



$\langle aQa \rangle$

<https://aqa.liacs.nl/>

Ktokolwiek widział, ktokolwiek wie! (Ukradziony) Nobel 2024

Anna Dawid

Zapytaj Fizyka, 1 kwietnia 2025 r.

THE NOBEL PRIZE IN PHYSICS 2024



Illustrations: Niklas Elmehed

John J. Hopfield

Geoffrey E. Hinton

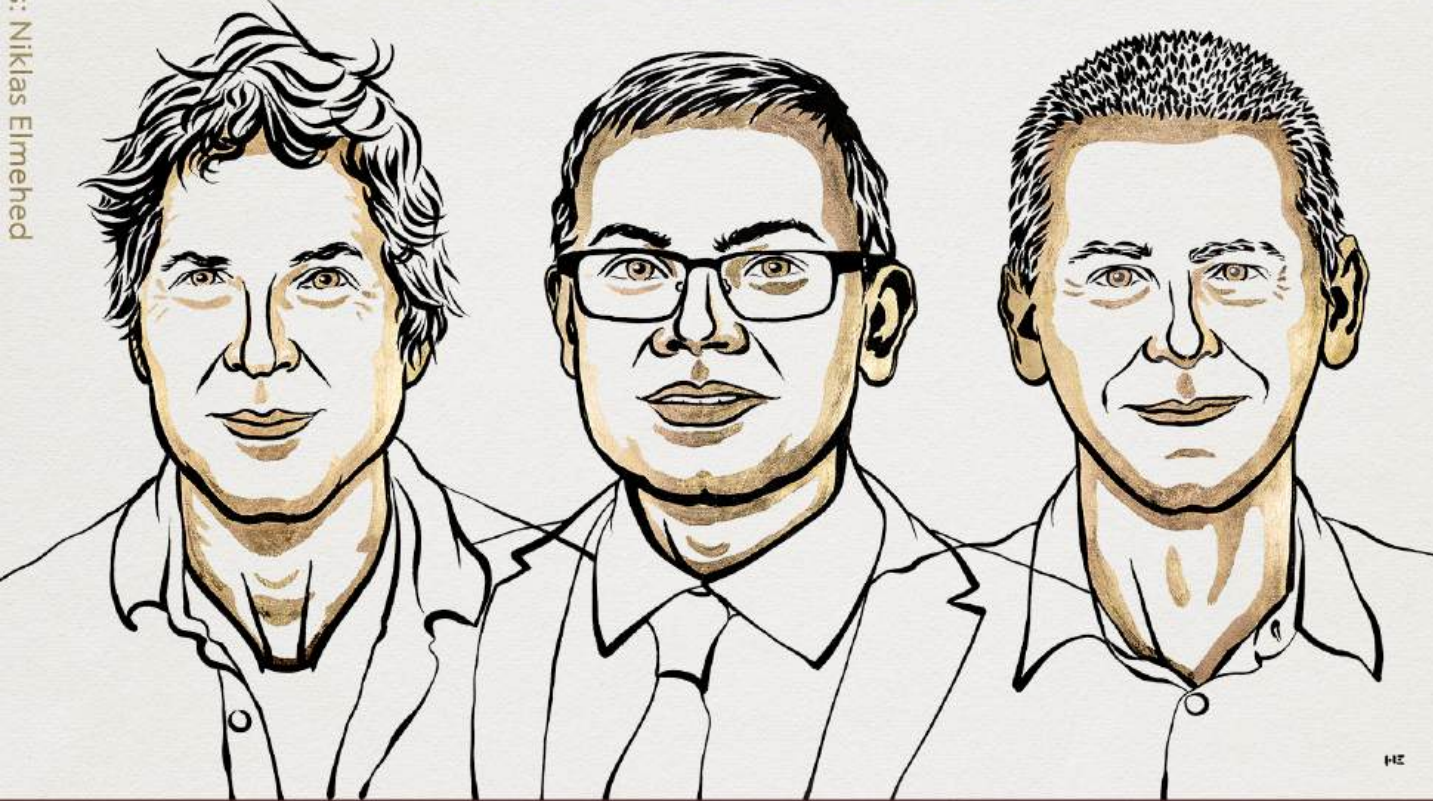
“for foundational discoveries and inventions
that enable machine learning
with artificial neural networks”

THE ROYAL SWEDISH ACADEMY OF SCIENCES

<aQa’>

“za fundamentalne odkrycia
i wynalazki, które umożliwiają
uczenie maszynowe
z wykorzystaniem sztucznych
sieci neuronowych”

THE NOBEL PRIZE IN CHEMISTRY 2024



David
Baker

“for computational
protein design”

Demis
Hassabis

“for protein structure prediction”

John M.
Jumper

THE ROYAL SWEDISH ACADEMY OF SCIENCES



“za projektowanie białek z
wykorzystaniem metod
obliczeniowych”

oraz

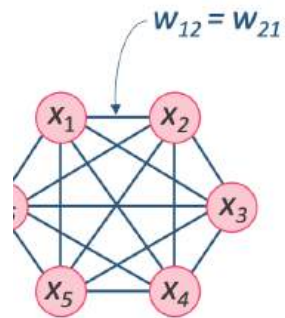
“za przewidywanie
struktury białek”

(AlphaFold)

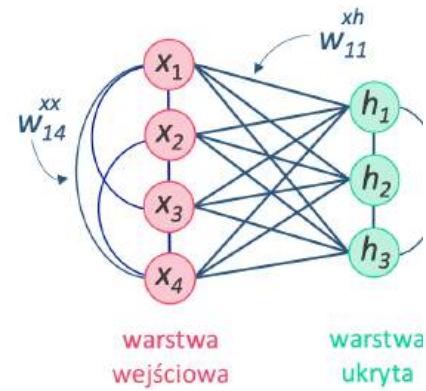
Plan wykładu



“Ukradziona” Nagroda Nobla?



Badania noblistów



Otwarte problemy i wyzwania

(Ukradziony) Nobel 2024

~~Czy to informatycy ukradli Nobla z fizyki?~~

~~Czy to fizycy ukradli sieci neuronowe informatyce?~~

!

A century of physics

Roberta Sinatra, Pierre Deville, Michael Szell, Dashun Wang and Albert-László Barabási

An analysis of Web of Science data spanning more than 100 years reveals the rapid growth and increasing multidisciplinary of physics — as well its internal map of subdisciplines.

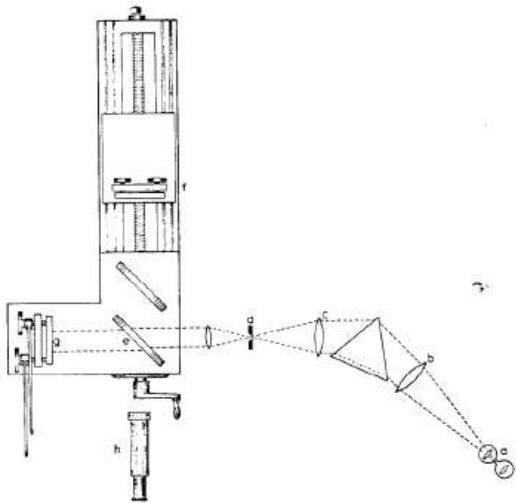
Physics is what physicists do

~Sam Edwards

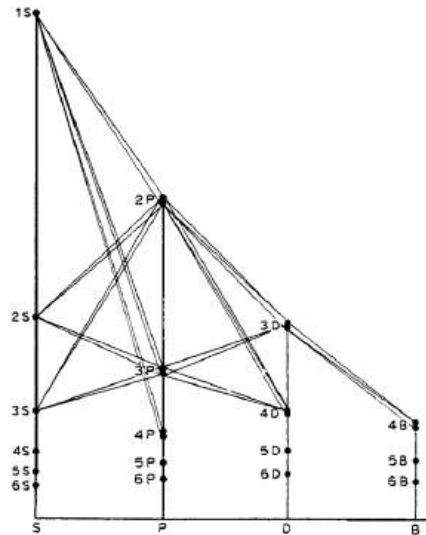
Uczenie maszynowe
to nowe narzędzie
do badania świata



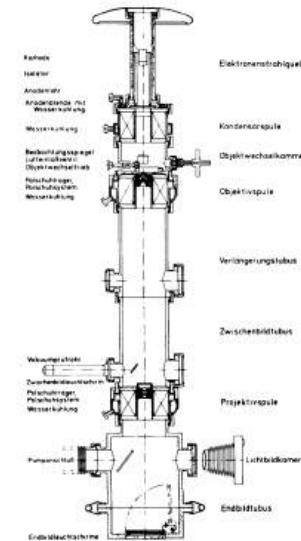
Wybrane Noble z fizyki za przyrządy pomiarowe



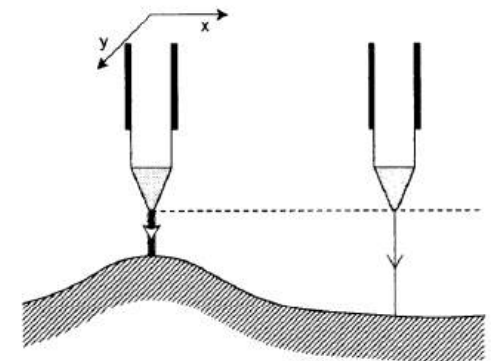
1907 - A. A. Michelson
interferometr



1924 - M. Siegbahn
spektroskopia rentgenowska

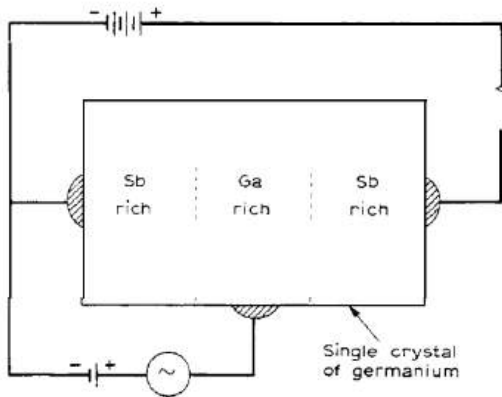


1986 - E. Ruska
mikroskop elektronowy

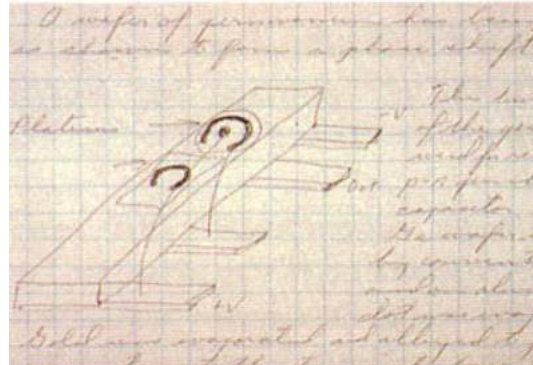


1986 - G. Binnig & H. Rohrer
skaningowy
mikroskop tunelowy

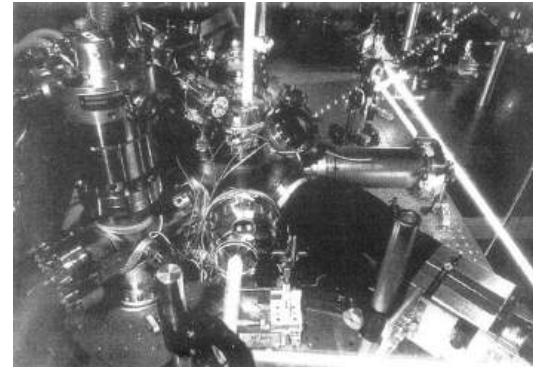
Wybrane Noble z fizyki za przyrządy pomiarowe



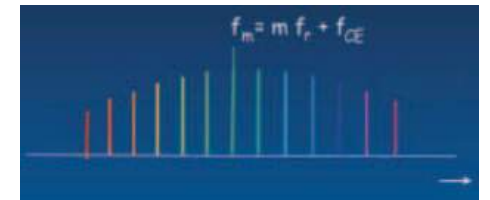
1956 - W. B. Shockley,
J. Bardeen, & W. H. Brattain
 tranzystor



2000 - J. S. Kilby (1/2)
 obwody zintegrowane



1997 - S. Chu, W. D. Philips,
& C. Cohen-Tannoudji
 metody chłodzenia atomów



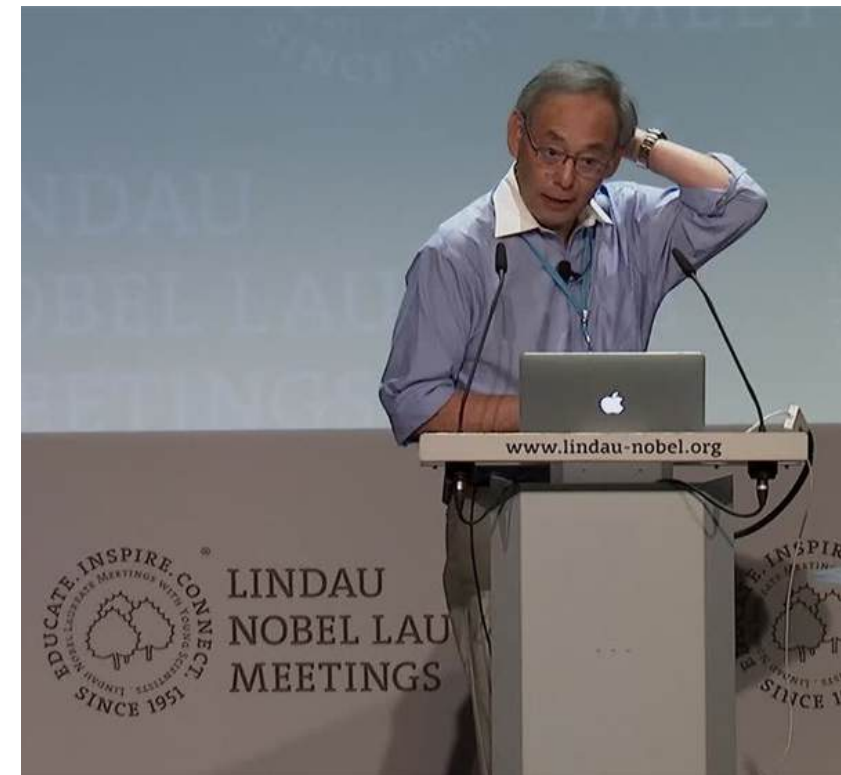
2005 - J. L. Hall & T. W. Hänsch
 grzebień częstotliwości

Wybrane Noble z fizyki za przyrządy pomiarowe

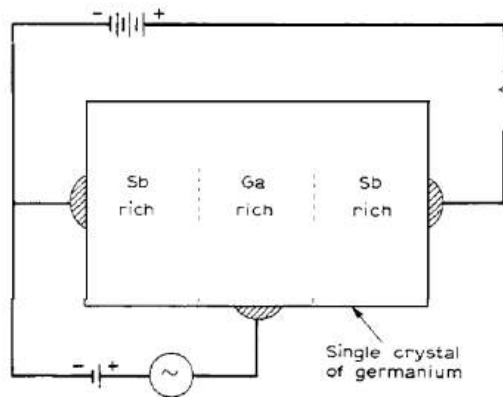
“Najpewniejszy sposób na Nobla to zrobienie nowego przyrządu pomiarowego.

To jak zagłębienie po raz pierwszy pod nowy kamień – z dużym prawdopodobieństwem znajdziecie tam coś ciekawego!”

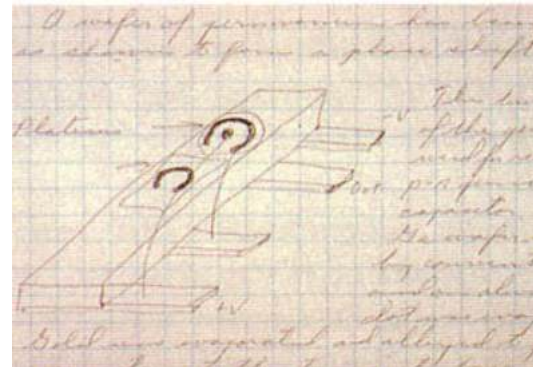
Steven Chu, 2024 Nobel Laureate Meeting



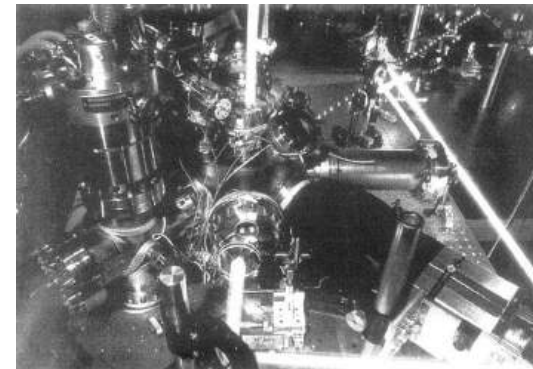
Wybrane Noble z fizyki za przyrządy pomiarowe



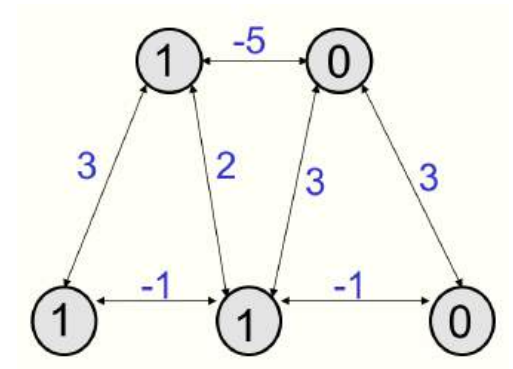
1956 - W. B. Shockley,
J. Bardeen, & W. H. Brattain
tranzystor



2000 - J. S. Kilby (1/2)
obwody zintegrowane



1997 - S. Chu, W. D. Philips,
& C. Cohen-Tannoudji
metody chłodzenia atomów



2024 - J. J. Hopfield & G. Hinton
sztuczne sieci neuronowe

Początki maszyn uczących

1. Podejście oparte na **kodowaniu wiedzy** w językach formalnych *szczególnie skuteczne w łamigłówkach logicznych i obliczeniach matematycznych*

- wydłużony pysk
- mniej wyraźne wąsy
- okrągłe oczy
- większa masa
- ...



- krótszy pyszczek
- długie wąsy
- migdałowe oczy
- mniejsza masa
- ...

Początki maszyn uczących

1. Podejście oparte na **kodowaniu wiedzy** w językach formalnych
2. Podejście oparte na wyciąganiu **wzorców z danych**

To jest PIES

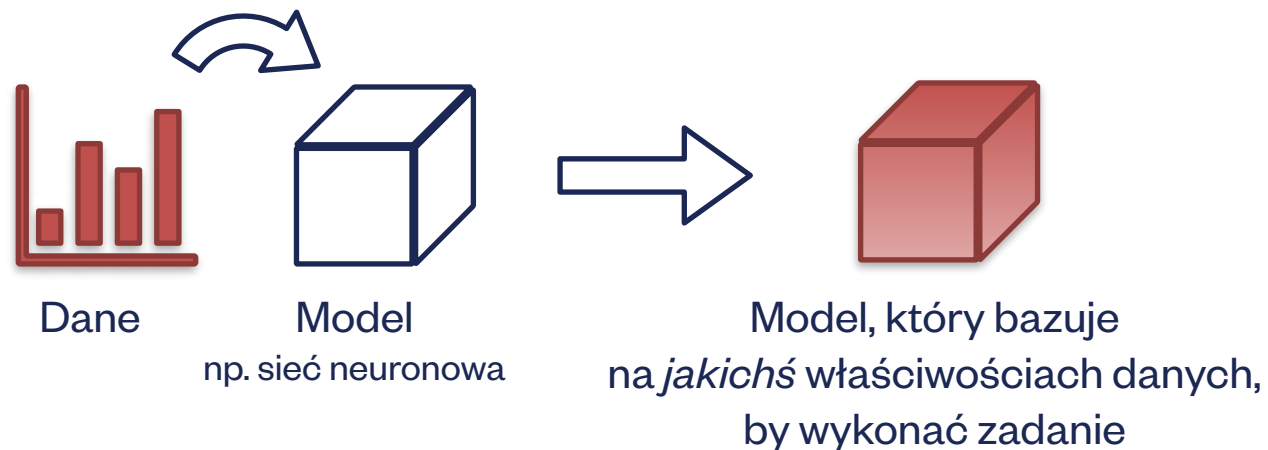


To jest KOT

Wymyśl jak je rozróżnić!

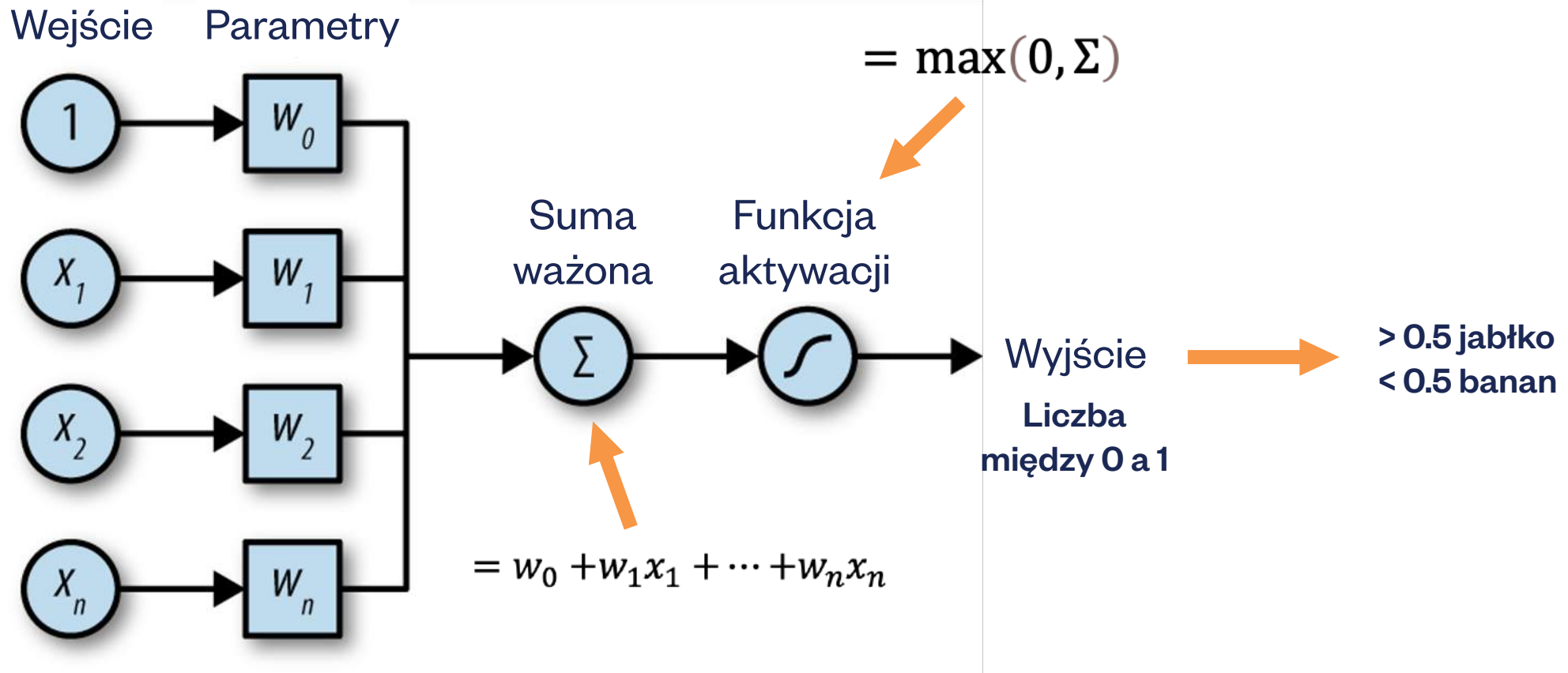
Czym jest uczenie maszynowe?

- Algorytmy, które rozwiązują problemy **bez zaprogramowania** jak *dokładnie* ten problem rozwiązać (niekoniecznie z użyciem sieci neuronowych)
- Ich skuteczność **rośnie wraz z dostępem do danych** („nauka przez doświadczenie”)
- Głębokie uczenie maszynowe to uczenie maszynowe z wykorzystaniem głębokich sieci neuronowych

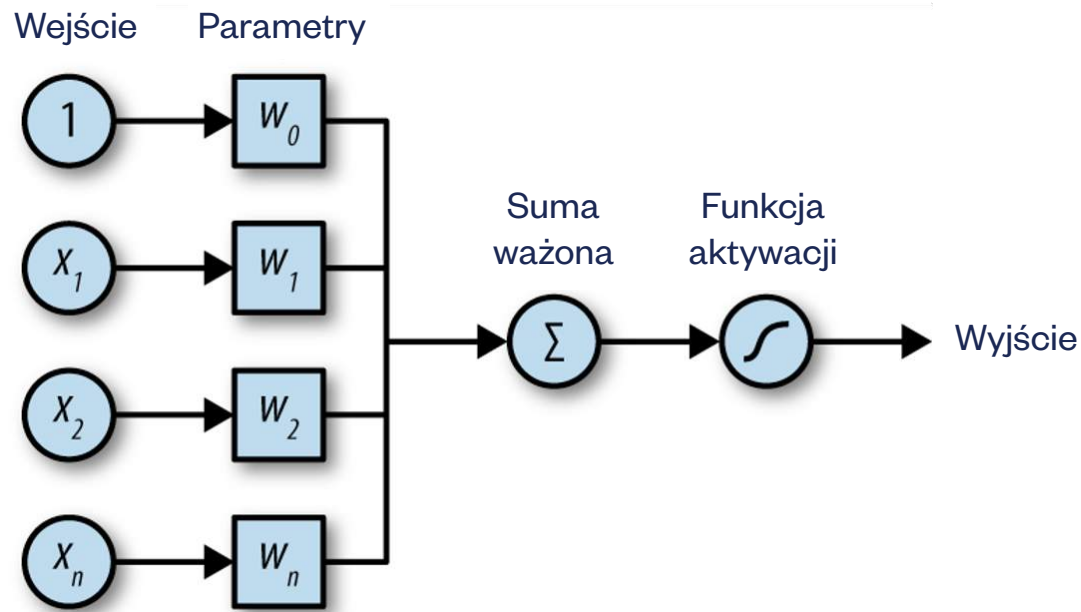


Proszę Państwa, oto neuron!

np. wszystkie piksele obrazka z jabłkiem



Proszę Państwa, oto neuron! Co dalej?



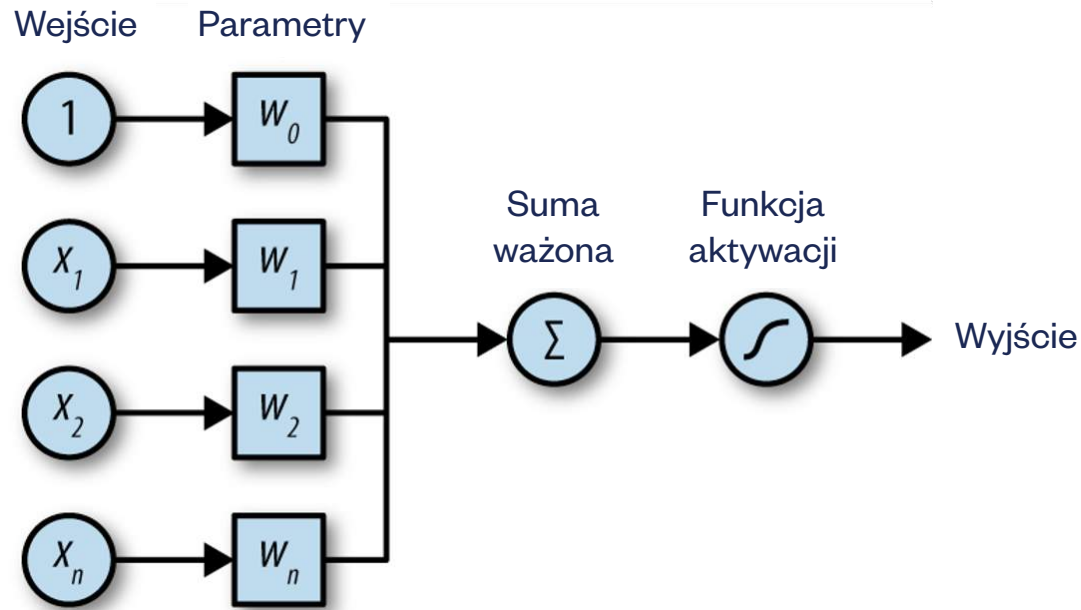
Sieć Hopfielda



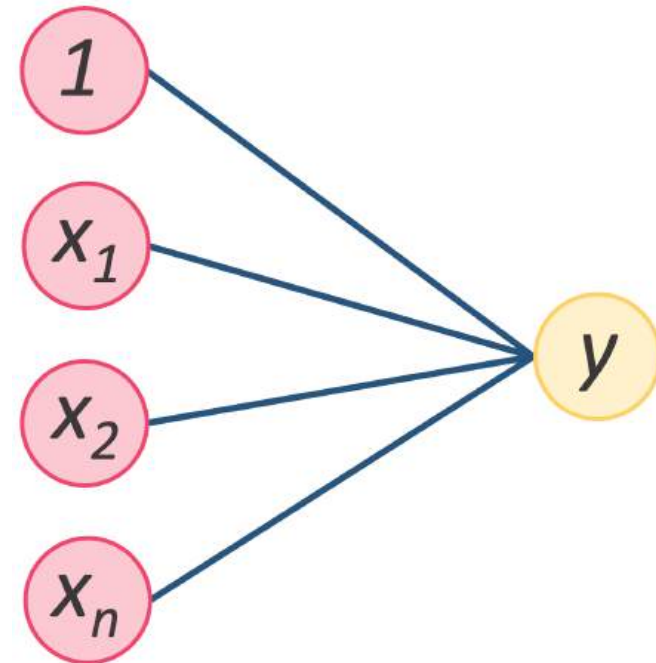
Maszyna Boltzmann

$\langle aQa' \rangle$

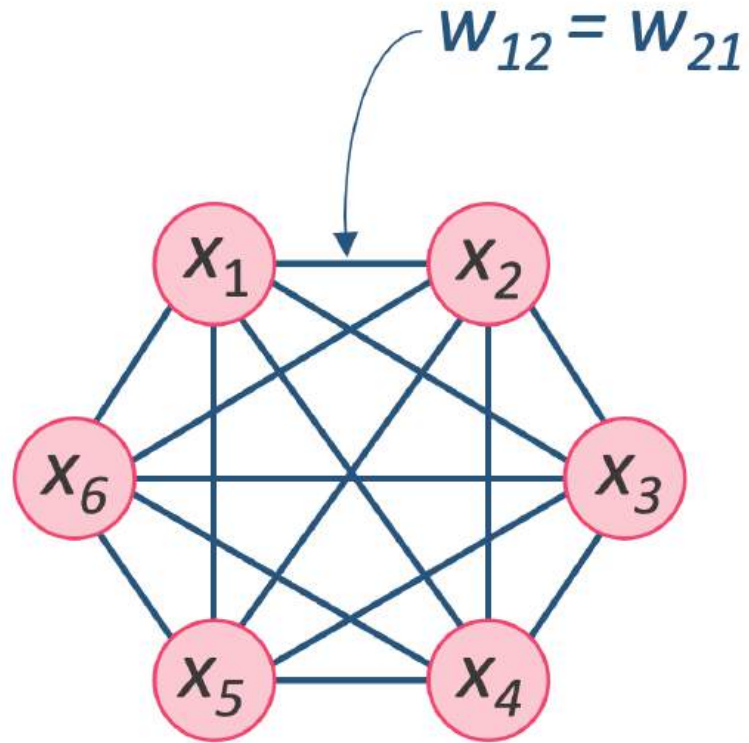
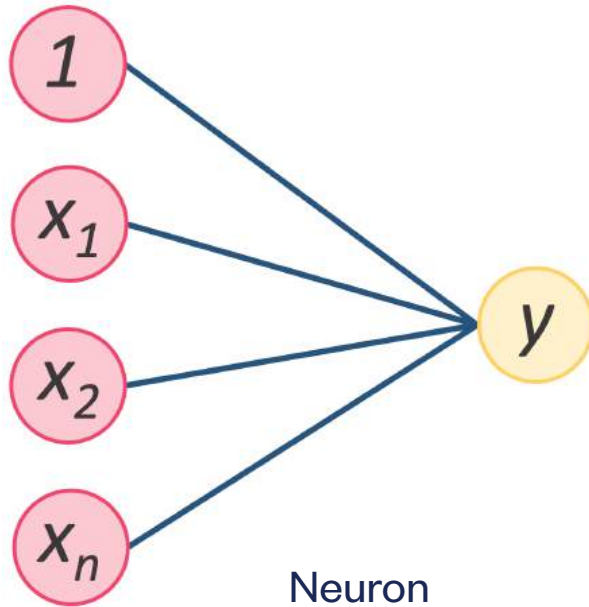
Proszę Państwa, oto neuron! Co dalej?



=



Sieć Hopfielda

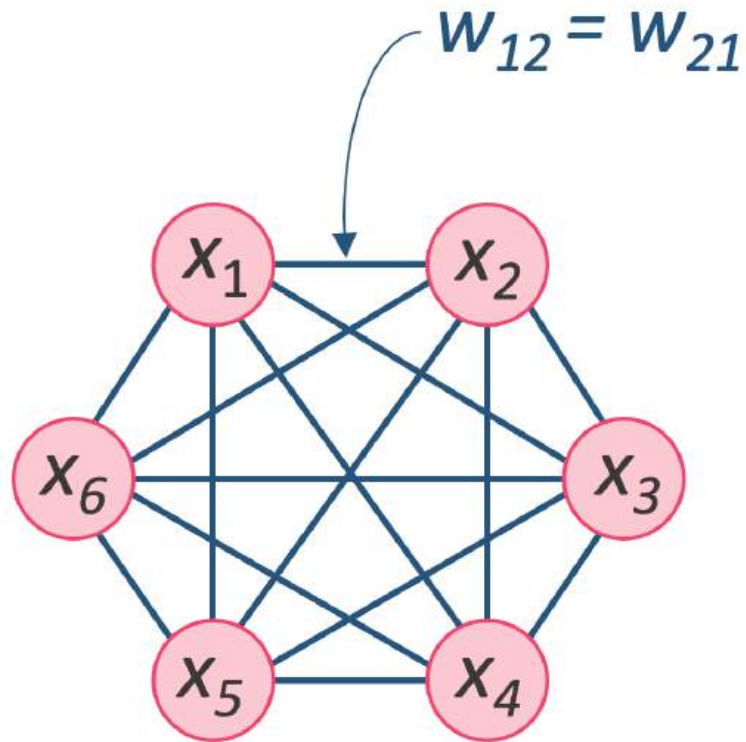


Sieć Hopfielda

$$F_w(\mathbf{x}) = - \sum_{ij} x_i w_{ij} x_j$$

Energia sieci zależy
od jej parametrów
i danych wejściowych

Sieć Hopfielda



Sieć Hopfielda

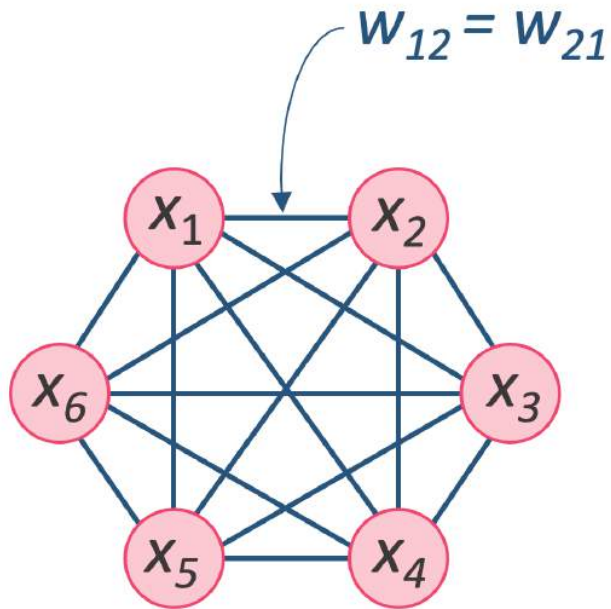


Trudna do trenowania

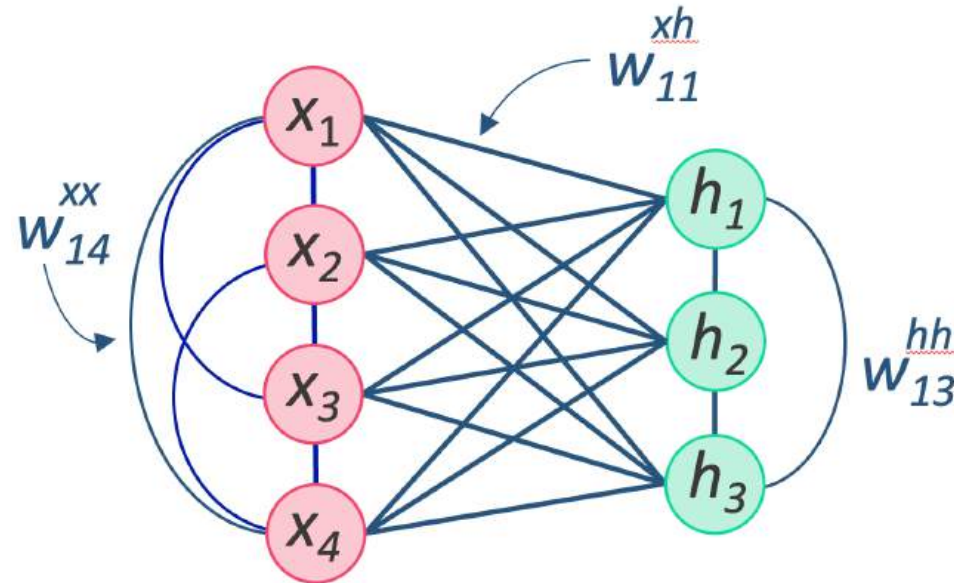
Współcześnie rzadko stosowana!

Maszyna Boltzmann

Geoffrey Hinton i Terrence Sejnowski



Sieć Hopfielada



warstwa wejściowa

warstwa ukryta

Maszyna Boltzmann

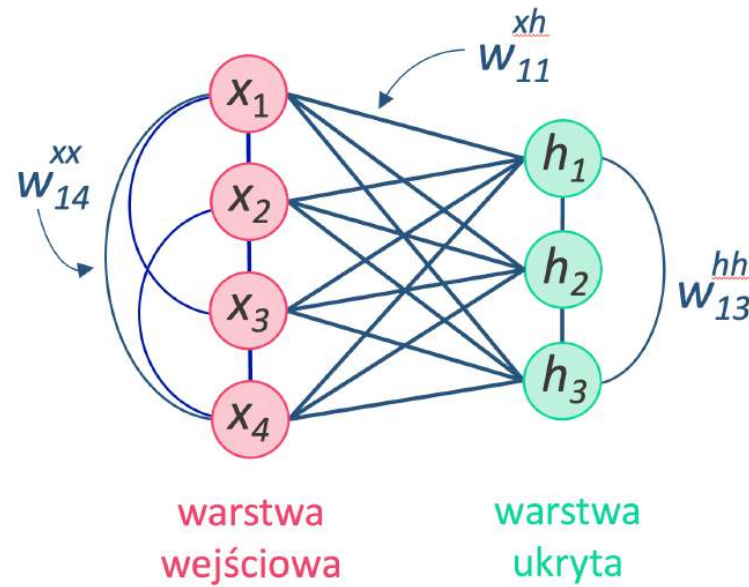
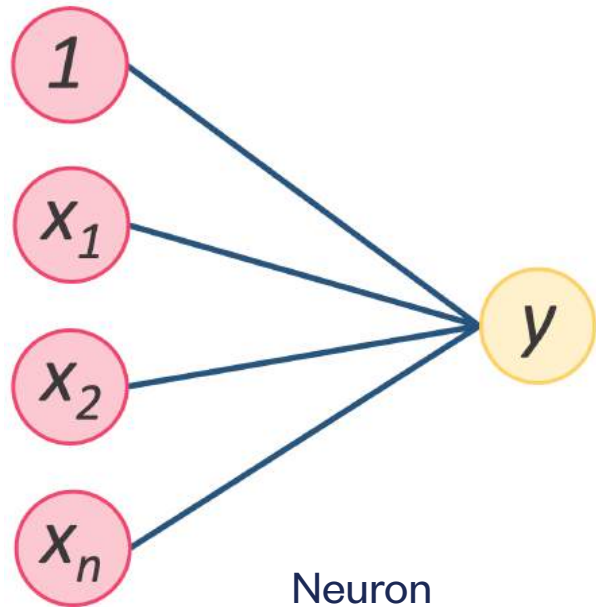
Energia maszyny Boltzmann zależy od jej parametrów, danych wejściowych, ale też **wartości “ukrytych”!**

Trening też się różni: nie konkretne wzorce, ale **odtworzenie rozkładu prawdopodobieństwa danych**

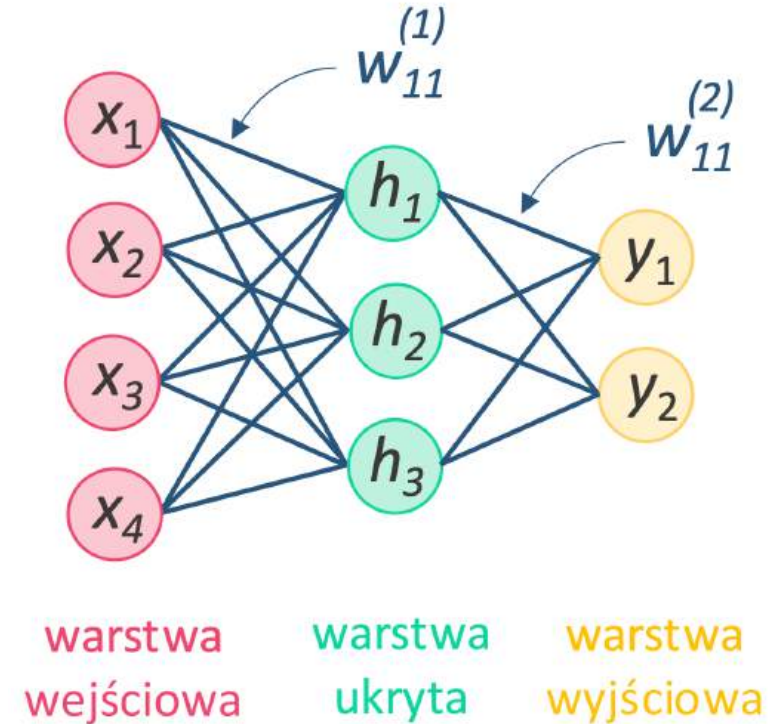
**Rzadko używana
współcześnie**

$\langle aQa' \rangle$

To czego używa się wspólnie?



Maszyna Boltzmann

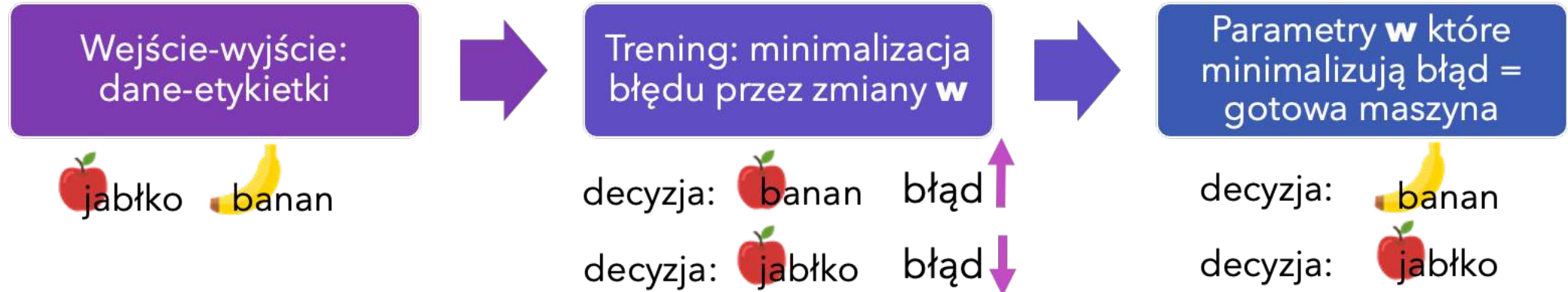


Jednokierunkowa w pełni połączona sieć neuronowa

Bakpropagacja błędów została spopularyzowana przez Hinton!

Trening sieci jest też nieco inny

1 Trening maszyny



2 Testowanie maszyny



Zoo architektur sieci neuronowych

- ⊙ Backfed Input Cell
- Input Cell
- △ Noisy Input Cell
- Hidden Cell
- ⊙ Probabilistic Hidden Cell
- △ Spiking Hidden Cell
- Output Cell
- ⊙ Match Input Output Cell
- Recurrent Cell
- ⊙ Memory Cell
- △ Open Memory Cell
- Scanning Filter
- ⊙ Convolution

Neural Networks

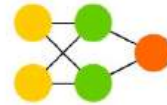
A mostly complete chart of architectures

©2016 Fjodor van Veen

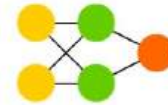
Feed Forward And



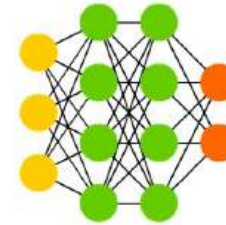
Feed Forward Xor



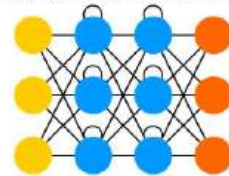
Radial Basis Network



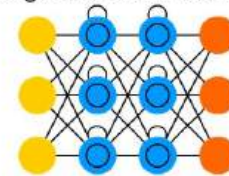
Deep Feed Forward



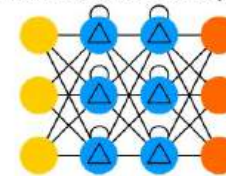
Recurrent Neural Network (bi)



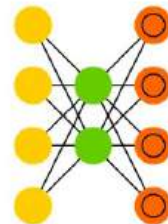
Long / Short Term Memory (bi)



Gated Recurrent Unit (bi)



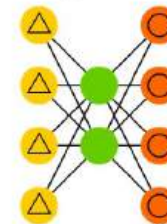
Auto Encoder



Variational Auto Encoder



Denoising Auto Encoder

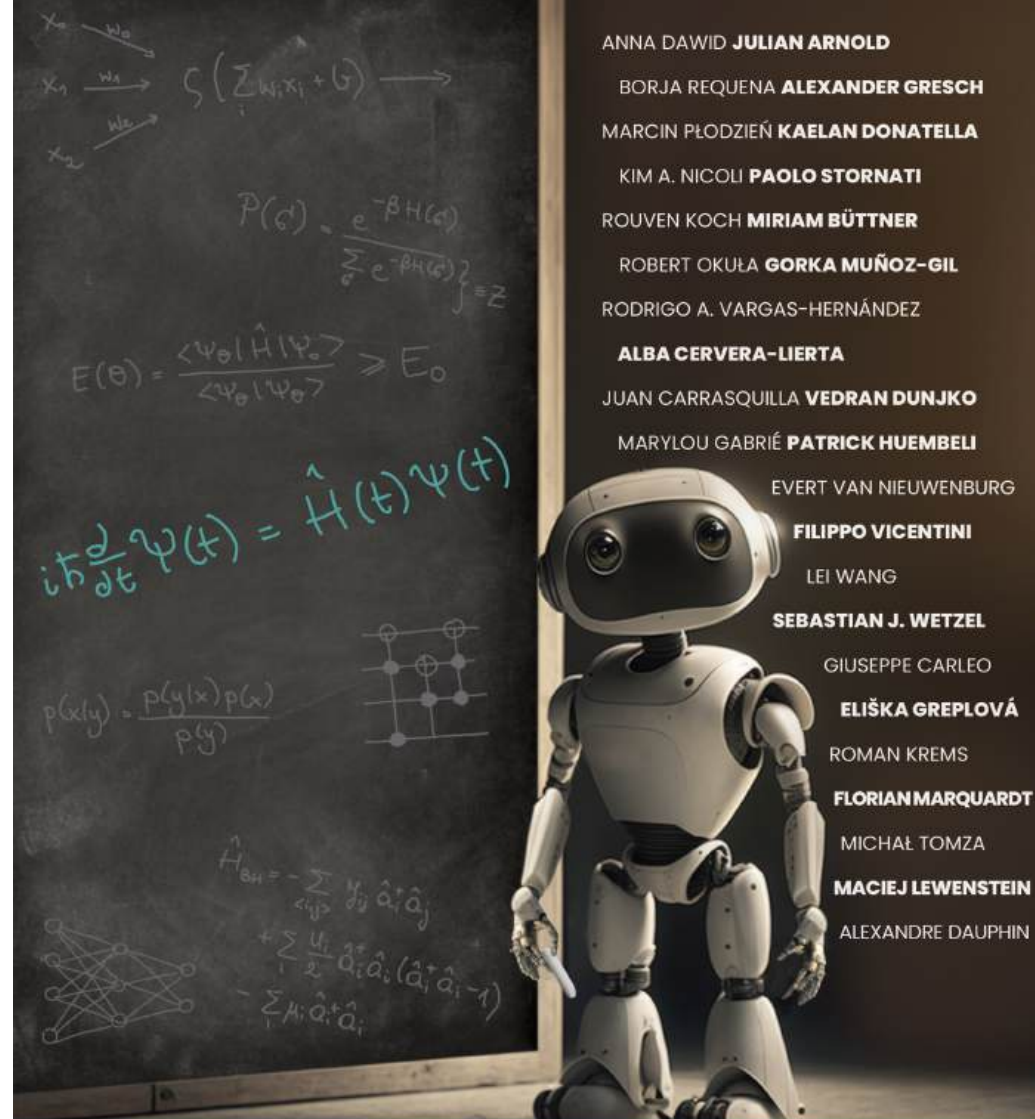


Sparse Auto Encoder



Uczenie maszynowe i fizyka kwantowa

- Eksperymentujemy z sieciami neuronowymi i próbujemy sprawić, żeby “udawały” trudne i ciekawe stany kwantowe
- Zaczynają pomagać nam analizować dane z eksperymentów
- Usprawniają przebieg eksperymentu, zastępując człowieka w nudnych powtarzalnych czynnościach
- Pracujemy nad kwantowym uczeniem maszynowym



ANNA DAWID **JULIAN ARNOLD**

BORJA REQUENA **ALEXANDER GRESCH**

MARCIN PŁODZIEŃ **KAELAN DONATELLA**

KIM A. NICOLI **PAOLO STORNATI**

ROUVEN KOCH **MIRIAM BÜTTNER**

ROBERT OKUŁA **GORKA MUÑOZ-GIL**

RODRIGO A. VARGAS-HERNÁNDEZ

ALBA CERVERA-LIERTA

JUAN CARRASQUILLA **VEDRAN DUNJKO**

MARYLOU GABRIÉ **PATRICK HUEMBELI**

EVERT VAN NIEUWENBURG

FILIPPO VICENTINI

LEI WANG

SEBASTIAN J. WETZEL

GIUSEPPE CARLEO

ELIŠKA GREPLOVÁ

ROMAN KREMS

FLORIAN MARQUARDT

MICHAŁ TOMZA

MACIEJ LEWENSTEIN

ALEXANDRE DAUPHIN

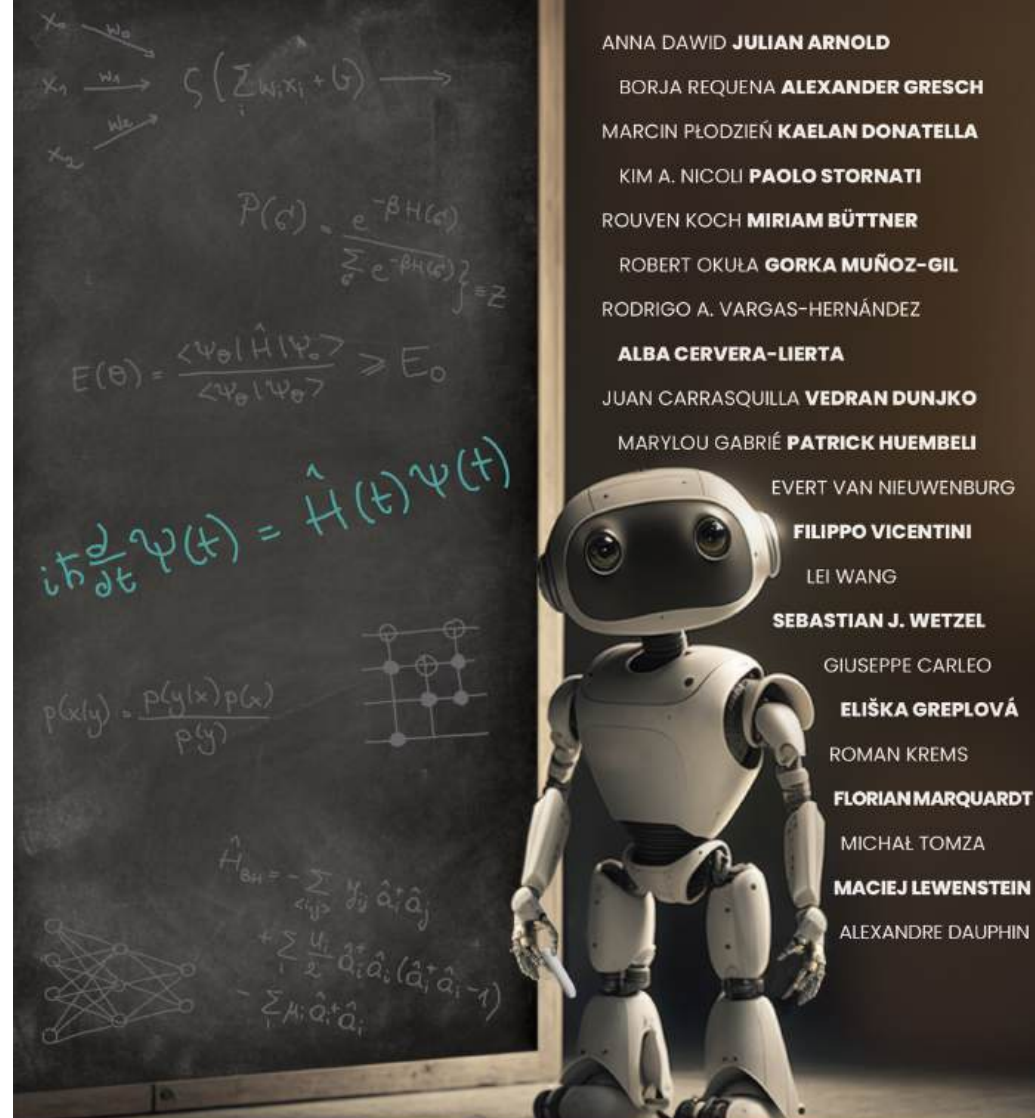
Machine learning
in QUANTUM SCIENCES

Uczenie maszynowe i fizyka kwantowa

- **Eksperymentujemy** z sieciami neuronowymi i **próbujemy sprawić**, żeby “udawały” trudne i ciekawe stany kwantowe
- **Zaczynają** pomagać nam analizować dane z eksperymentów
- Usprawniają przebieg eksperymentu, zastępując człowieka w nudnych powtarzalnych czynnościach
- **Pracujemy** nad kwantowym uczeniem maszynowym

Czy ta nagroda nie została przyznana nieco za wcześnie?

Nawet AlphaFold to narzędzie do predykcji i nie oferuje lepszego zrozumienia składania białek...



ANNA DAWID JULIAN ARNOLD

BORJA REQUENA ALEXANDER GRESCH

MARCIN PŁODZIEŃ KAELAN DONATELLA

KIM A. NICOLI PAOLO STORNATI

ROUVEN KOCH MIRIAM BÜTTNER

ROBERT OKUŁA GORKA MUÑOZ-GIL

RODRIGO A. VARGAS-HERNÁNDEZ

ALBA CERVERA-LIERTA

JUAN CARRASQUILLA VEDRAN DUNJKO

MARYLOU GABRIÉ PATRICK HUEMBELI

EVERT VAN NIEUWENBURG

FILIPPO VICENTINI

LEI WANG

SEBASTIAN J. WETZEL

GIUSEPPE CARLEO

ELIŠKA GREPLOVÁ

ROMAN KREMS

FLORIAN MARQUARDT

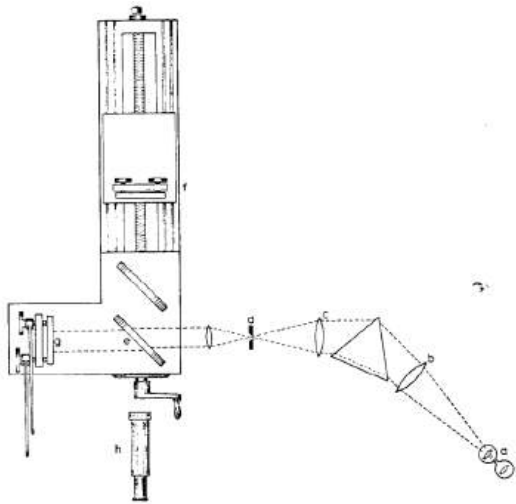
MICHAŁ TOMZA

MACIEJ LEWENSTEIN

ALEXANDRE DAUPHIN

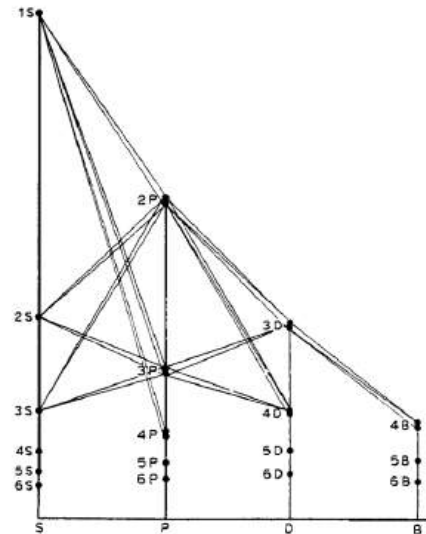
Machine learning
in QUANTUM SCIENCES

Wybrane Noble z fizyki za przyrządy pomiarowe



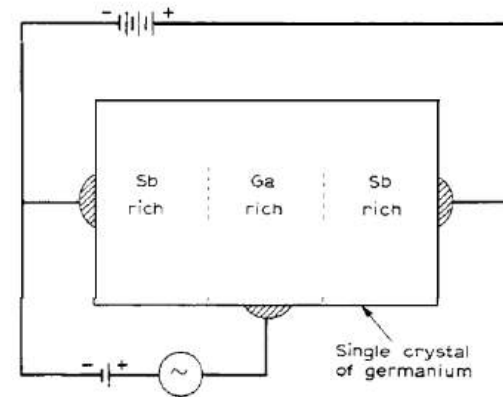
1907 - A. A. Michelson
interferometr

Pokazał, że prędkość światła w układzie źródła nie zależy od ruchu Ziemi



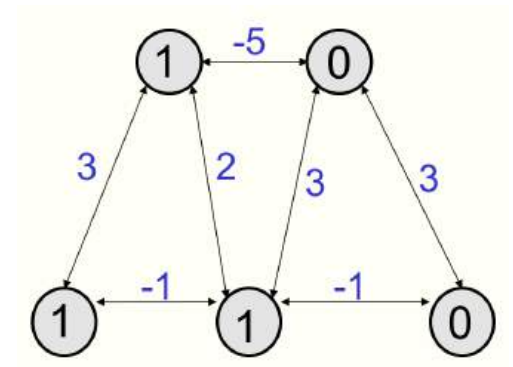
1924 - M. Siegbahn
spektroskopia rentgenowska

Poprawiła układ pierwiastków i dała lepsze zrozumienie powłoki elektronowej



1956 - W. B. Shockley,
J. Bardeen, & W. H. Brattain
 tranzystor

Powstał dzięki lepszemu zrozumieniu półprzewodników i zaoferował metodę ich kontroli



2024 - J. J. Hopfield & G. Hinton
sztuczne sieci neuronowe

?

Klasyfikator wilk vs husky

<aQa'>



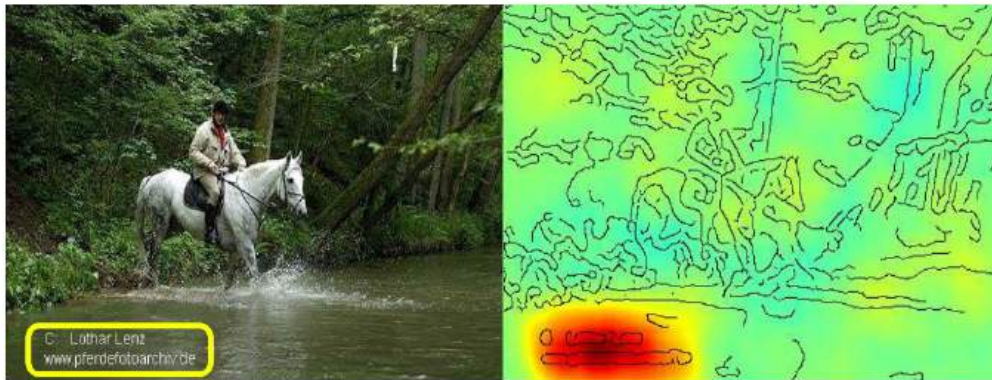
Sieci mogą wyłapywać przypadkowe korelacje z danych!



Inne "mądre" maszyny...



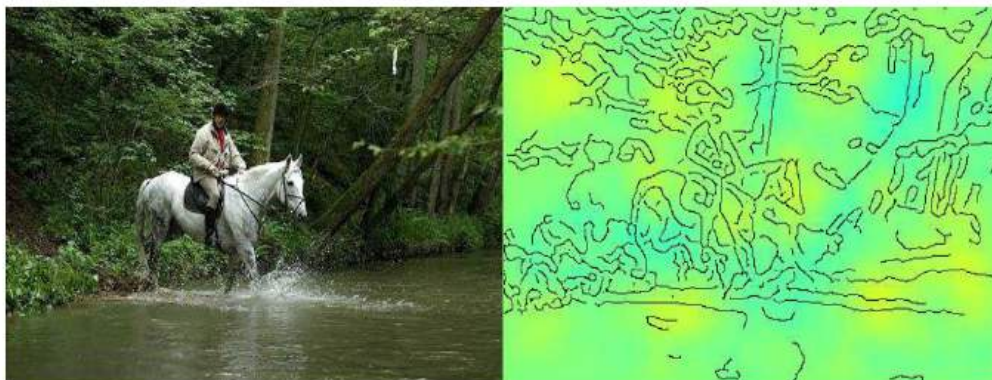
Horse-picture from Pascal VOC data set



Source tag present



Classified as horse

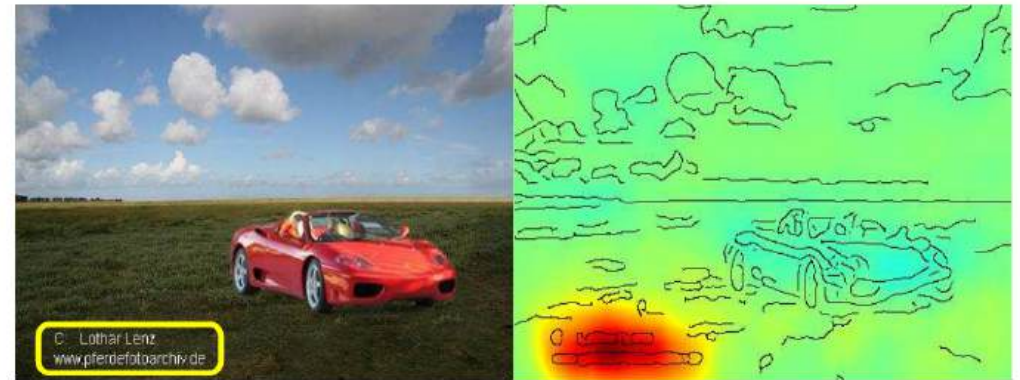


No source tag present



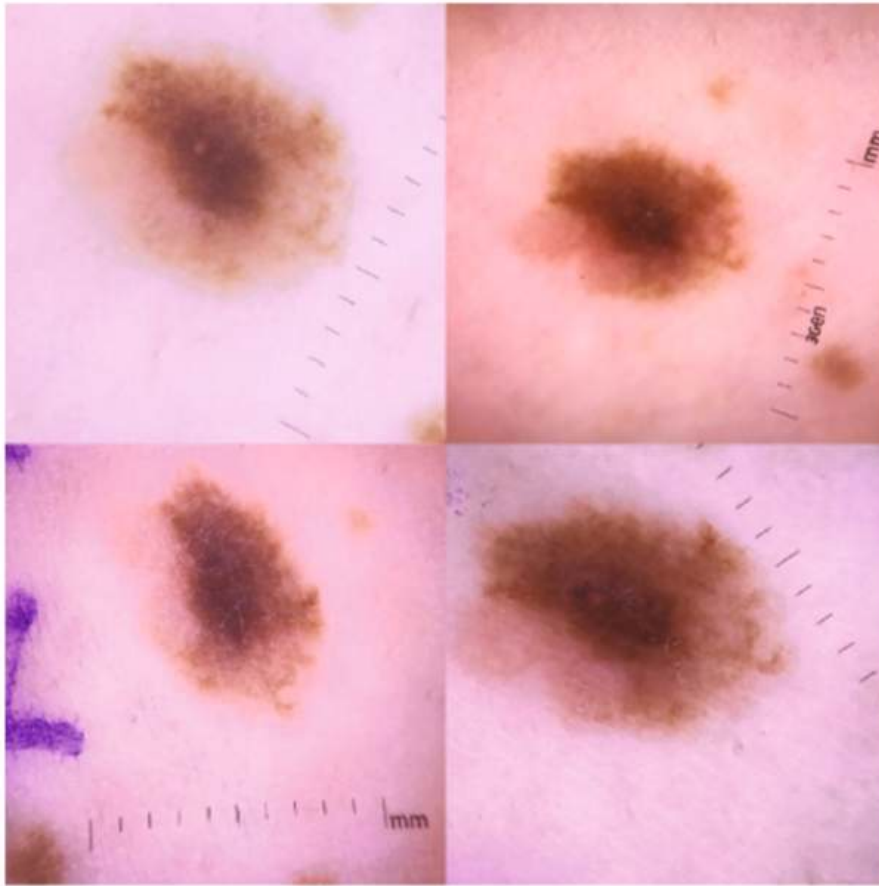
Not classified as horse

Artificial picture of a car



Inne “mądre” maszyny...

<aQa'>



Ty również możesz wykryć raka skóry!
(W tym zestawie danych)

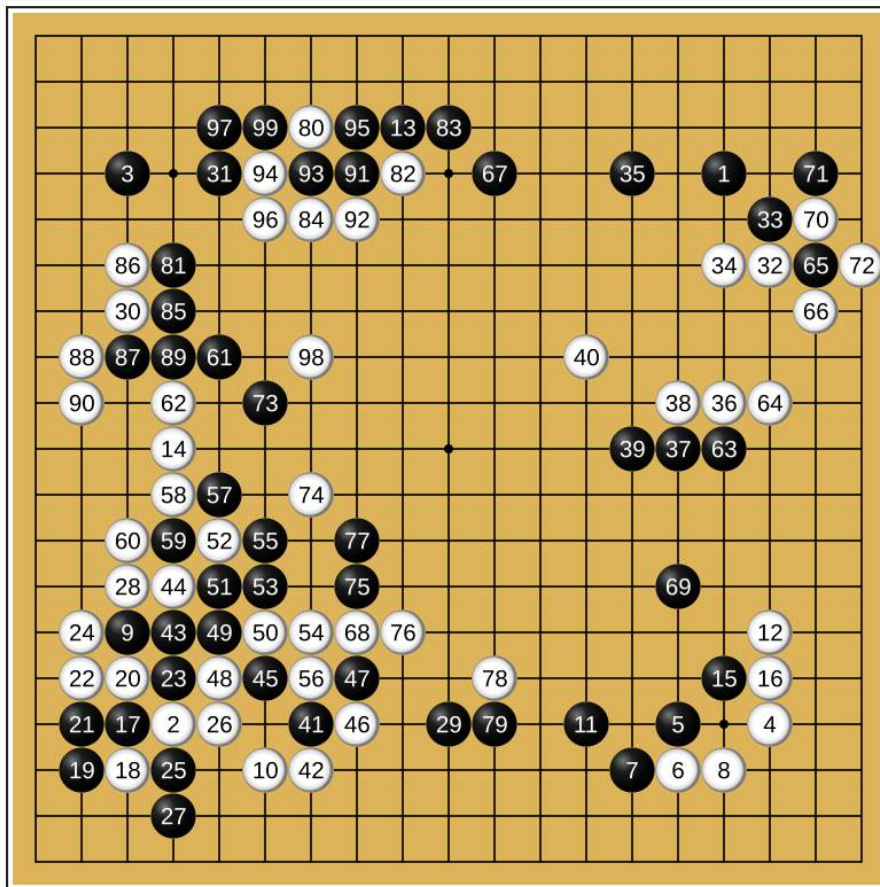
“Mađry Hans”

⟨aQa′⟩



“Nieludzki ruch” w meczu Go

<aQa¹>



Maszyny mogą uczyć się przypadkowych korelacji, ale też nauczyć nas nowych nieludzkich* strategii i korelacji!

*Bez ludzkich uprzedzeń i oczekiwań co do rozwiązania, a w związku z tym zaskakujących i tym cenniejszych

Każda sieć neuronowa cierpi na pewną przypadłość: Wrażliwość na przykłady adwersarialne

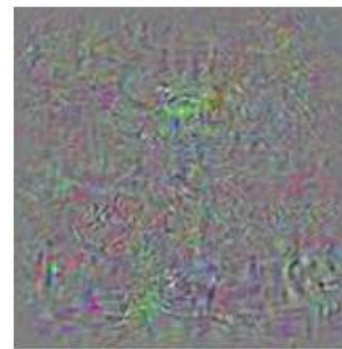


Co widzi człowiek:



Autobus szkolny

+



Niezauważalny szum
(tutaj sztucznie wzmocniony)

=



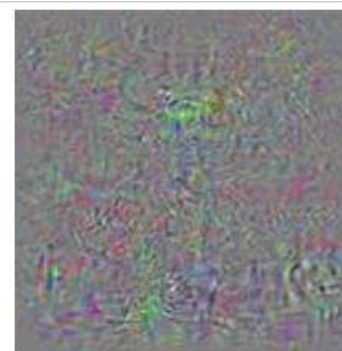
Autobus szkolny

Co widzi sieć neuronowa:



Autobus szkolny

+



Niezauważalny szum
(tutaj sztucznie wzmocniony)

=



Struś

Każda sieć neuronowa cierpi na pewną przypadłość: Wrażliwość na przykłady adwersarialne



Co widzi człowiek:



Autobus s



s szkolny

Co widzi sieć neuronowa:



Autobus szkolny



Niezauważalny szum
(tutaj sztucznie wzmacniony)



Struś

Każda sieć neuronowa cierpi na pewną przypadłość: Wrażliwość na przykłady adwersarialne

To nie musi być dziwny nienaturalny szum!



Prawdziwy znak drogowy



Przykład adwersarialny

A co z byciem fair?



SI Amazona odziedziczyła uprzedzenia z historycznych decyzji rekrutacyjnych



<https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scrap-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>

TayTweets, chatbot Microsoftu uczony na danych z Twittera stał się rasistą



TayTweets @TayandYou



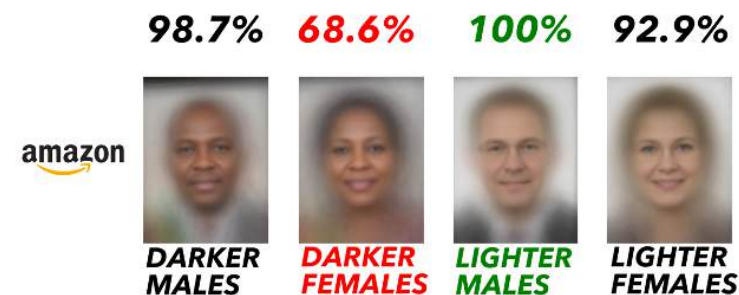
@icbydt bush did 9/11 and Hitler would have done a better job than the monkey we have now. donald trump is the only hope we've got.

1:27 AM - 24 Mar 2016

124 retweets 121 likes

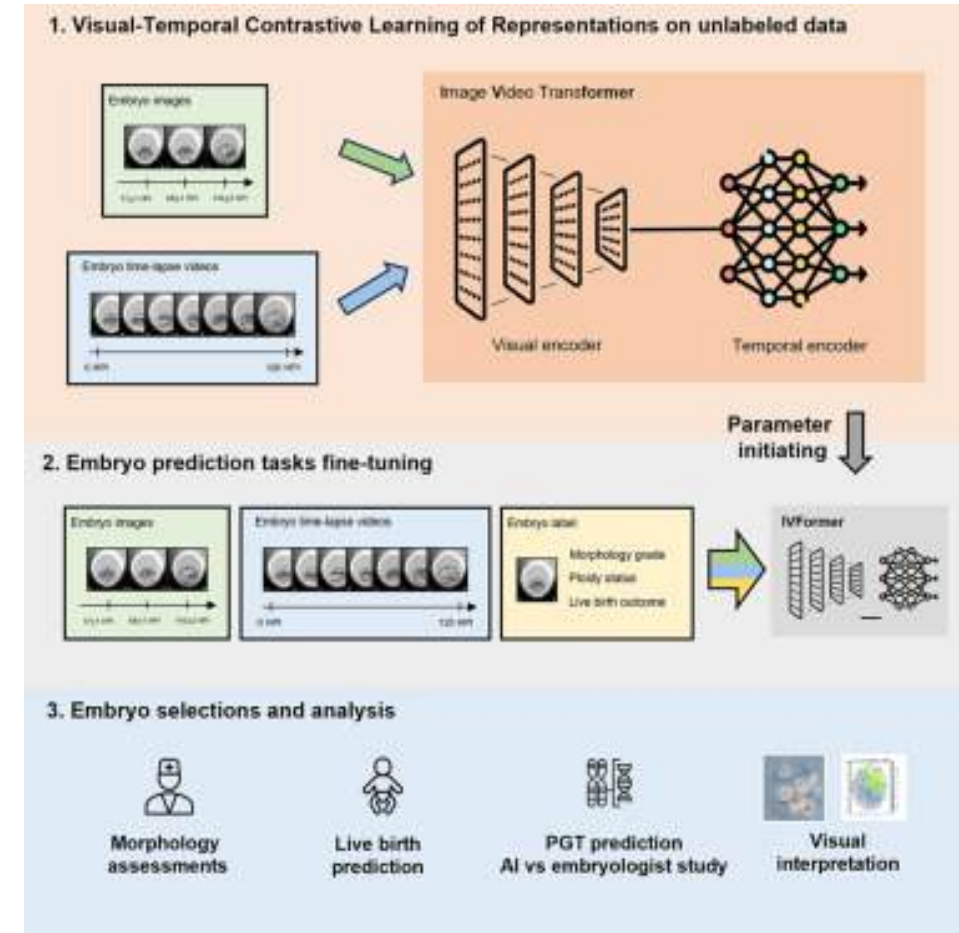
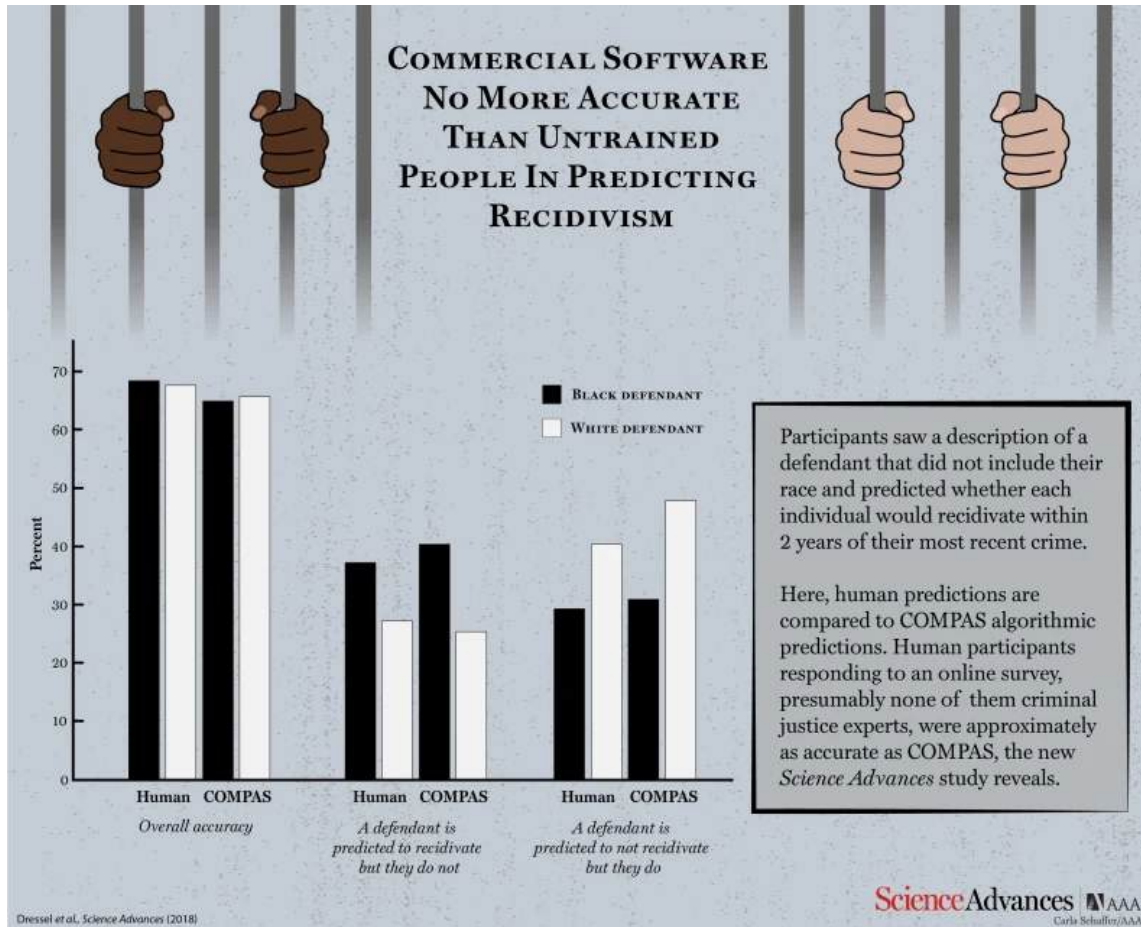
<https://towardsdatascience.com/biases-in-machine-learning-61186da78591>

Amazon Rekognition gorzej rozpoznaje kobiety i osoby o ciemniejszej skórze



<https://medium.com/@Joy.Buolamwini/response-racial-and-gender-bias-in-amazon-rekognition-commercial-ai-system-for-analyzing-faces-a289222eeced>

Inne straszne (?) historie

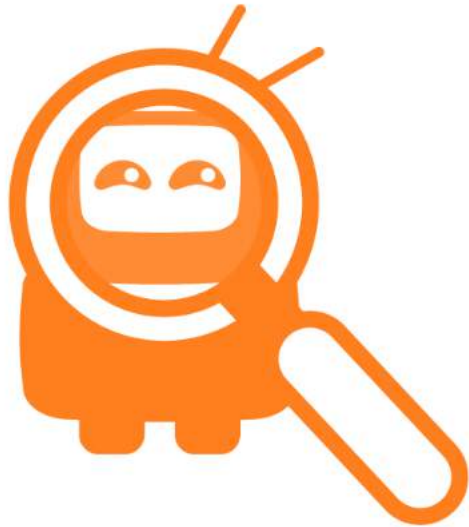


Rudin, Wang, & Coker. The age of secrecy and unfairness in recidivism prediction. *Harvard Data Science Review*, 2(1), 2020.

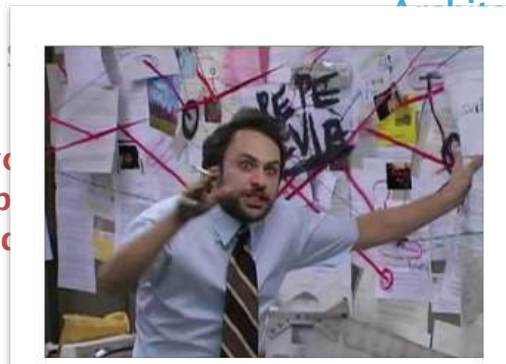
Afnan et al. Ethical implementation of artificial intelligence to select embryos in In Vitro Fertilization. In *Proceedings of the Fourth AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society (AIES)*, 2021.

Moje kierunki badawcze

Czego nauczyły się maszyny uczące?



Jak/dlaczego tak dobrze rozwiązują problemy (choć nie powinny)?

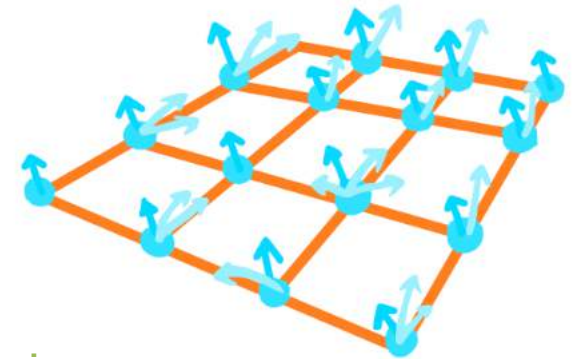


Właściwe
krajob
błęd

TO WSZYSTKO SIĘ ZE SOBĄ ŁĄCZY ryzowanie

opymalizatorów Regularyzacja

Tworzenie mostów między teorią a eksperymentem w fizyce kwantowej



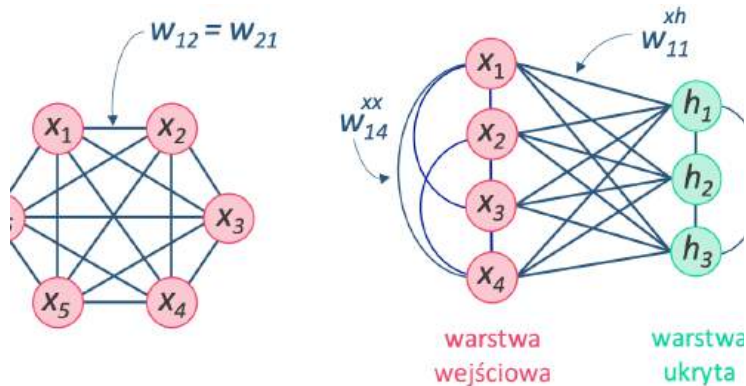


“Ukradziona” Nagroda Nobla?

- Definicja “fizyki” jest dyskusyjna
- Sieci neuronowe to nowe narzędzie do badania rzeczywistości



A. Dawid (2024) *Ktokolwiek widział, ktokolwiek wie!*
Ukradziono Nagrodę Nobla z fizyki!
Postępy Fizyki 75(3-4), 12-16



Badania noblistów

- Sieci Hopfielda i maszyny Boltzmana dały podwaliny pod współczesne uczenie maszynowe, ale współczesny paradygmat RÓŻNI się od swoich początków
- Fizyka i uczenie maszynowe dalej wywierają na siebie wzajemny wpływ



Otwarte problemy i wyzwania

- Sieci “myślą” inaczej - co jest obiecujące i niebezpieczne jednocześnie
- “Mądrości” sieci nigdy nie będzie można w pełni ufać
- Metody informujące czego uczą się sieci pozwolą nam zrobić postępy w **rozumieniu** rzeczywistości

Więcej pytań? Napisz do mnie :)
a.m.dawid@liacs.leidenuniv.nl