, 4308 (2014)

Higgs benchmark



(a)

Signal



(b)

Background

8% improvement over the best approaches till publication