THE AUTOMATED PHYSICIST

EXPERIMENTAL PARTICLE PHYSICS IN THE ERA OF AI

Harrison B. Prosper

Florida State University

Colloquium

Michigan State University, February 8, 2018

Outline

- A Very Brief History of ML/AI
- The State of the Art
- The Large Hadron Collider
- The Automated Physicist

A VERY BRIEF HISTORY OF ML/AI

Aristotle in the
The School of Athens
Raphael, 1509
Wikimedia Commons

384 B.C. to 322 B.C

Example:

A = She is a physicist B = She is smart

Major premise: If A is TRUE, then B is TRUE

Minor premise: **She is a physicist** is TRUE

Conclusion: Therefore, **She is smart** is TRUE

Note, however, according to Aristotle, we cannot conclude that if **She is smart** is TRUE, **She is a physicist** is TRUE!

$$\boldsymbol{AB} = A, \qquad A = 1 \rightarrow B = 1, \text{but } B = 1 \nrightarrow A = 1$$

Moveable type (Gutenberg Bible, 1456)

By NYC Wanderer (Kevin Eng) - originally posted to Flickr as Gutenberg Bible

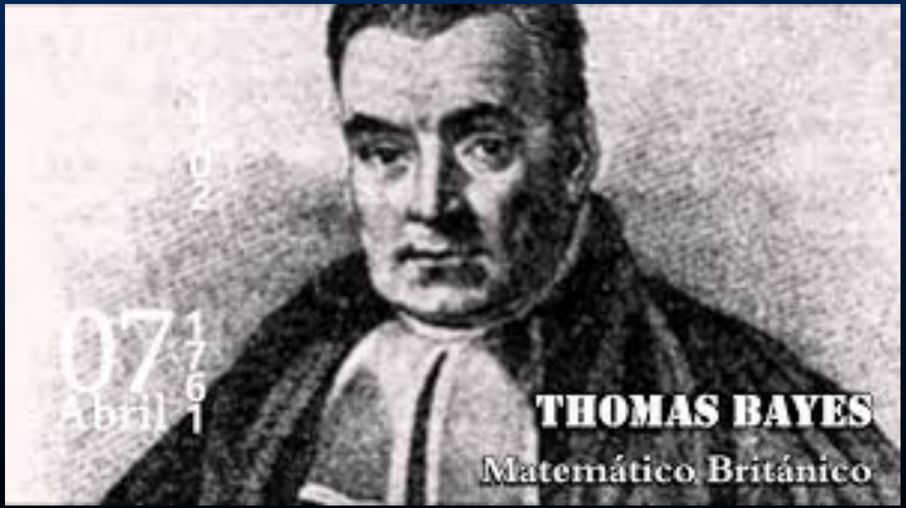
17th century

- Many philosophical ideas about knowledge, reason, and the nature of Man.

18th century

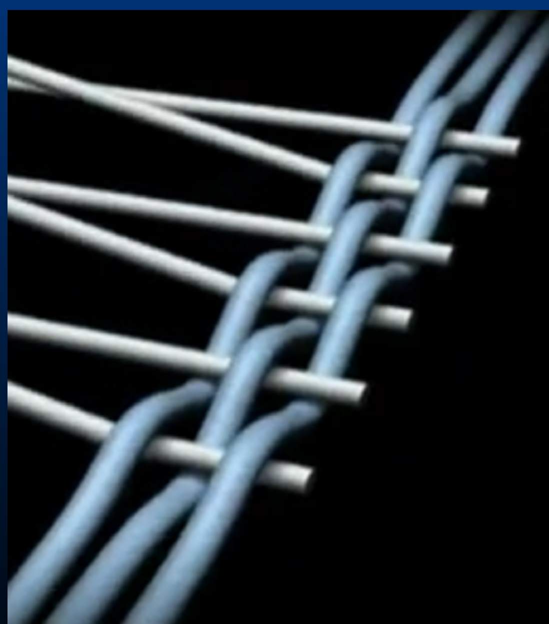
- 1763 – Thomas Bayes publishes important theorem.

$$P(\boldsymbol{H}|Data) = \frac{P(D|\boldsymbol{H})P(\boldsymbol{H})}{P(Data)}$$

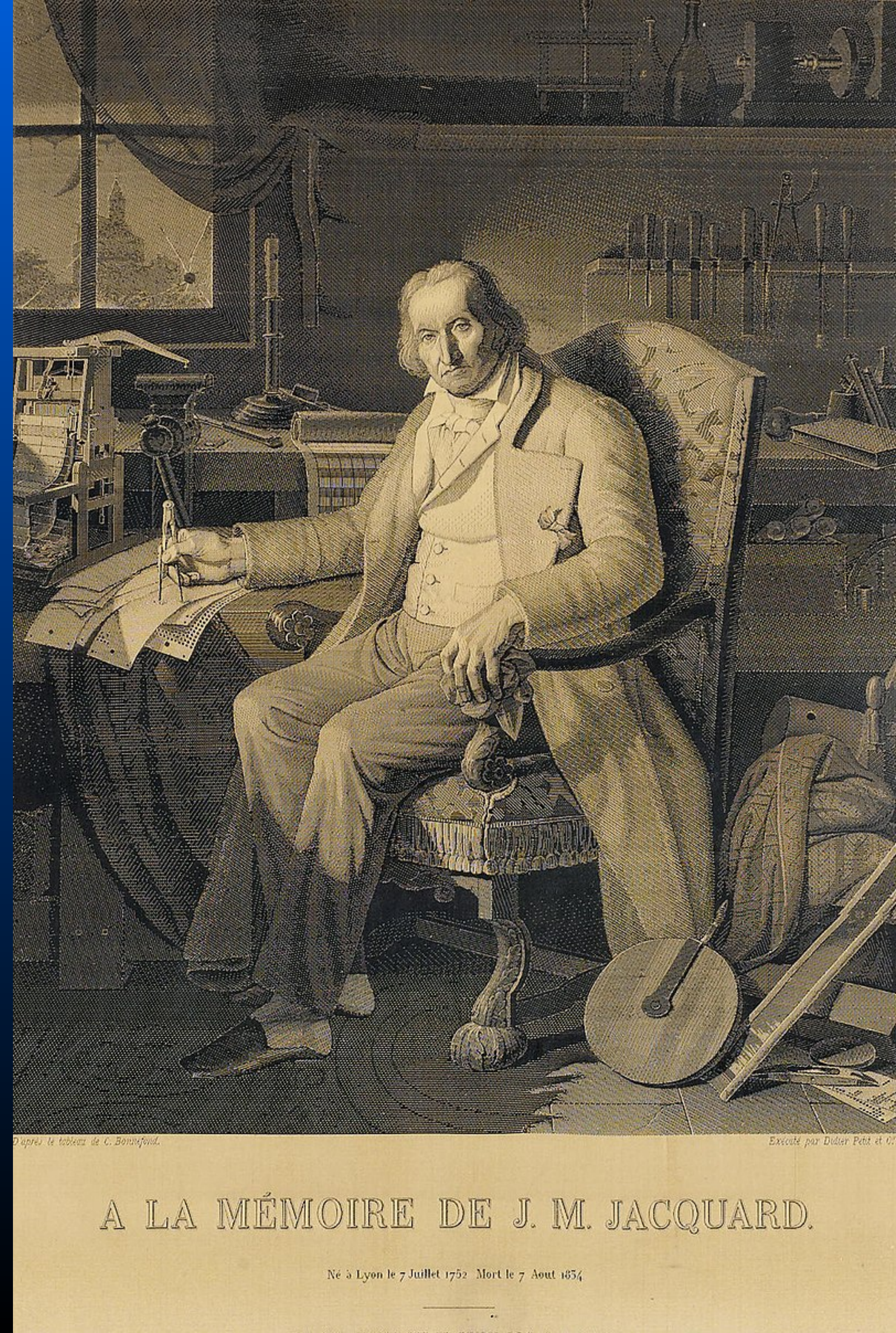


19th century

- 1801 – Joseph-Marie Jacquard invents first programmable machine.



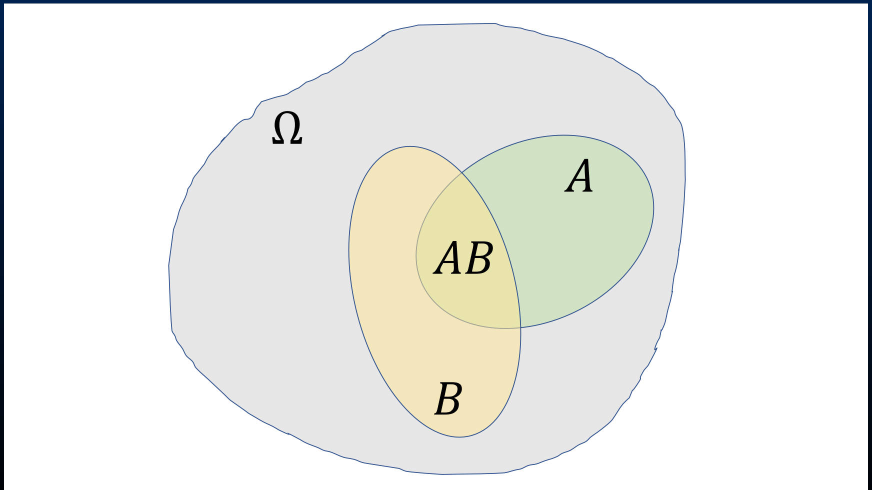
Wikimedia commons

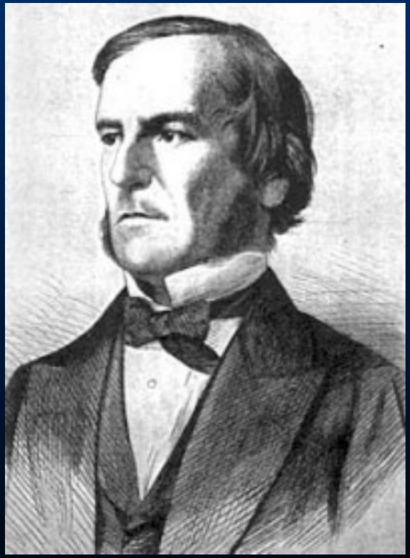


19th century

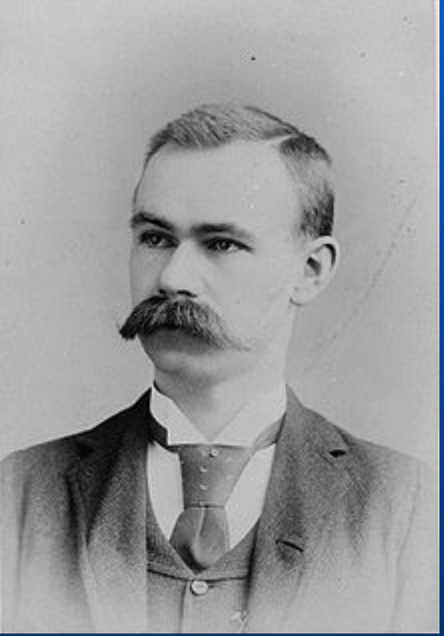
- 1832 – Charles Babbage designs first programmable calculator.

- 1854 – George Boole invents algebra of logic.





1815 -1864



1890 US Census

Herman Hollerith
(1860 – 1929)

Wikimedia commons

A Very Brief History

20th century (1900 – 1950)

- 1936 – Alan Turing proposes a universal computing machine.
- 1943 – Warren McCulloch and Walter Pitts invent *neural networks* (NN).
- 1950 – Turing Test, an operational definition of an artificially intelligent agent.

A Very Brief History

20th century (1950 – 2000)

- Many important developments:

1. First industrial robot (George Devol's Unimate).
2. Development of specialized computer languages.
3. First robot able visually to locate and assemble objects (Edinburgh University).
4. Werbos invents *backpropagation algorithm.*
5. First autonomous robot rover on Mars (Sojourner, NASA, July 1997).

1997 World chess champion Gary Kasparov defeated by IBM's Deep Blue

Stan Honda/AFP/Getty Images

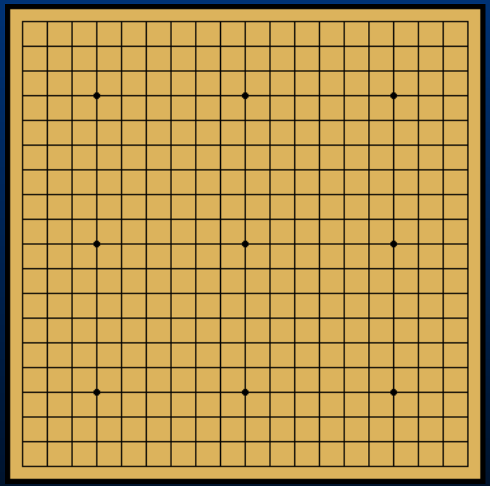
Computer Wins on ‘Jeopardy!’: Trivial, It’s Not
New York Times, Feb. 17, 2011

Carol Kaelson/Jeopardy Productions Inc., via Associated Press

Ken Jennings: “I felt obsolete”
TED Talk

Machine 4, Human 1

2016 – Google's DeepMind AlphaGo program beats Go champion Lee Sodol.



Photograph: Yonhap/Reuters

A Very Brief History of ML/AI

“Michigan State professors protest their replacement by iPhone 9000s”

New York Times, Feb. 8, 2078

MACHINE LEARNING:
THE STATE
OF
THE ART

“That is positively the dopiest idea I have heard.”

Richard Feynman,

Thinking Machines Corporation, summer 1983.

The State of the Art

REVIEW

doi:10.1038/nature14539

Deep learning

Yann LeCun[1,2], Yoshua Bengio[3] & Geoffrey Hinton[4,5]

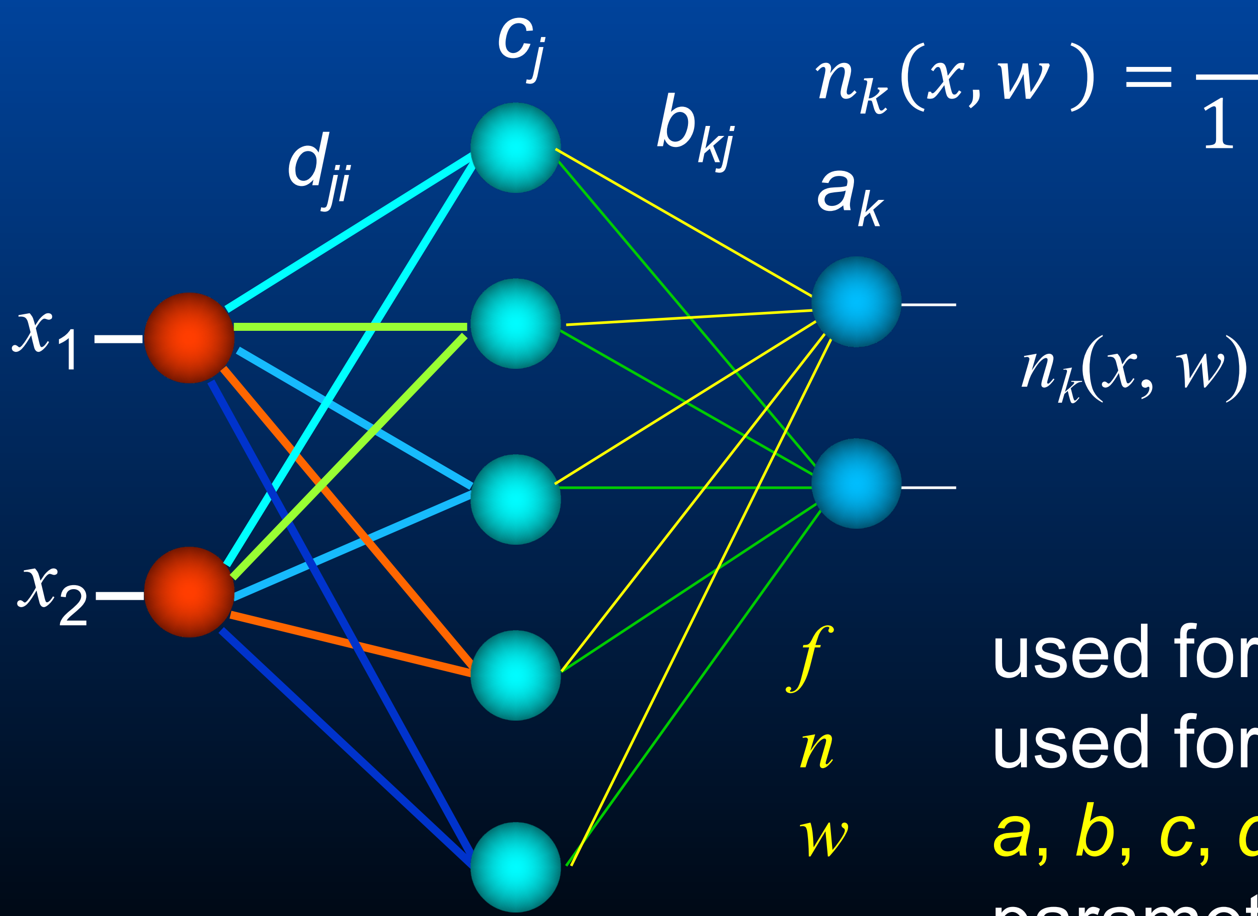
Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have dramatically improved the state-of-the-art in speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep learning discovers intricate structure in large data sets by using the backpropagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer. Deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent nets have shone light on sequential data such as text and speech.

436 | NATURE | VOL 521 | 28 MAY 2015

NN

$$f_k(x,w) = \boldsymbol{a_k} + \sum_{j=1}^{H} \boldsymbol{b_{kj}} \boldsymbol{h}\left(\boldsymbol{c_j} + \sum_{i=1}^{I} \boldsymbol{d_{ji}} x_i \right)$$

$$n_k(x,w) = \frac{1}{1+\exp[-f_k(x,w)]}$$



f used for regression

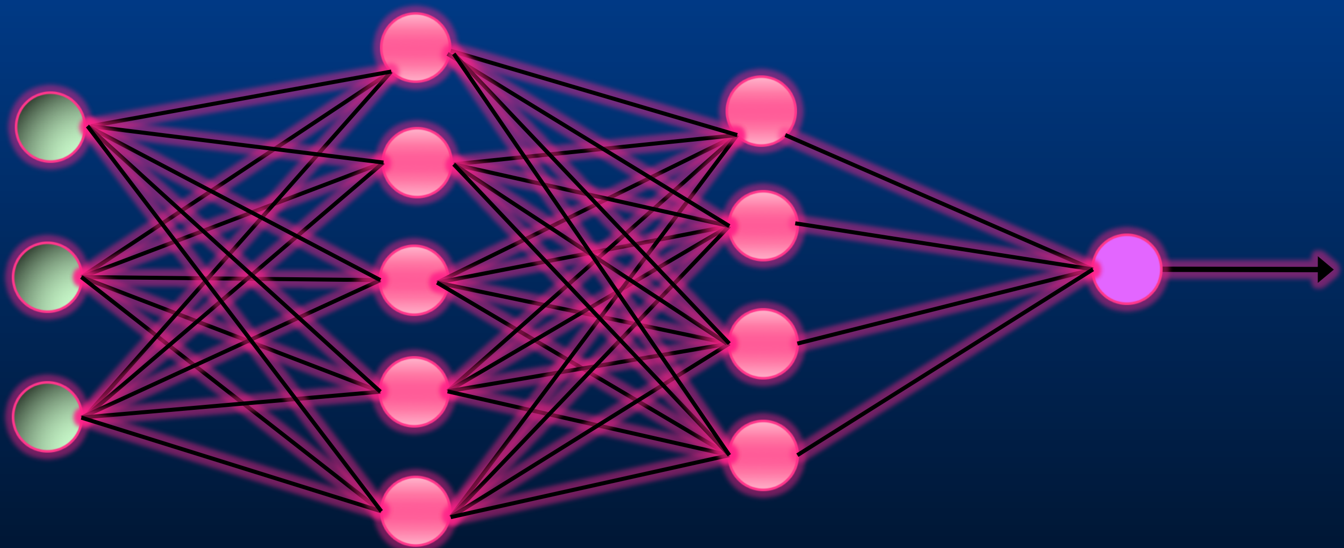
n used for classification

w *a*, *b*, *c*, *d* are free parameters to be fitted.

Deep Neural Networks

A deep neural network (DNN) with two “hidden” layers.

input layer **hidden layer 1** **hidden layer 2** **output layer**



$\boldsymbol{x}$

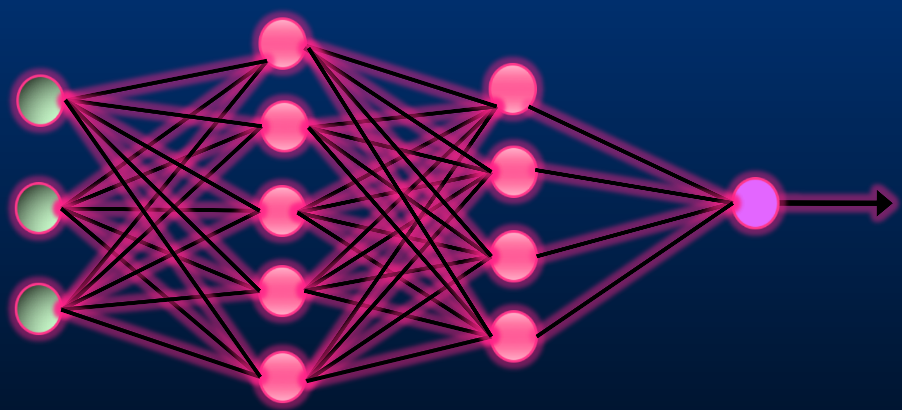
$\boldsymbol{h}_1(\mathbf{c}_1 + \mathbf{d}_1\boldsymbol{x})$

$\boldsymbol{h}_2(\mathbf{c}_2 + \mathbf{d}_2\boldsymbol{h}_1)$

$\mathbf{o} = \boldsymbol{h}_3(\mathbf{c}_3 + \mathbf{d}_3\boldsymbol{h}_2)$

Deep Awakening

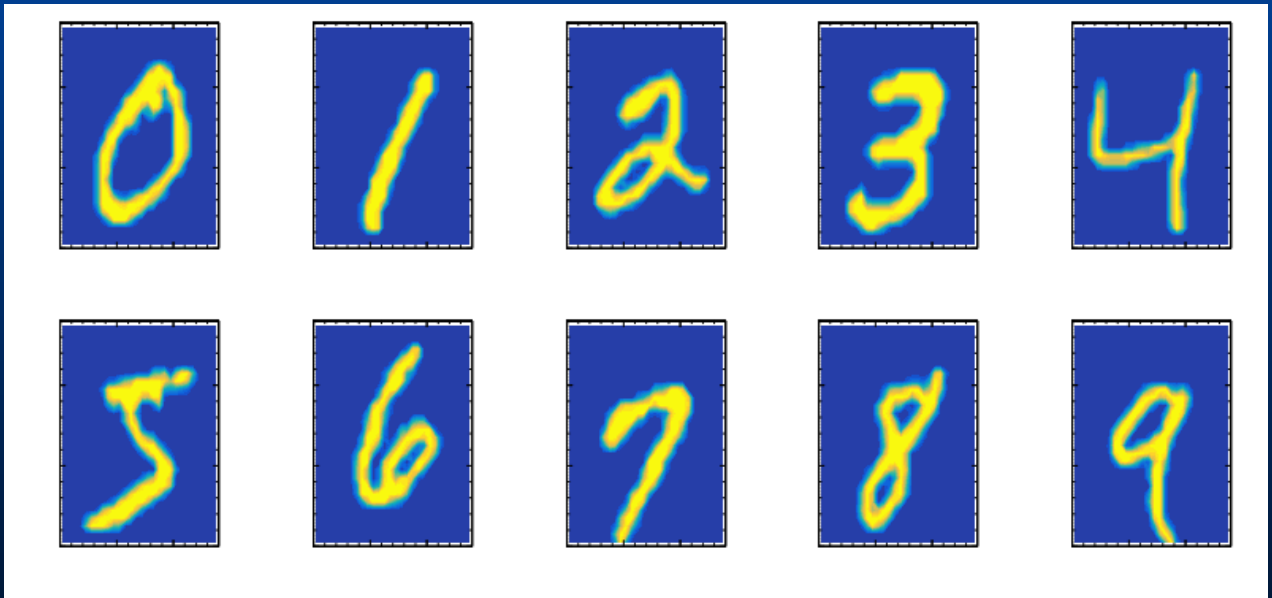
In 2006, University of Toronto researchers Hinton, Osindero, and Teh* developed a sophisticated practical method to train deep neural networks.



Geoffrey Hinton

* Hinton, G. E., Osindero, S. and Teh, Y., A fast learning algorithm for deep belief nets, Neural Computation 18, 1527-1554.

But, it turns out that sophistication may be overrated*!

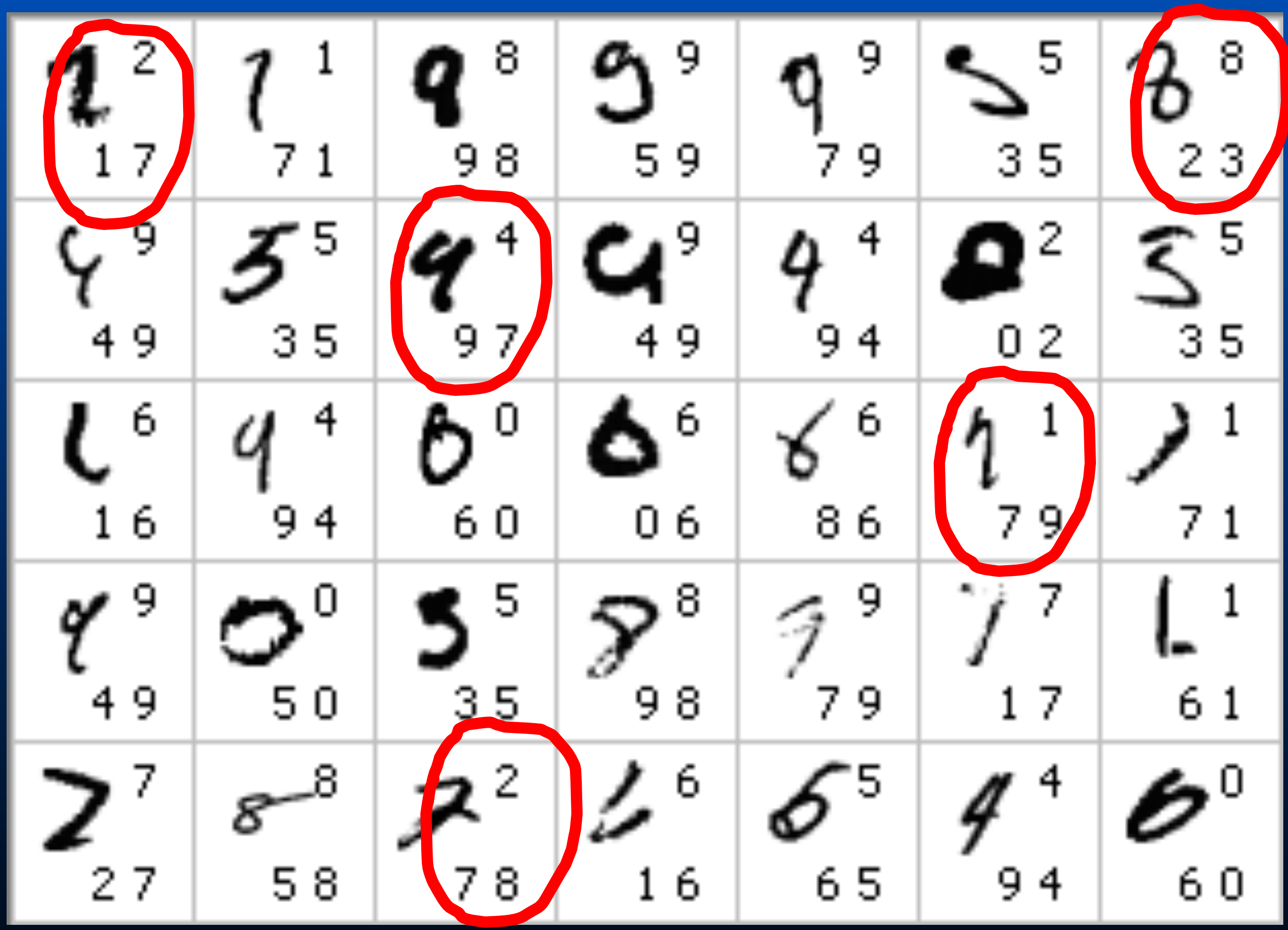


*Cireşan DC, Meier U, Gambardella LM, Schmidhuber J. ,

Deep, big, simple neural nets for handwritten digit recognition.

Neural Comput. 2010 Dec. 22 (12): 3207-20.

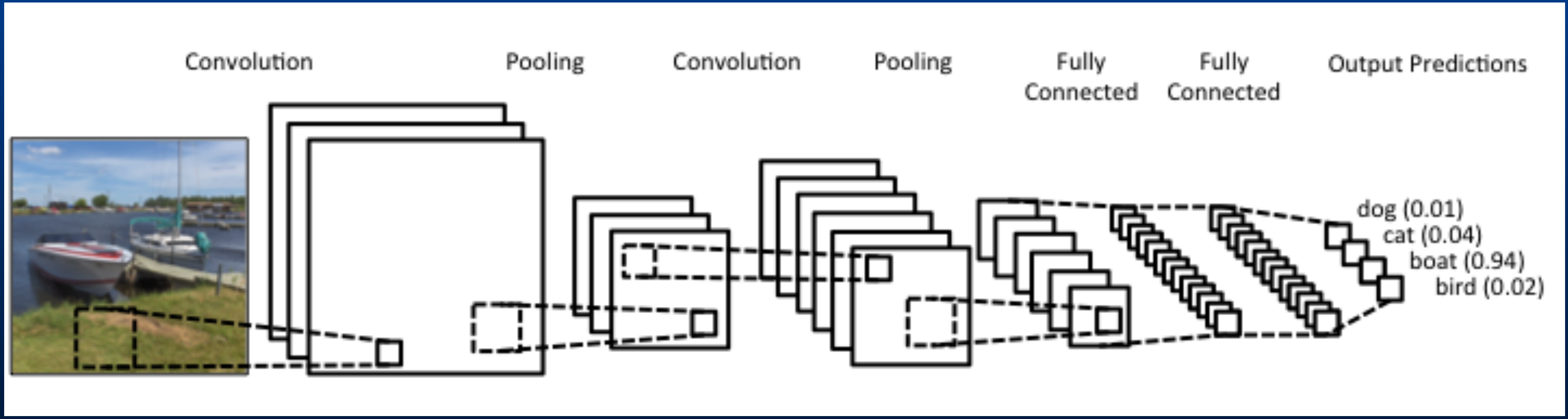
(784, 2500, 2000, 1500, 1000, 500, 10)



Upper right: correct answer; lower left answer of highest DNN output; lower right answer of next highest DNN output.

Deep Neural Networks

Many of the breakthroughs in tasks such as face recognition use a DNN called a *convolutional neural network* (CNN).

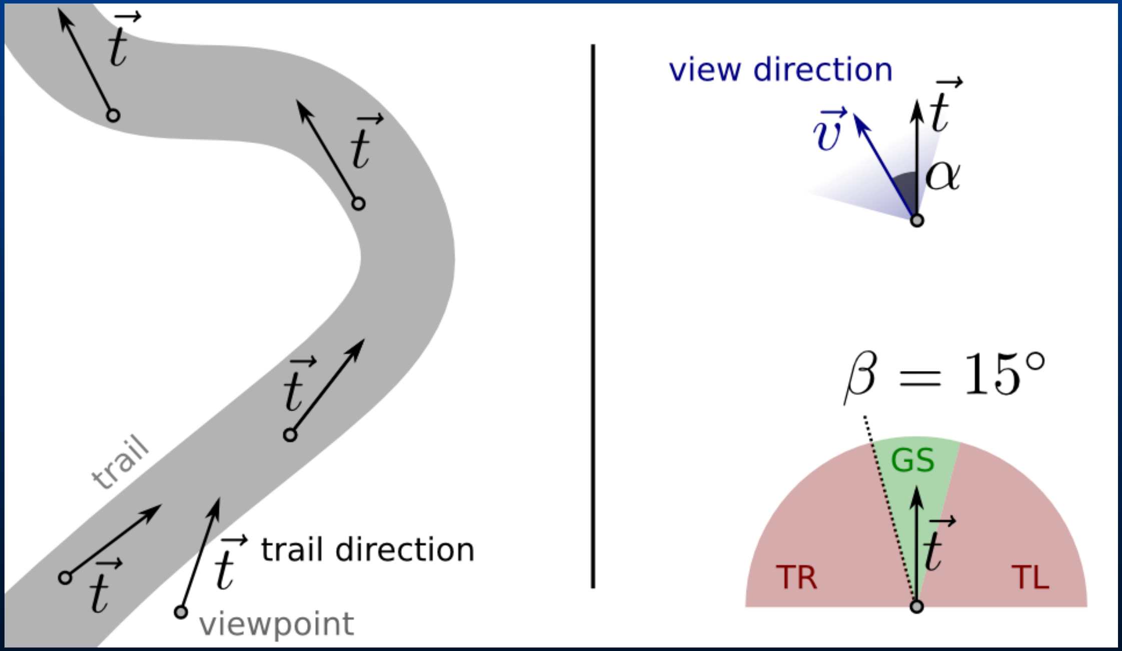


Source: https://www.clarifai.com/technology
http://yann.lecun.com/

Follow the Yellow Brick Road!

Giusti *et al.* treat the problem of trail navigation as a classification problem!

Data: 8 hours of 1920 x 1080 30fps video using 3 GoPro cameras.



IEEE Robotics and Automation Letters (Volume: 1, Issue: 2, July 2016)

ARTICLE

doi:10.1038/nature24270

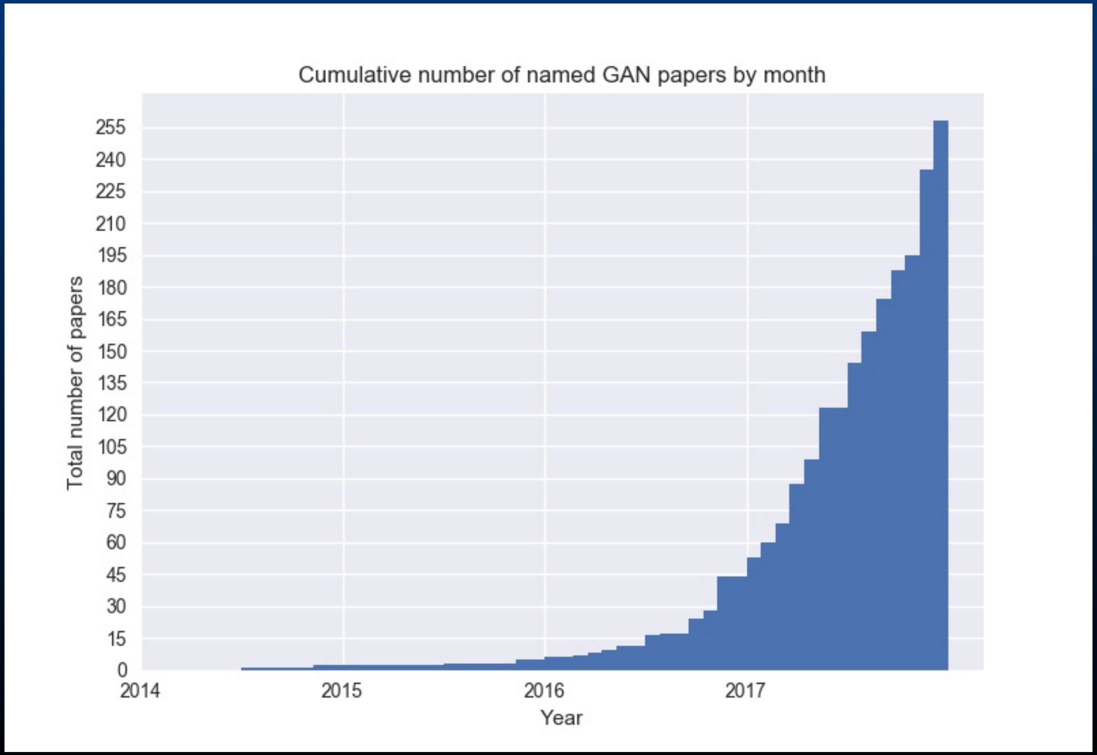
Mastering the game of Go without human knowledge

David Silver[1]*, Julian Schrittwieser[1]*, Karen Simonyan[1]*, Ioannis Antonoglou[1], Aja Huang[1], Arthur Guez[1], Thomas Hubert[1], Lucas Baker[1], Matthew Lai[1], Adrian Bolton[1], Yutian Chen[1], Timothy Lillicrap[1], Fan Hui[1], Laurent Sifre[1], George van den Driessche[1], Thore Graepel[1] & Demis Hassabis[1]

A long-standing goal of artificial intelligence is an algorithm that learns, *tabula rasa*, superhuman proficiency in challenging domains. Recently, AlphaGo became the first program to defeat a world champion in the game of Go. The tree search in AlphaGo evaluated positions and selected moves using deep neural networks. These neural networks were trained by supervised learning from human expert moves, and by reinforcement learning from self-play. Here we introduce an algorithm based solely on reinforcement learning, without human data, guidance or domain knowledge beyond game rules. AlphaGo becomes its own teacher: a neural network is trained to predict AlphaGo's own move selections and also the winner of AlphaGo's games. This neural network improves the strength of the tree search, resulting in higher quality move selection and stronger self-play in the next iteration. Starting *tabula rasa*, our new program AlphaGo Zero achieved superhuman performance, winning 100–0 against the previously published, champion-defeating AlphaGo.

Generative **A**dversarial **N**etworks

"Generative Adversarial Networks is the most interesting idea in the last ten years in machine learning"
Yann LeCun,
Chief AI Scientist, Facebook



Ian Goodfellow
OpenAI

MACHINE LEARNING:

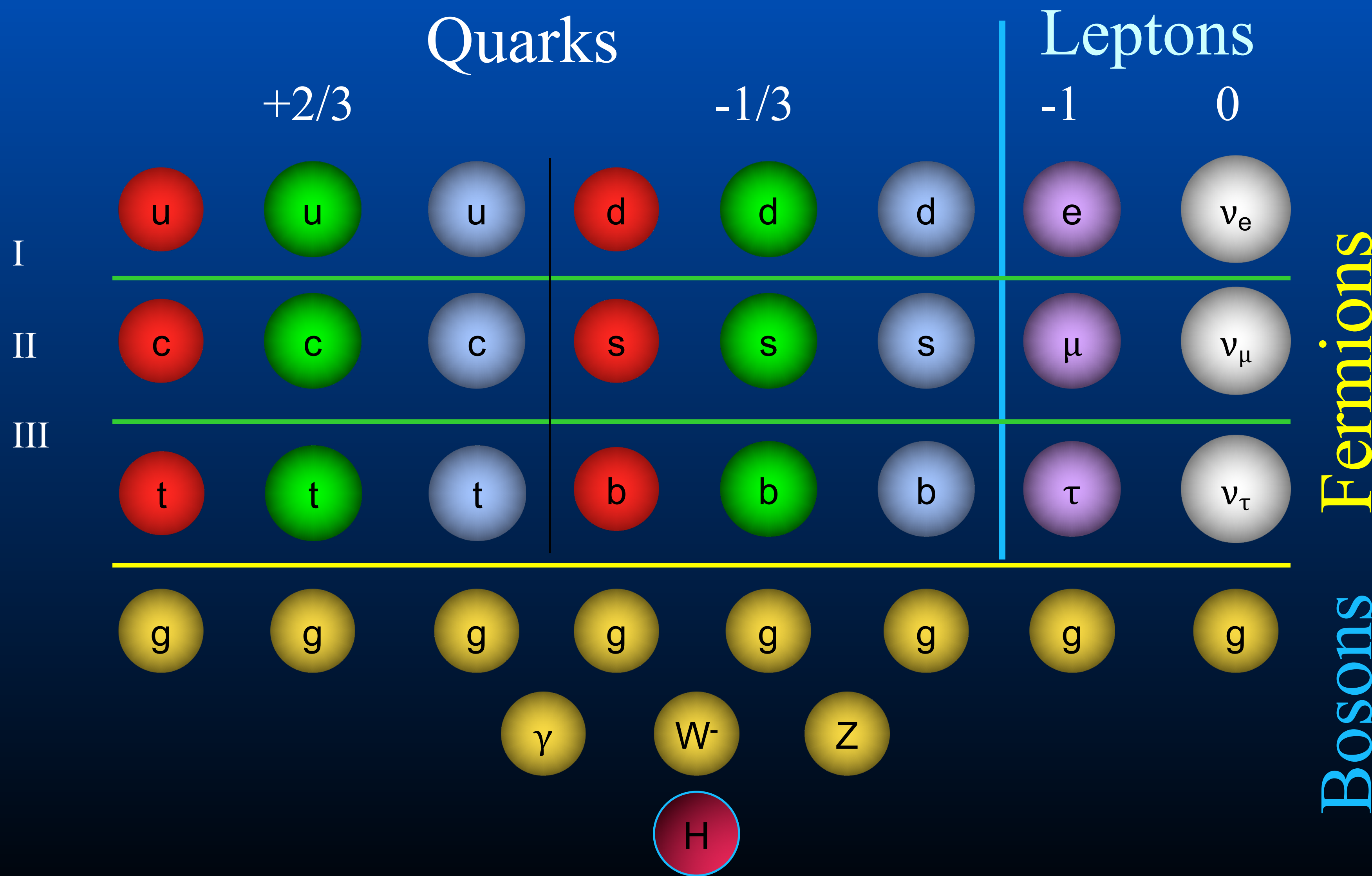
THE LARGE HADRON COLLIDER

"There are, therefore, agents in nature able to make the particles of bodies stick together by very strong attractions. And it is the business of experimental philosophy to find them out"

Sir Isaac Newton


Collision energy
13 TeV
Total stored energy
720 MJ
Collision rate
1GHz
Length
26.7 km
One Ring to rule them all,
One Ring to find them,
One Ring to bring them all
And in the darkness bind them.
The Large Hadron Collider
Jörg Wenninger

The Standard Model – 2018



Parameters of the Standard Model

Symbol	Description	Renormalization	Value
m_e	Electron mass		511 keV
m_μ	Muon mass		105.7 MeV
m_τ	Tau mass		1.78 GeV
m_u	Up quark mass	$\mu_{MS} = 2$ GeV	1.9 MeV
m_d	Down quark mass	$\mu_{MS} = 2$ GeV	4.4 MeV
m_s	Strange quark mass	$\mu_{MS} = 2$ GeV	87 MeV
m_c	Charm quark mass	$\mu_{MS} = m_c$	1.32 GeV
m_b	Bottom quark mass	$\mu_{MS} = m_b$	4.24 GeV
m_t	**Top quark mass**	On-shell scheme	**172.7 GeV**
θ_{12}	CKM 12-mixing angle		13.1°
θ_{23}	CKM 23-mixing angle		2.4°
θ_{13}	CKM 13-mixing angle		0.2°
δ	CKM CP-violating Phase		0.995
g_1 or g'	U(1) gauge coupling	$\mu_{MS} = m_Z$	0.357
g_2 or g	SU(2) gauge coupling	$\mu_{MS} = m_Z$	0.652
g_3 or g_s	SU(3) gauge coupling	$\mu_{MS} = m_Z$	1.221
θ_{QCD}	QCD vacuum angle		~0
v	Higgs VEV		246 GeV
m_H	**Higgs mass**		**125 GeV**

21st Century Physics

- Are the 19 parameters of the Standard Model random numbers, or can they be explained?
- What makes a top quark a top quark, an electron and electron, and a neutrino a neutrino?

  Chris Quigg,
- What is dark matter?
- What is dark energy?

THE AUTOMATED PHYSICIST

Machine Learning in Particle Physics

Evolution of the use of machine learning in physics:

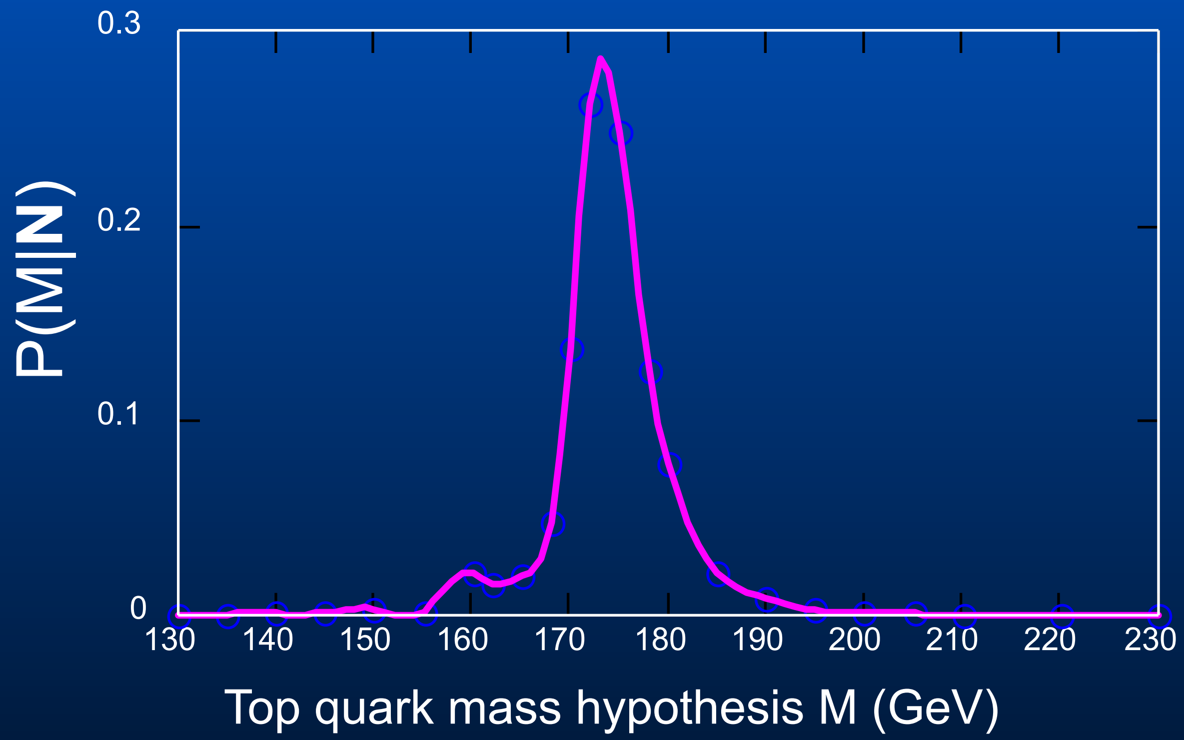
- **traditional:** classification & regression
- **emerging:** inference & generation

Kyle Cranmer, ACAT 2017

Machine Learning in HEP, The Early Days

- 1988 Denby, Comp. Phys. Comm.49:429 (1988)
- 1990 Bhat, Lönnblad, Meier, Sugano, *Snowmass;* Lönnblad, Peterson, Rögnvaldsson, *Phys. Rev. Lett*. 65:1321 (1990)
- 1992 Peterson, CHEP 92, Denby, FERMILAB-CONF-92-269-E (1992)
- 1994 Bhat PC (for the DØ Collaboration), APS Meeting, Albuquerque, NM
- 1997 Moneti (CLEO Collaboration) Nuclear Physics B (Proc. Suppl.) 59:17 (1997)

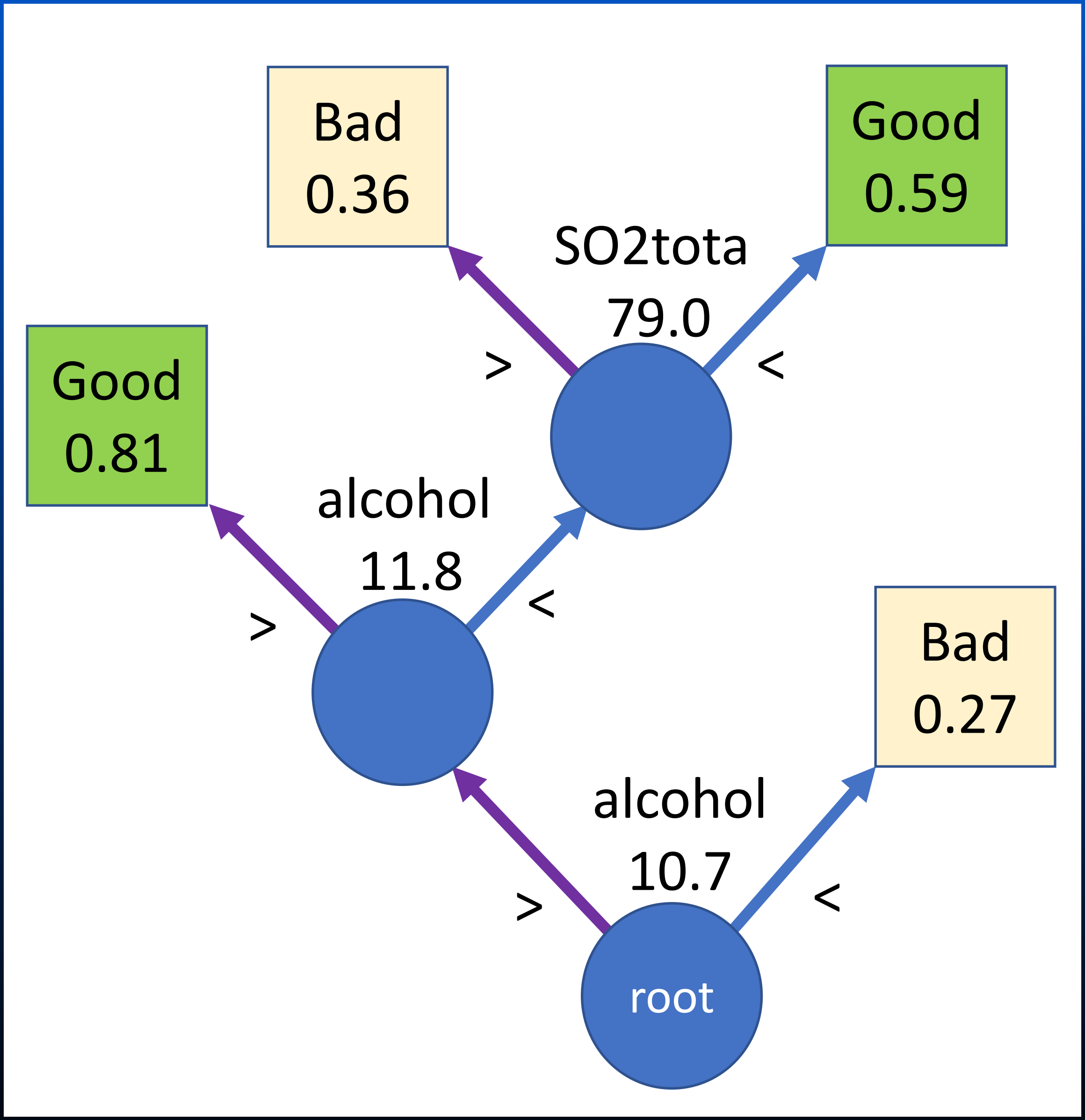
Top Quark Mass (DØ, 1997)



mass	= 173.5 ± 4.5 GeV (172.4 ± 0.5 GeV)
signal	= 33 ± 8 events
background	= 50.8 ± 8.3 events

Pushpalatha Bhat, HBP

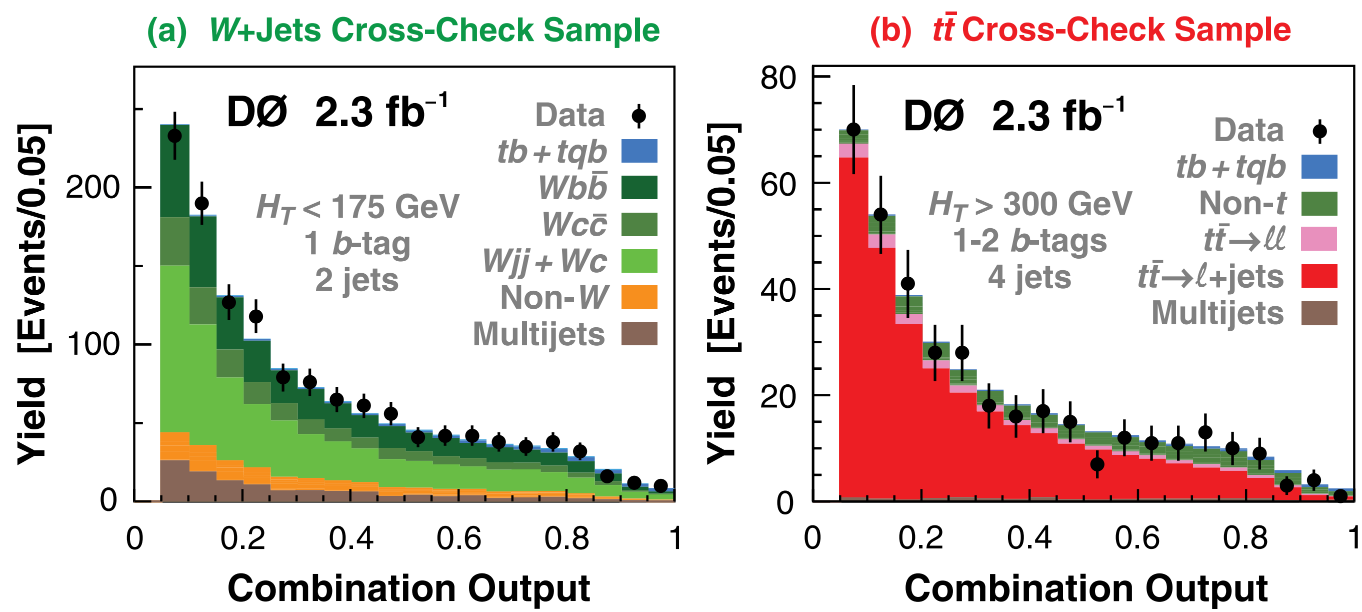

2007
2007
Corton-Charlemagne
GRAND CRU
J-F COCHE-DURY

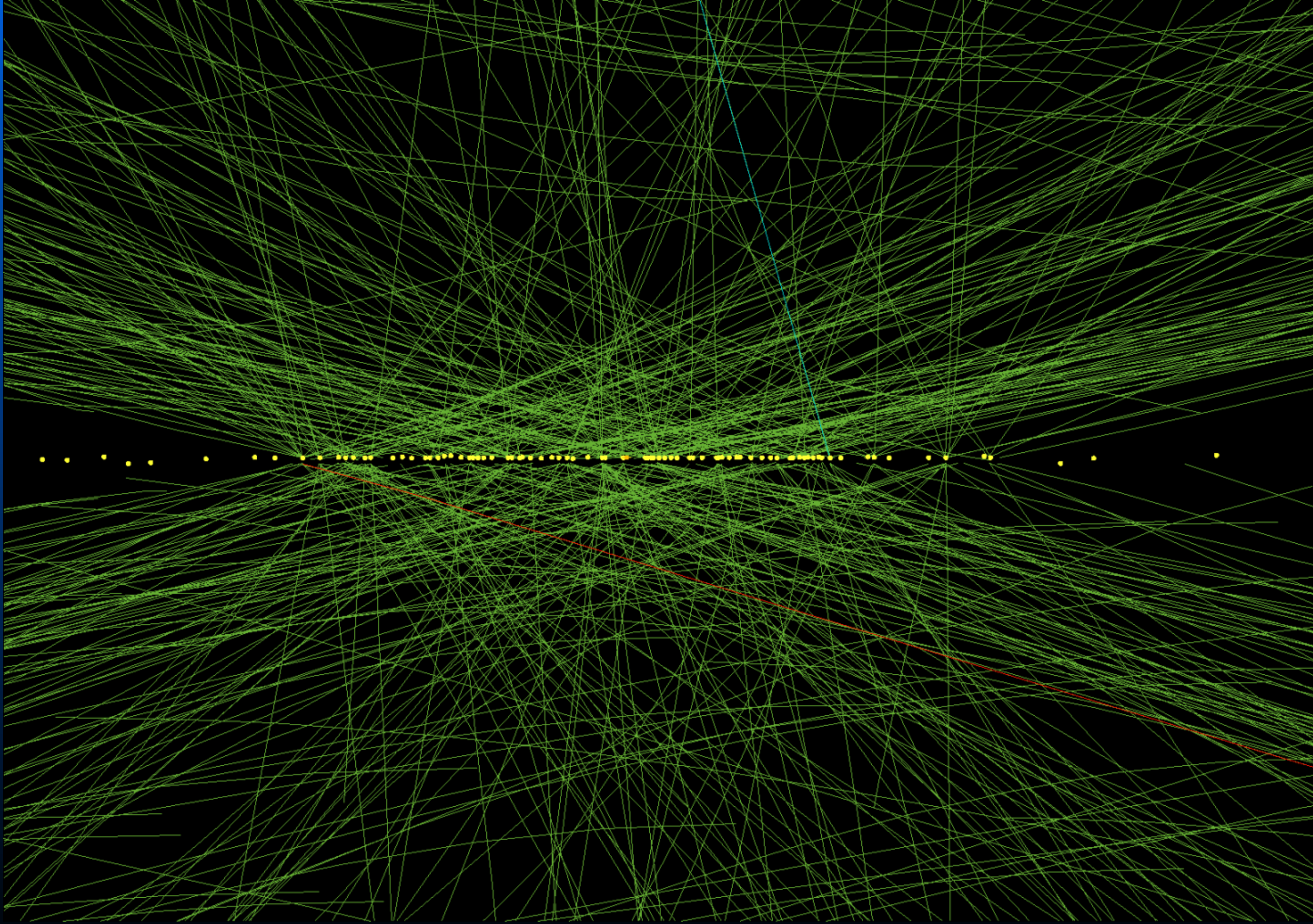

Bad
0.36
Good
0.59
SO2tota
79.0
>
<
Good
0.81
alcohol
11.8
>
<
Bad
0.27
alcohol
10.7
>
<
root

PRL 103, 092001 (2009) PHYSICAL REVIEW LETTERS week ending 28 AUGUST 2009

Observation of Single Top-Quark Production

V. M. Abazov,[36] B. Abbott,[74] M. Abolins,[64] B. S. Acharya,[29] M. Adams,[50] T. Adams,[48] E. Aguilo,[6] M. Ahsan,[58] G. D. Alexeev,[36] G. Alkhazov,[40] A. Alton,[64,*] G. Alverson,[62] G. A. Alves,[2] L. S. Ancu,[35] T. Andeen,[52] M. S. Anzelc,[52] M. Aoki,[49] Y. Arnoud,[14] M. Arov,[59] M. Arthaud,[18] A. Askew,[48,†] B. Åsman,[41] O. Atramentov,[48,†] C. Avila,[8]





CMS-PHO-EVENTS-2012-006

The Automated Physicist

- Automatically construct an algorithm to identify the particles from the main collision point.

- Automatically compress particle data (p_T, η, ϕ and identity) into a smaller set of numbers for further analysis.

- Automatically search for and characterize deviations between simulated and real data.

- Automatically construct summary reports.

McKinsey&Company

MCKINSEY GLOBAL INSTITUTE

A FUTURE THAT WORKS: AUTOMATION, EMPLOYMENT, AND PRODUCTIVITY

JANUARY 2017

EXECUTIVE SUMMARY

"Almost half the activities people are paid almost $16 trillion in wages to do in the global economy have the potential to be automated by adapting currently demonstrated technology, according to our analysis of more than 2,000 work activities across 800 occupations."

McKinsey & Company,

A FUTURE THAT WORKS: AUTOMATION, EMPLOYMENT, AND PRODUCTIVITY
Executive Summary January 2017


THANK YOU!