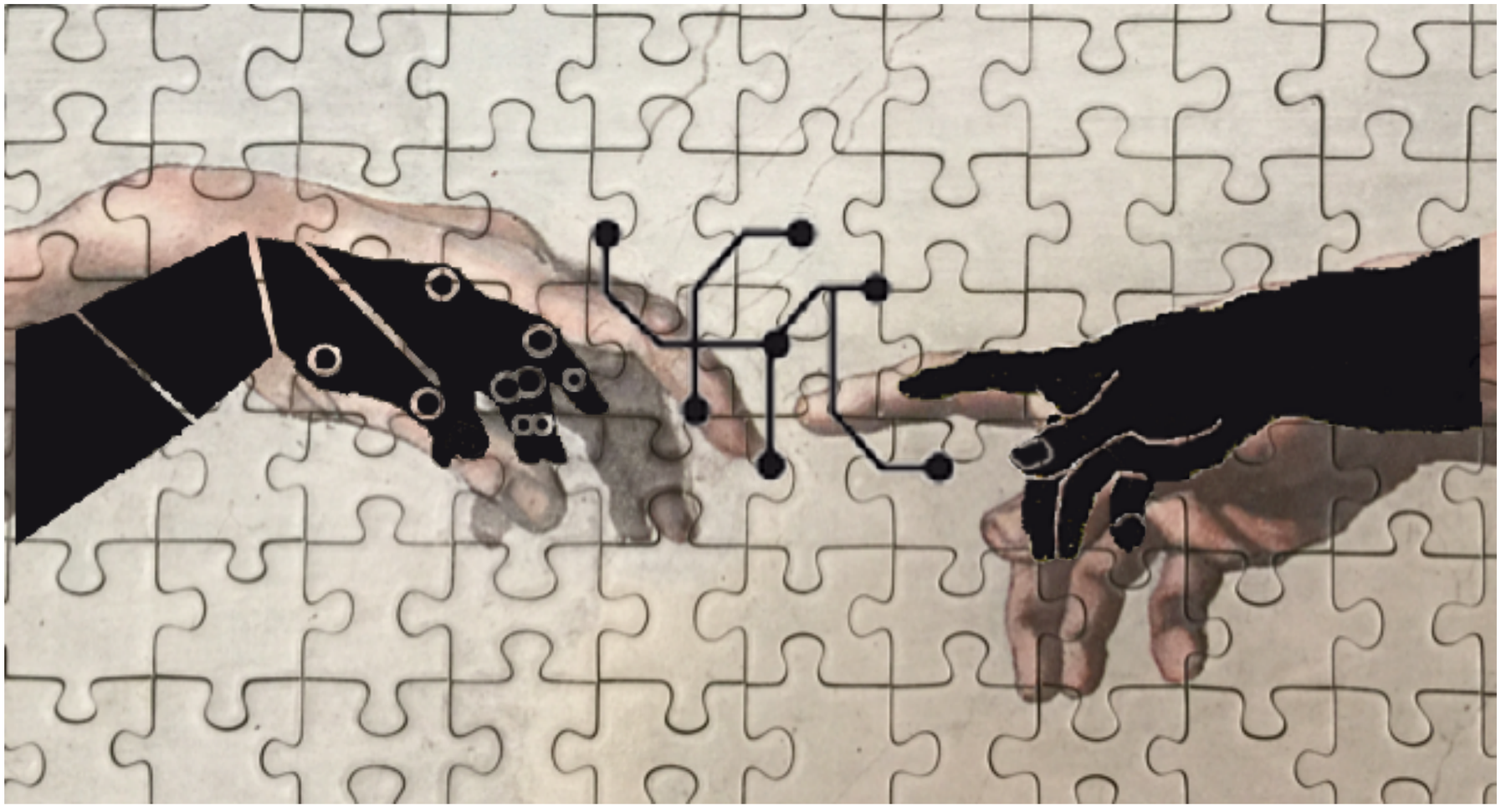
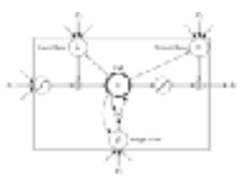
AI Artificial Intelligence +

The future of AI?

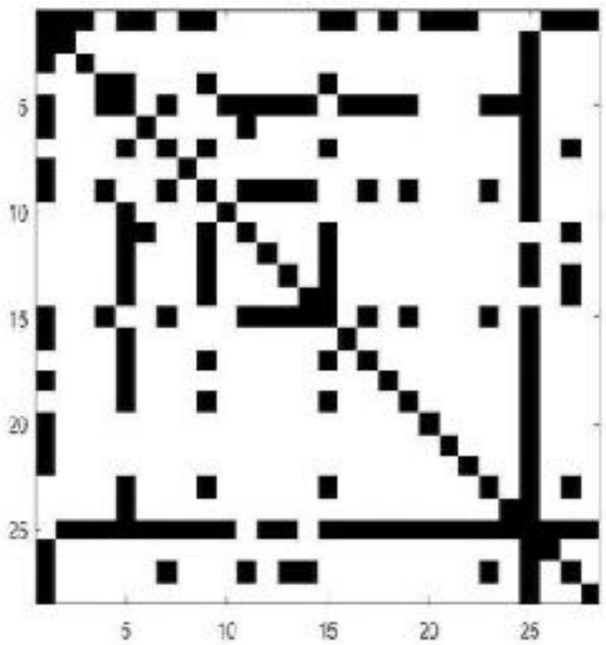


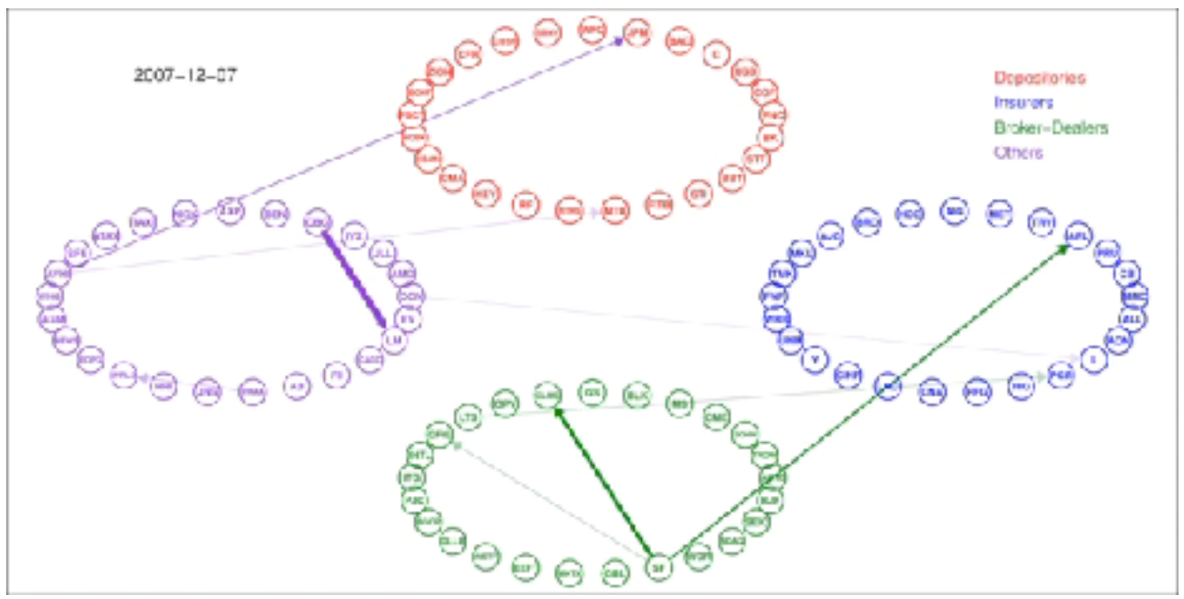
Motivation



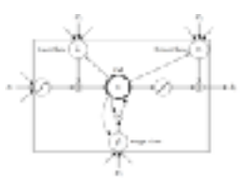
AI Artificial Intelligence

Micro to Macro



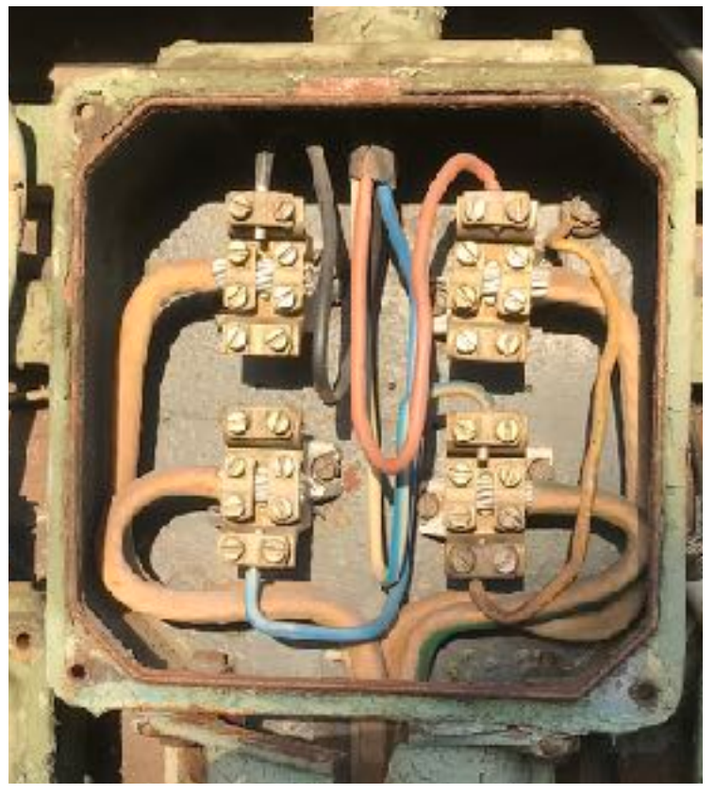


- AI technologies connect, empower and expose individuals
- Dynamic interactions ask for smart data science
- Digital society creates new economic opportunities

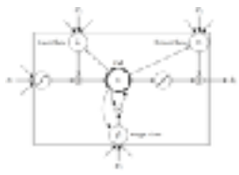


Basic element of Artificial Intelligence

…is a buzzer or a switch!



Combined with other switches …

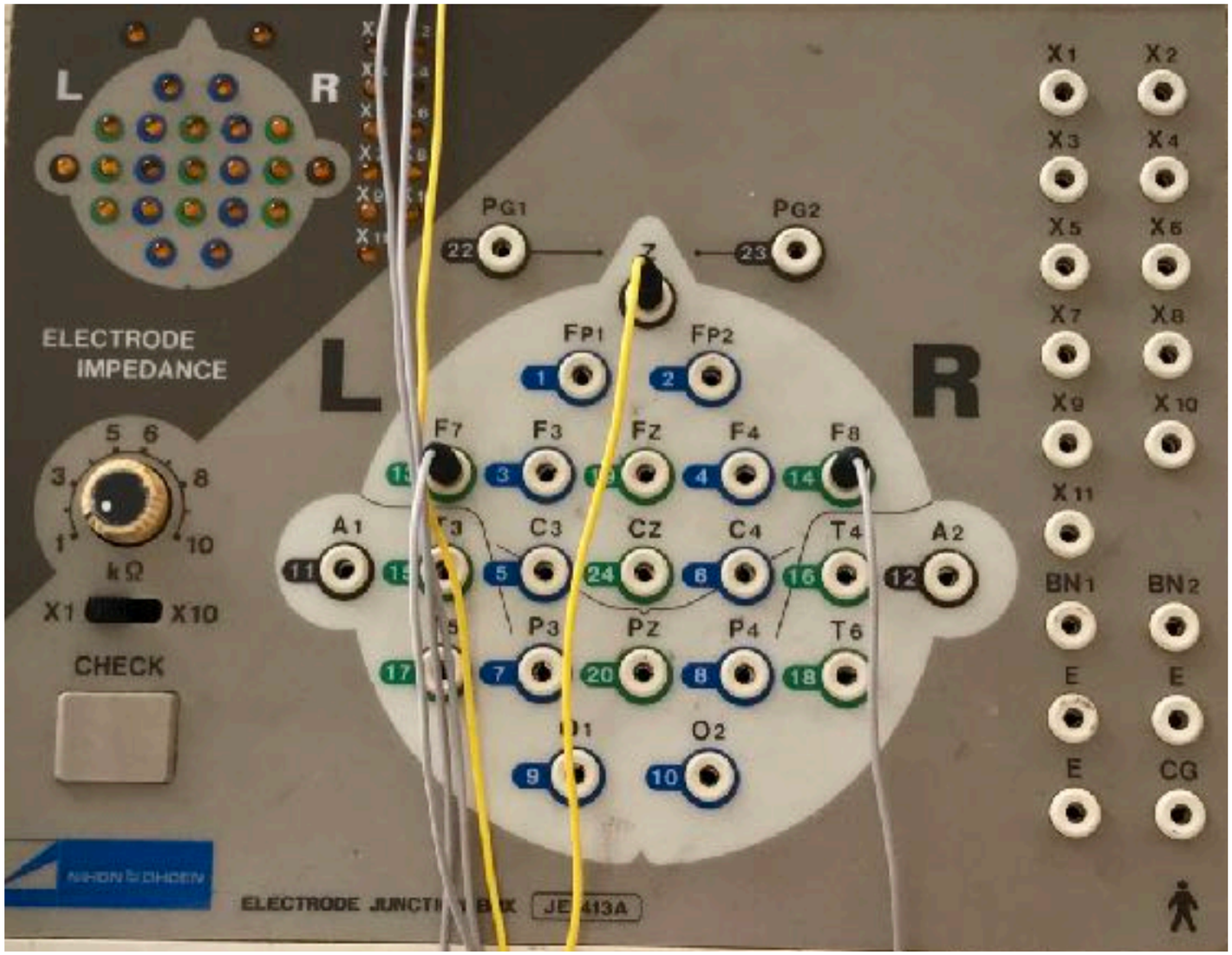


Basic elements combined

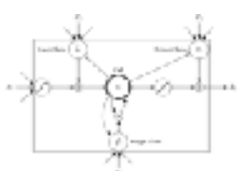
…it can do complex stuff!

…like recording an EEG

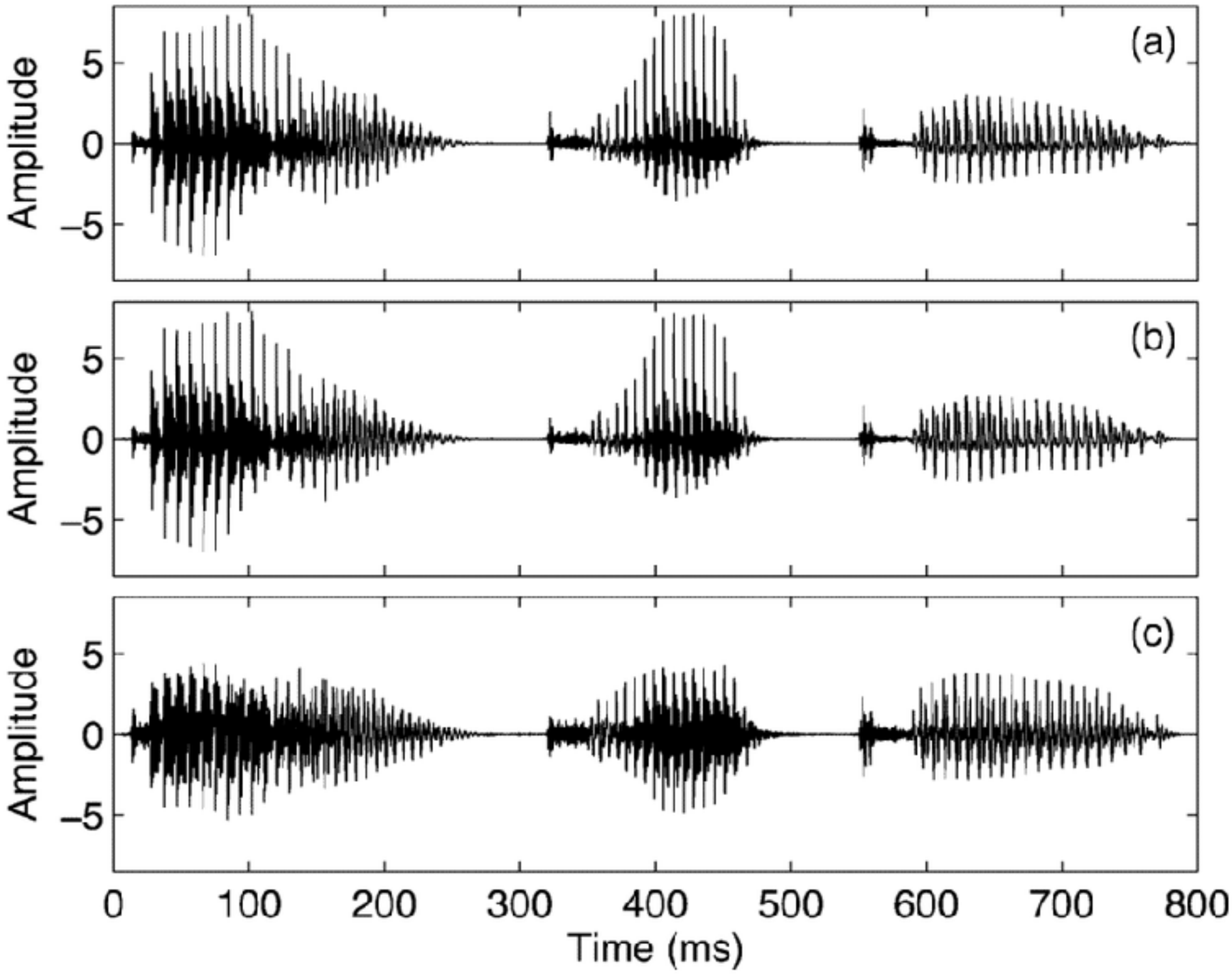
…classifying it

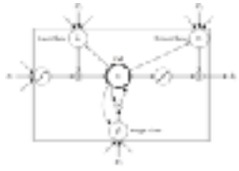


…and propose diagnostic follow ups!

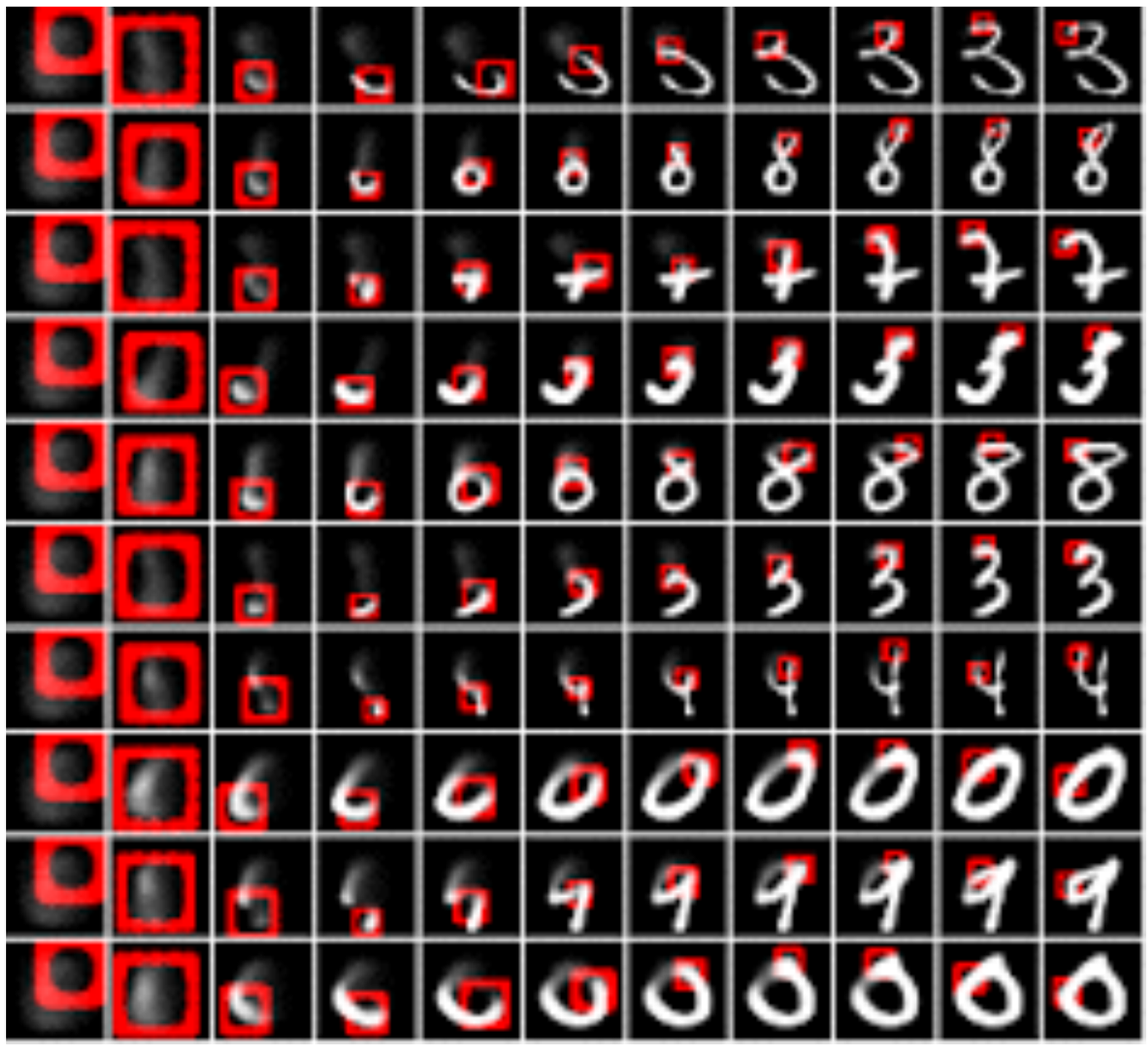


AI is good at: EEG, speech, handwriting, …

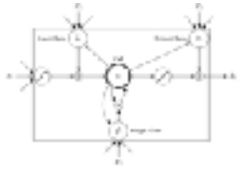




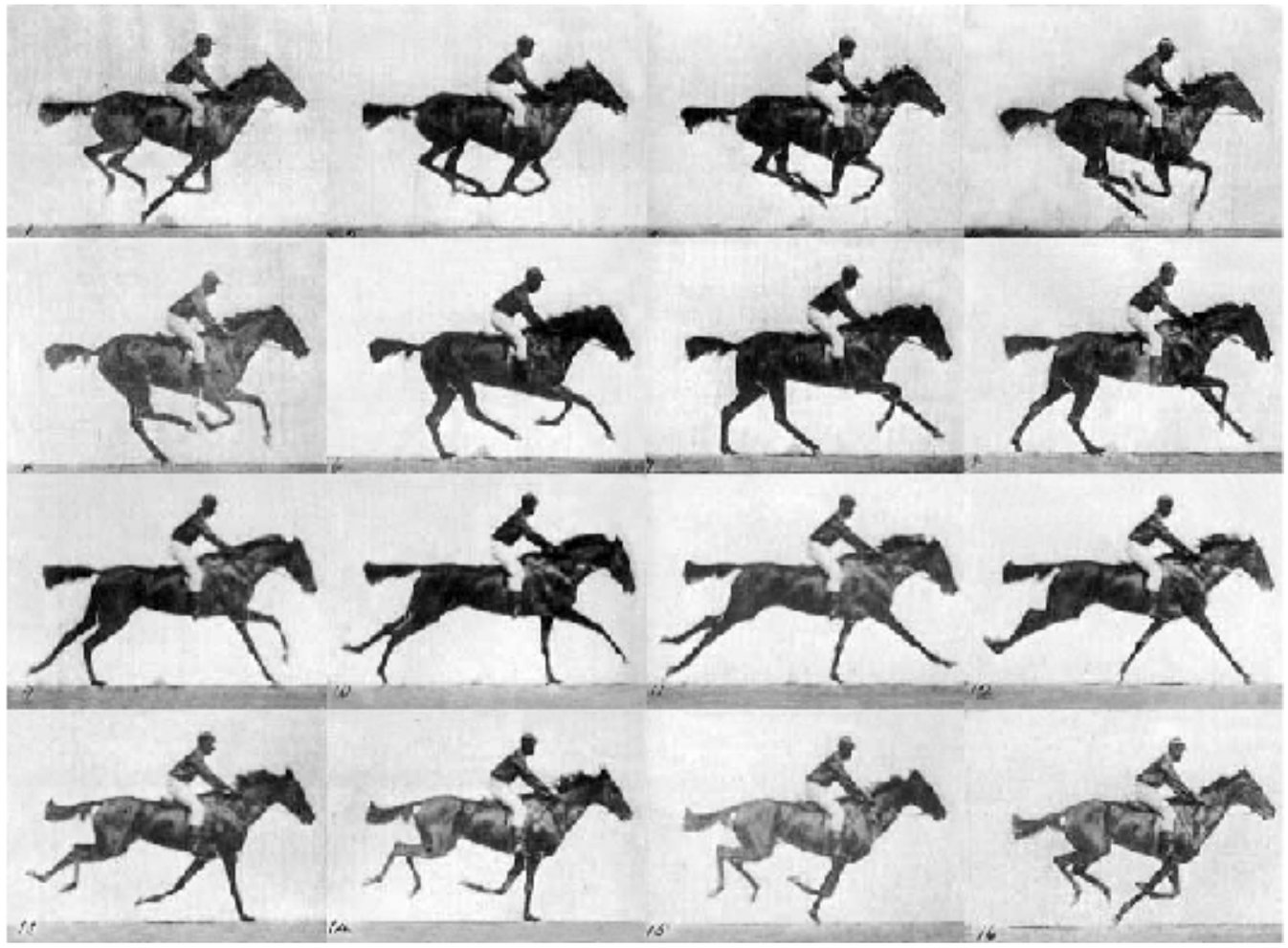
AI is good at: speech, handwriting, moves, …

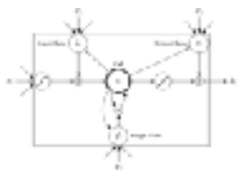


A trained DRAW network generating MNIST digits:
Each row shows successive stages in the generation of a single digit

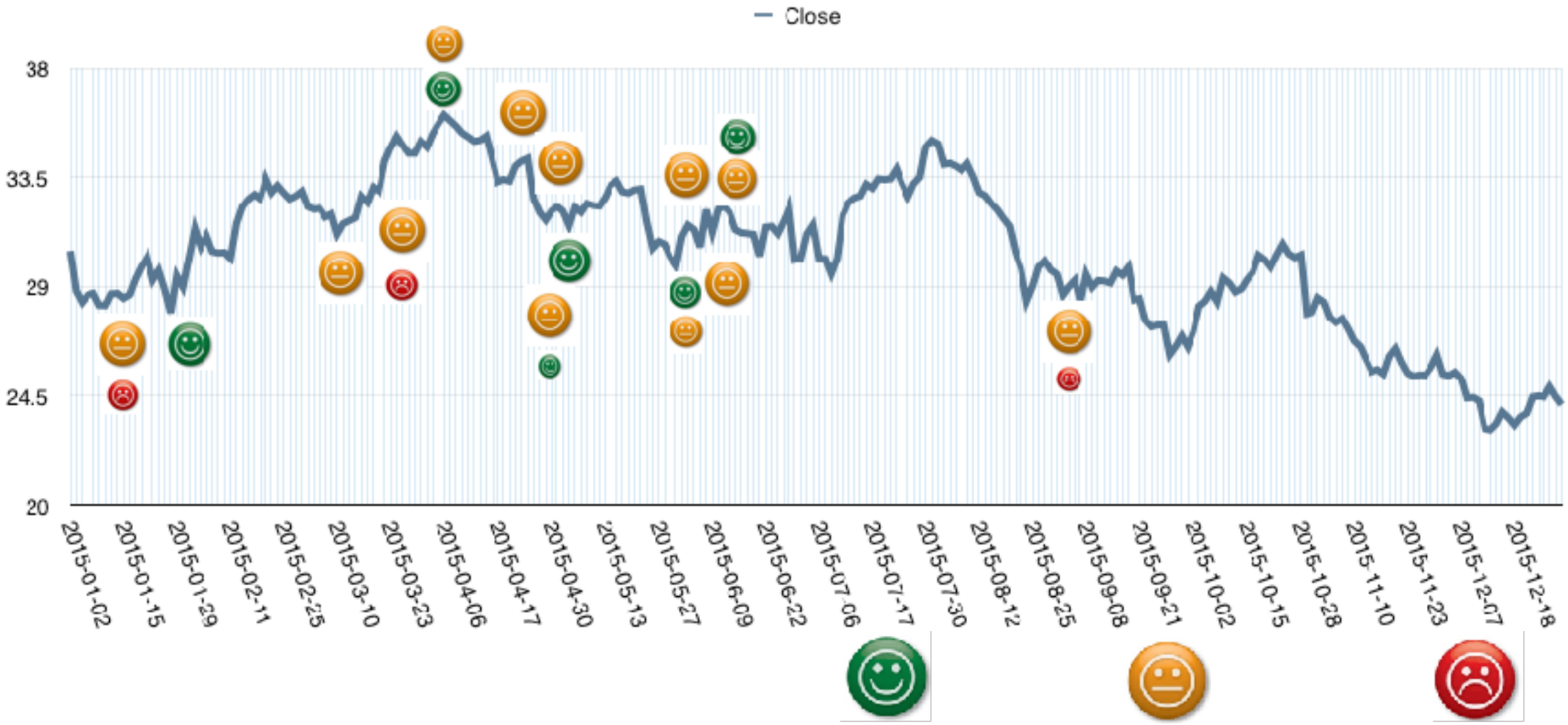


AI is good at: moves, x-rays, faces, …

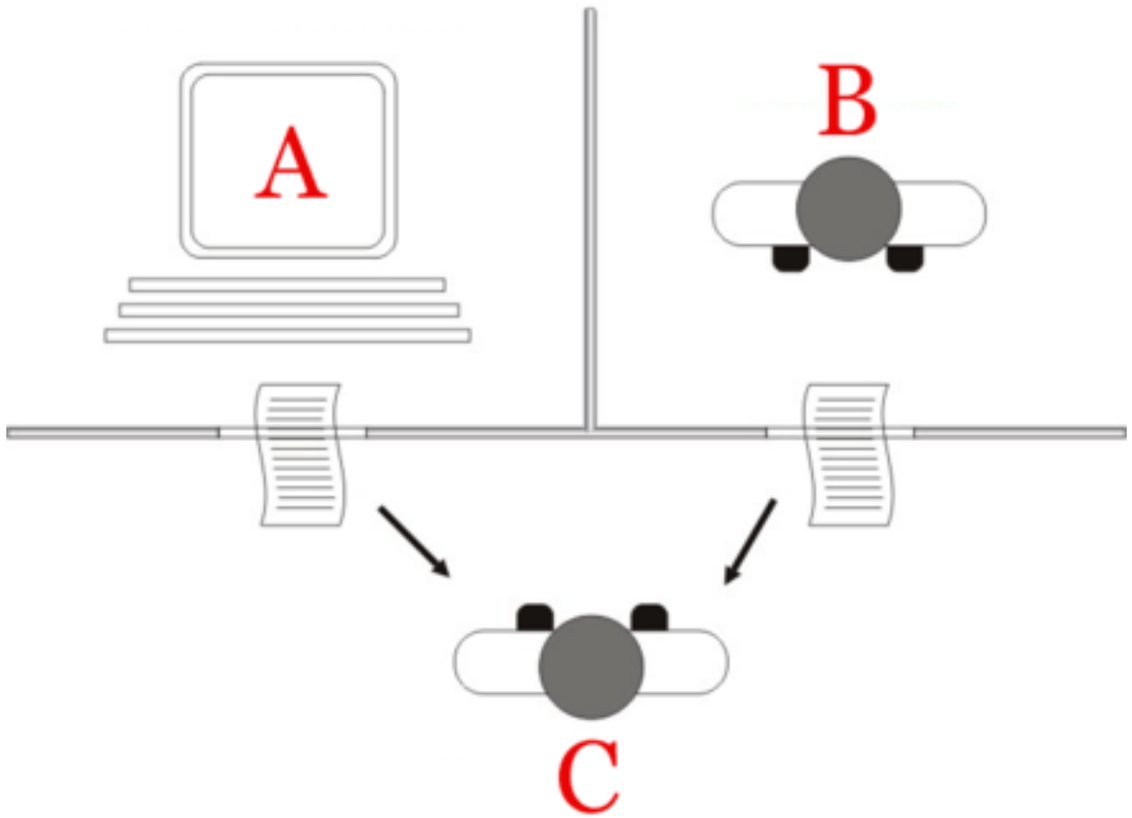


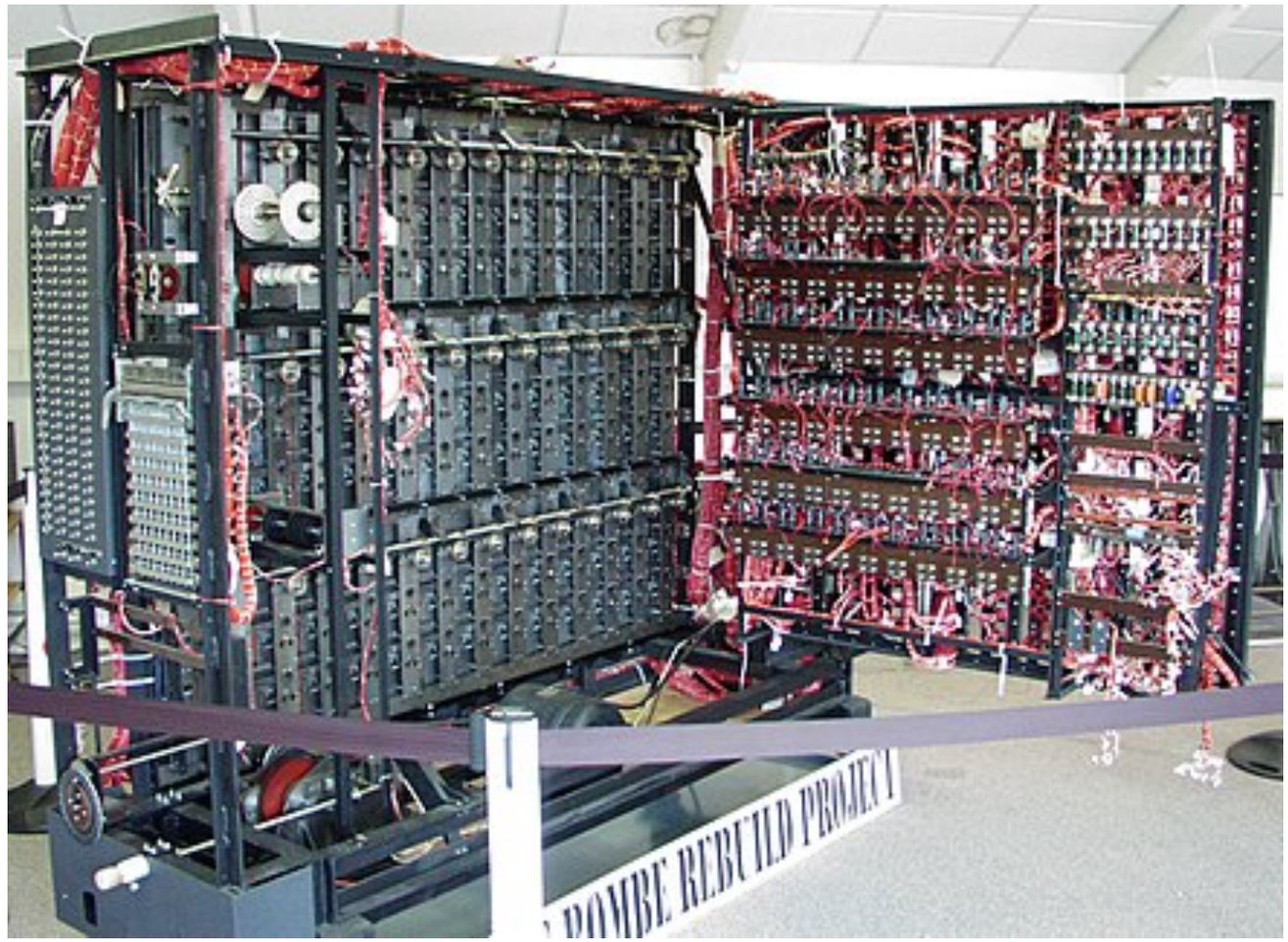


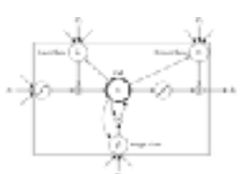
AI is good at: x-rays, face values, birds, …



AI is not human: Turing Test

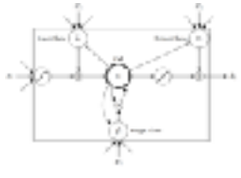




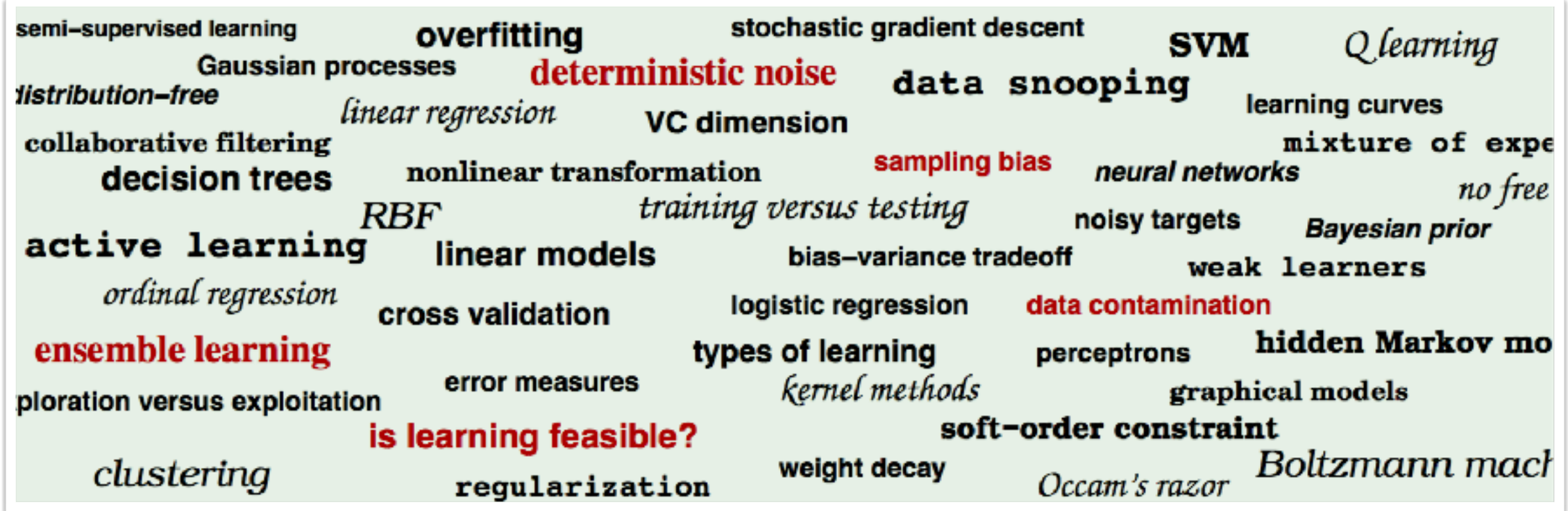


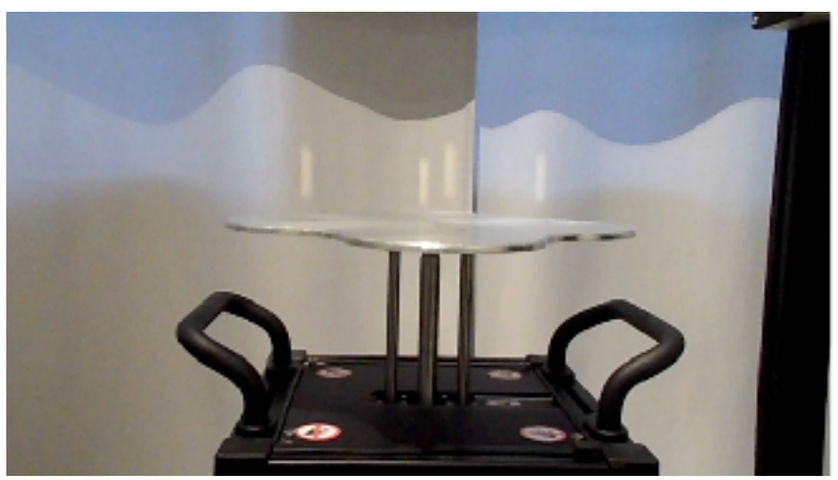
Outline

1. Motivation ✓
2. AI applications
3. AI Details
4. AI -> ?
5. References



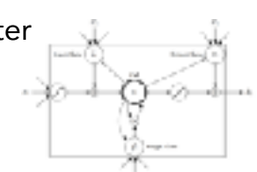
ML Machine Learning an old idea !





W Phillips (curve), hydro engineer/economist, MONIAC Monetary National Income Analogue Computer

work.caltech.edu/library/181.html

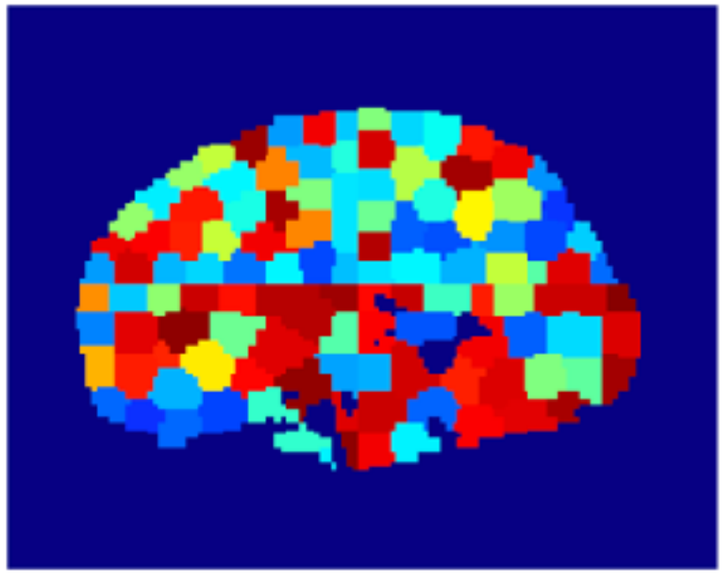


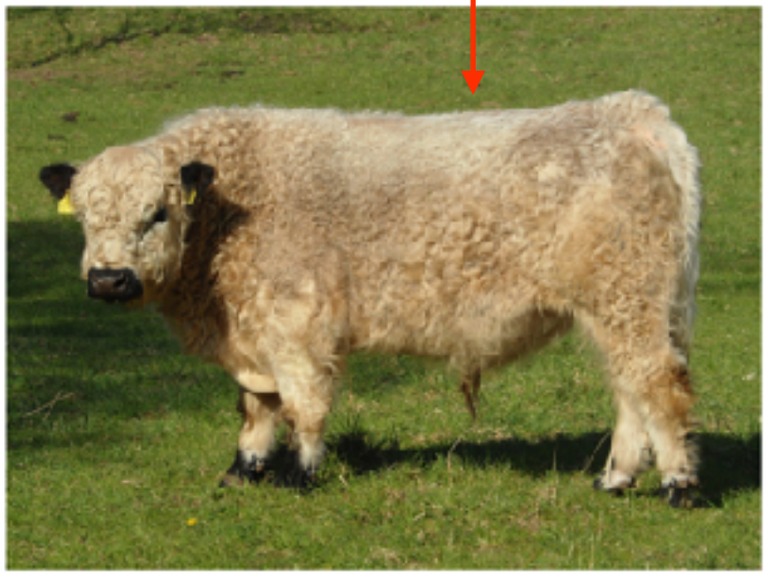
Risk Perception

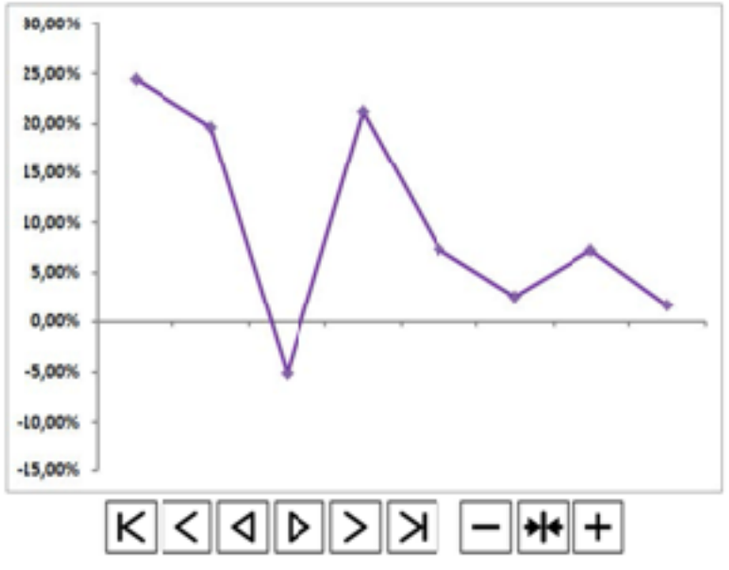
- RMI brain records 1M data points per 2 sec
- RMI on RPID tasks correlates with RP
- Stationary inputs, AI copies your RP!

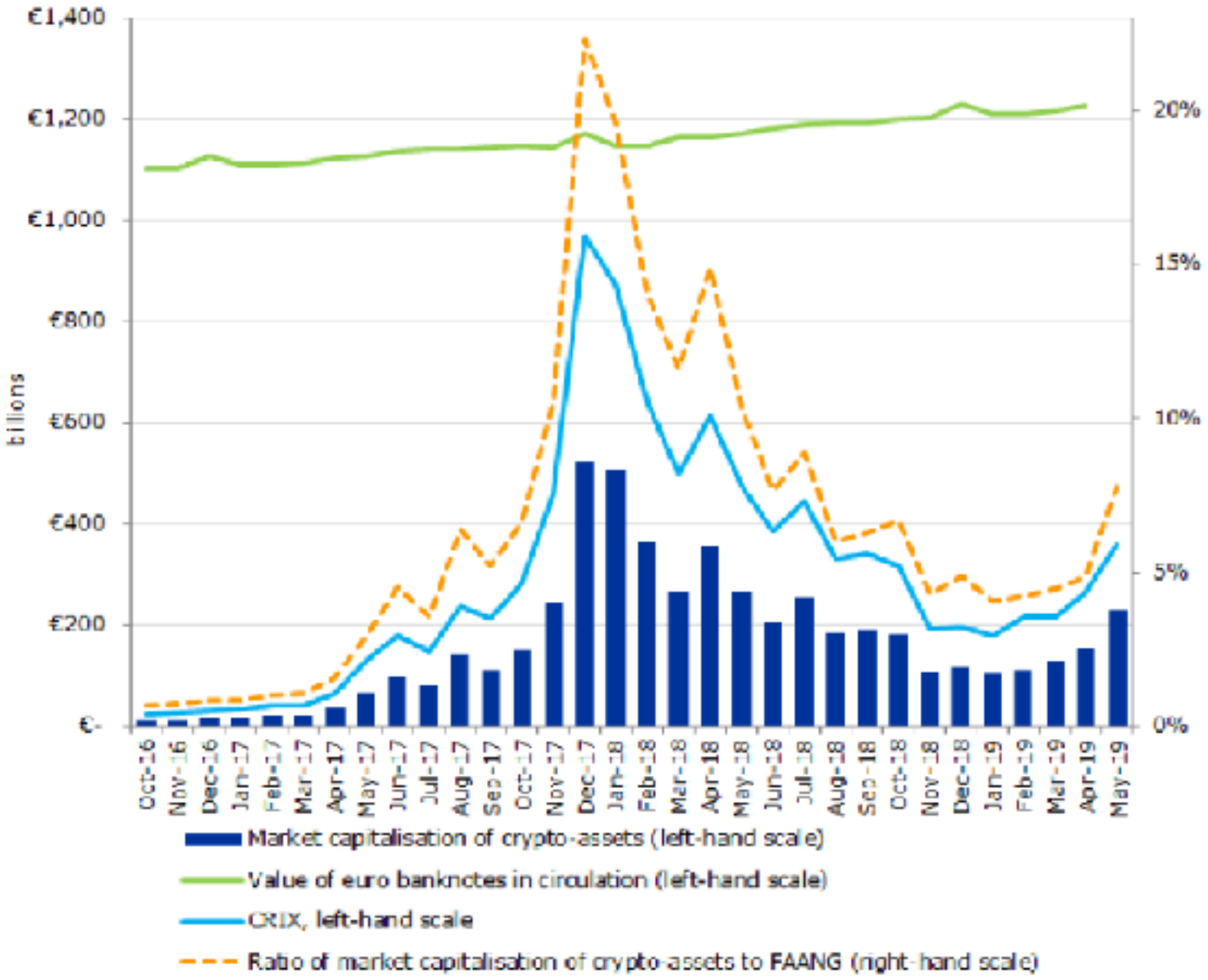
bull?

Chen Y, Härdle WK, Qiang H, Majer, P (2018)
Risk Related Brain Regions Detected with 3D Image FPCA, Statistics and Risk Modeling,



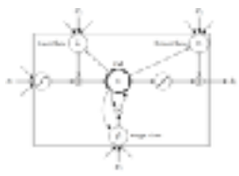






crix.berlin, **thecrix.de**

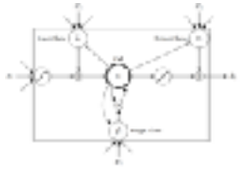
- CRyptocurrency IndeX / **CRIX**
- An ECB benchmark index
- Used by ETFs and Investment funds



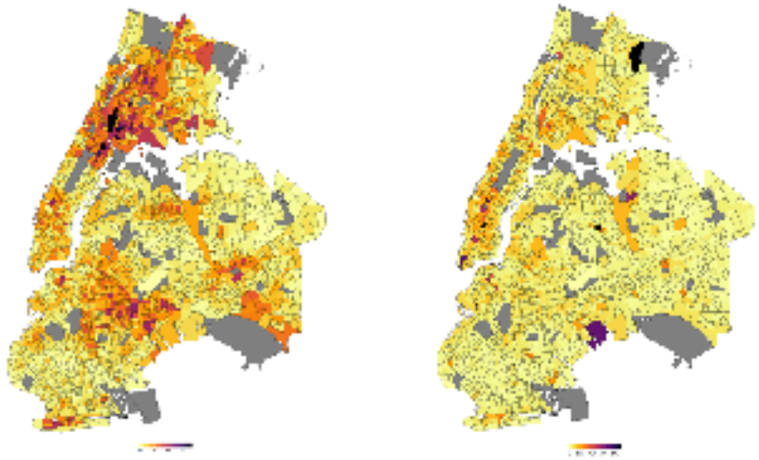
The unstoppable rise of robo-advisors

- Online, automated portfolio management service
- Based on risk tolerance and time horizon select investments
- Automatic portfolio rebalancing, tax-loss harvesting
- Translates into higher net returns for investors.
- FT (2016) estimates the market grows to $16.3 trillion worldwide

Figure 3: Quarterly returns of CRIX (blue), LSTM (green) and RNN (orange) portfolios

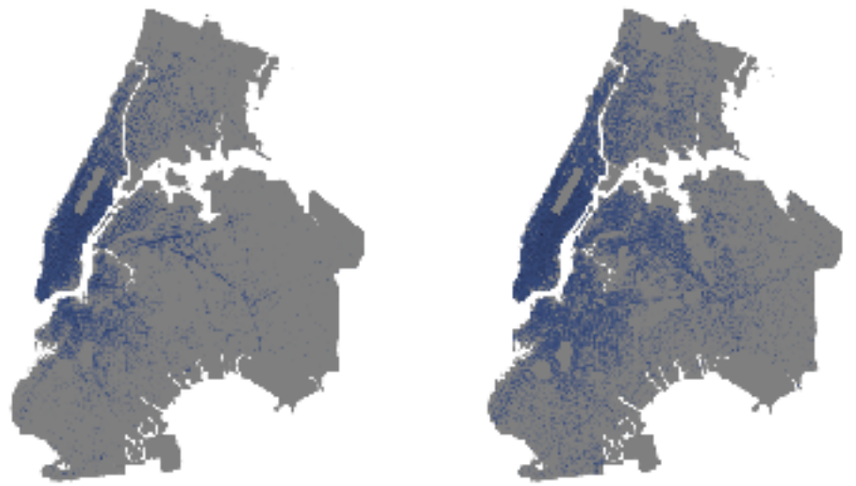


Predictive Policing



(a) Violent Crime (b) Property Crime

Figure 1: Number of crime incidents between June and November 2013. In the property crime map, the area around Penn station (largest outlier with 2002 incidents) is excluded for more consistent colour scaling.

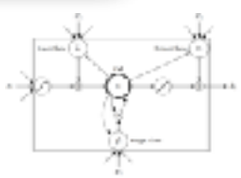


(a) Pickups (b) Dropoffs

Figure 2: Coordinates of completed taxi trips in New York City in week 43 in 2013.

- Crime Forecasting
- Social Media Data
- Spatial E'trics

Vomfell L, Härdle WK, Lessmann, S (2018) Improving Crime Count Forecasts Using Twitter and Taxi Data, Decision Support Systems, https://doi.org/10.1016/j.dss.2018.07.003

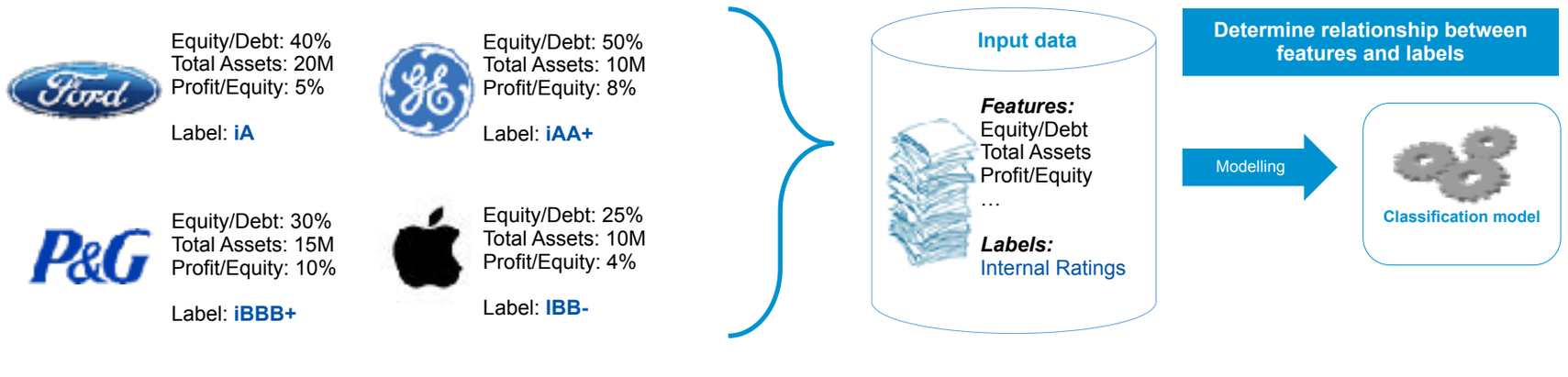


Machine Learning Based Credit Rating Methodology

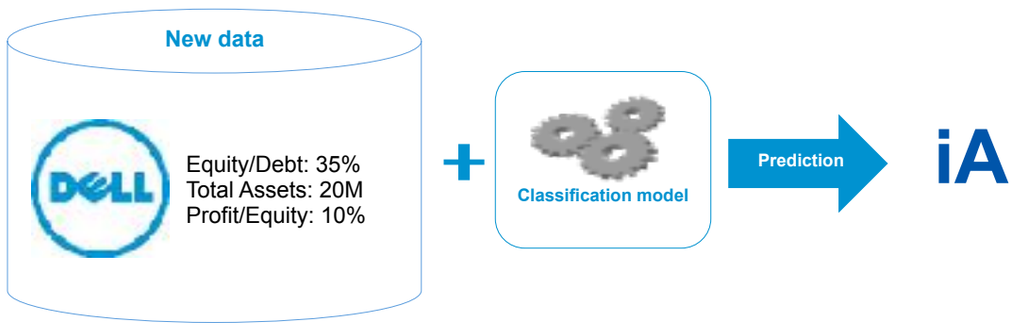
Corporate Credit Rating Prediction

… as a Supervised Learning Problem

Use labeled data points to train the model[1]



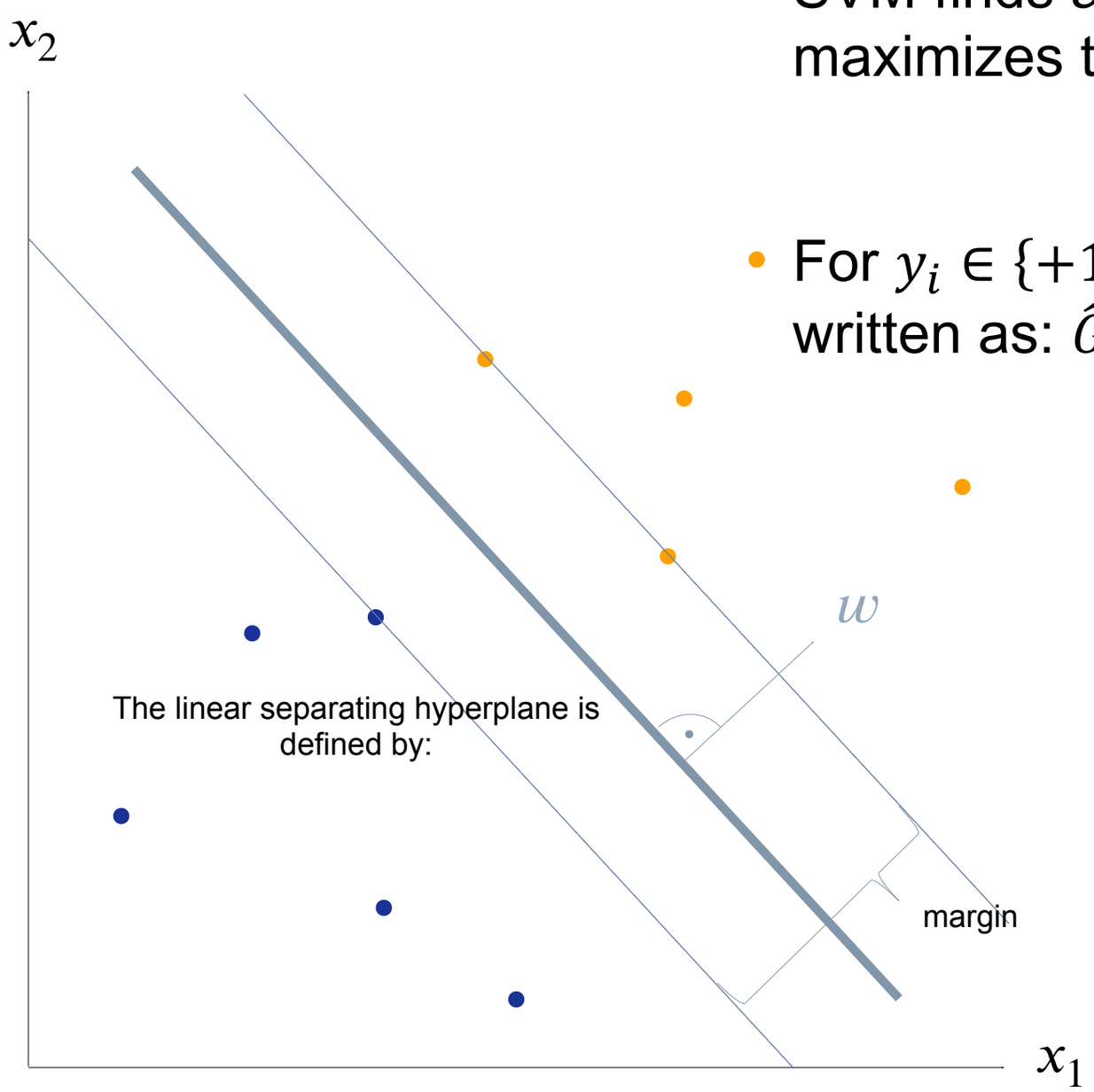
Use trained model to predict the label of new data points[1]



Internal DB Rating	Assigned PD in %
iAAA	> 0.00, ≤ 0.01
iAA+	> 0.01, ≤ 0.02
iAA	> 0.02, ≤ 0.03
iAA-	> 0.03, ≤ 0.04
iA+	> 0.04, ≤ 0.05
iA	> 0.05, ≤ 0.07
iA-	> 0.07, ≤ 0.11
iBBB+	> 0.11, ≤ 0.18
iBBB	> 0.18, ≤ 0.30
iBBB-	> 0.30, ≤ 0.50
iBB+	> 0.50, ≤ 0.83
iBB	> 0.83, ≤ 1.37
iBB-	> 1.37, ≤ 2.27
iB+	> 2.27, ≤ 3.75
iB	> 3.75, ≤ 6.19
iB-	> 6.19, ≤ 10.22
iCCC+	> 10.22, ≤ 16.87
iCCC	> 16.87, ≤ 27.84
iCCC-	> 27.84, ≤ 99.99

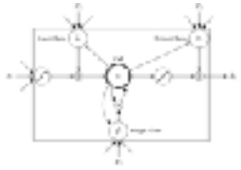
[1]Depicted company characteristics and internal ratings were set arbitrarily and do not reflect real data.

Support Vector Machine

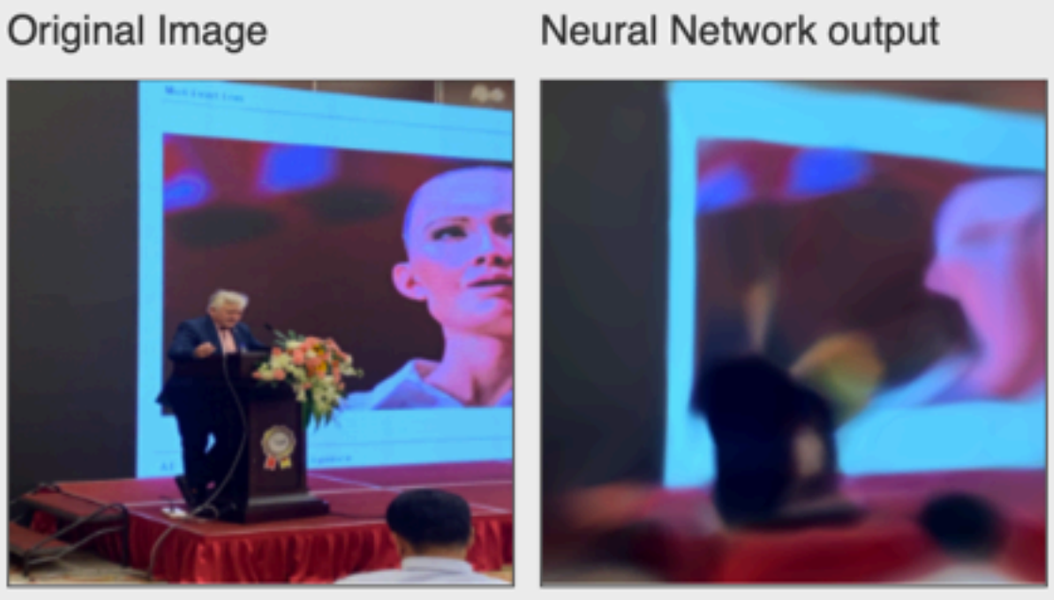


SVM finds a separating hyperplane that maximizes the margin between the two classes

- For $y_i \in \{+1, -1\}$, the decision function may be written as: $\hat{G}(x) = \text{sgn}(w^T x - b)$



AI draws pictures

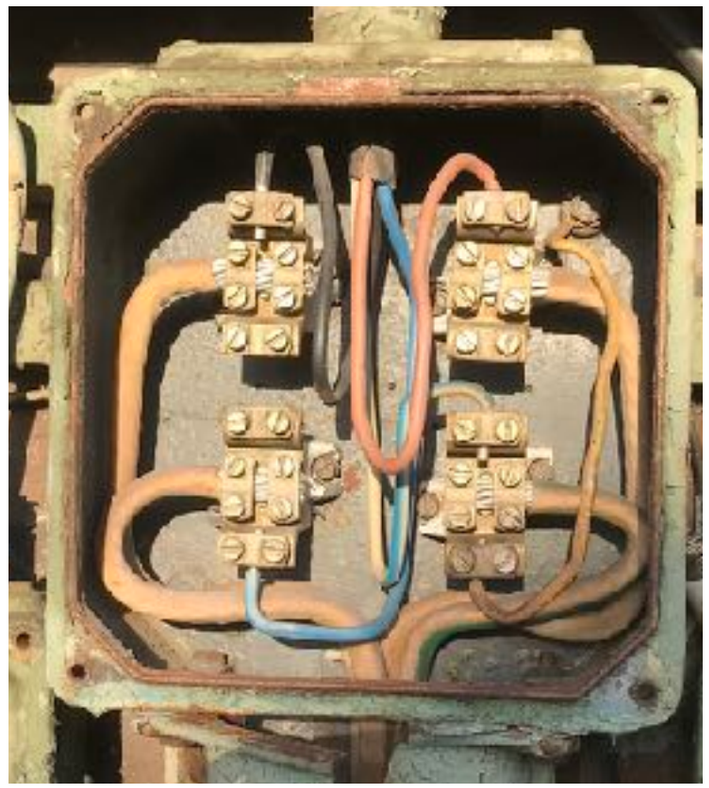


loss: 0.012254625661909253
iteration: 2957

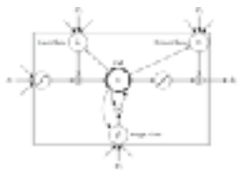
https://cs.stanford.edu/people/karpathy/convnetjs/

Basic element of Artificial Intelligence

…is a buzzer or a switch!

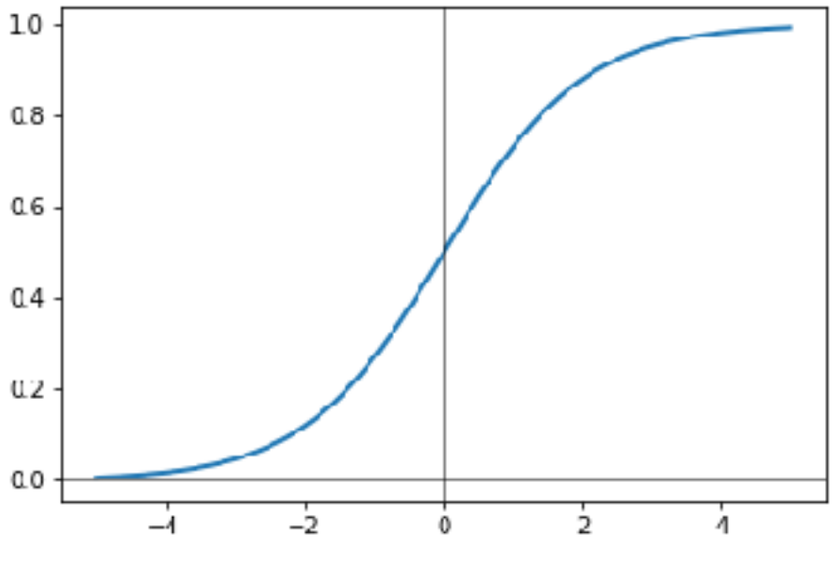


Combined with other switches …



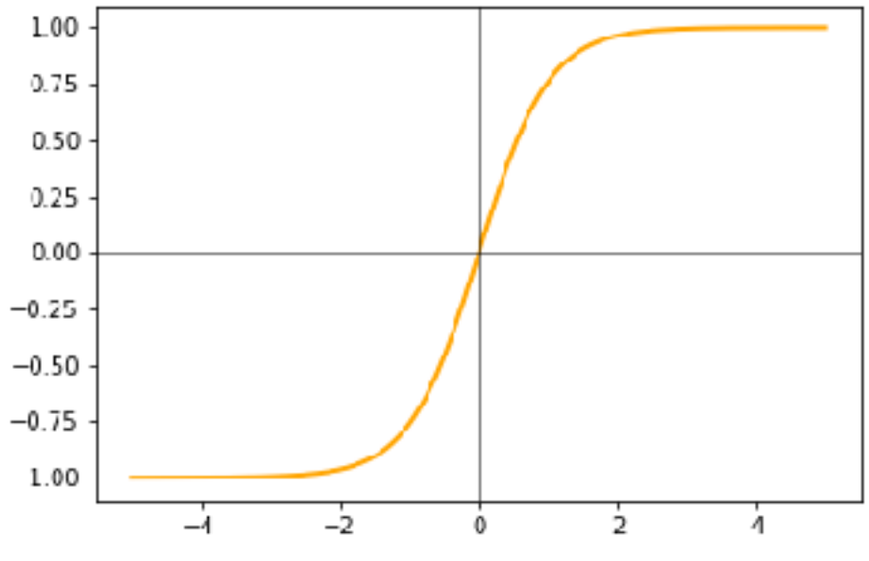
LSTM memory block: activation functions

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$



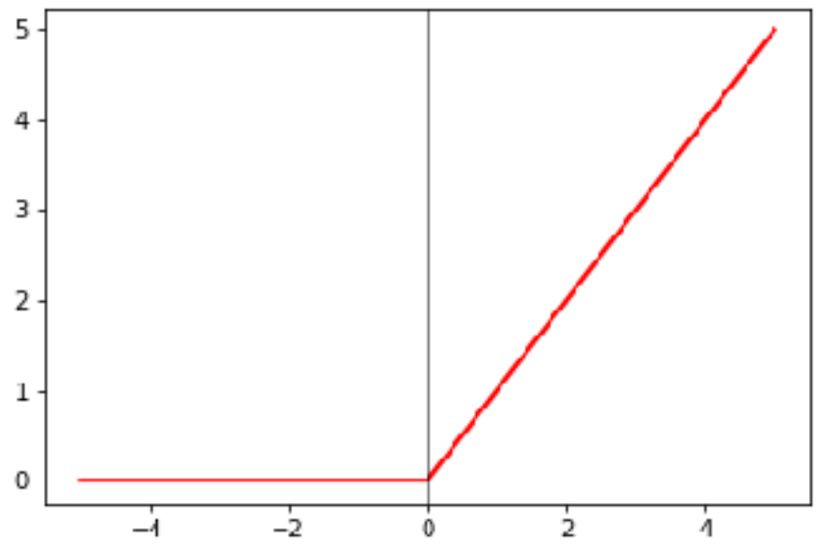
For forget gates

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$

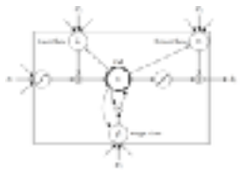


for input and output gates

Or

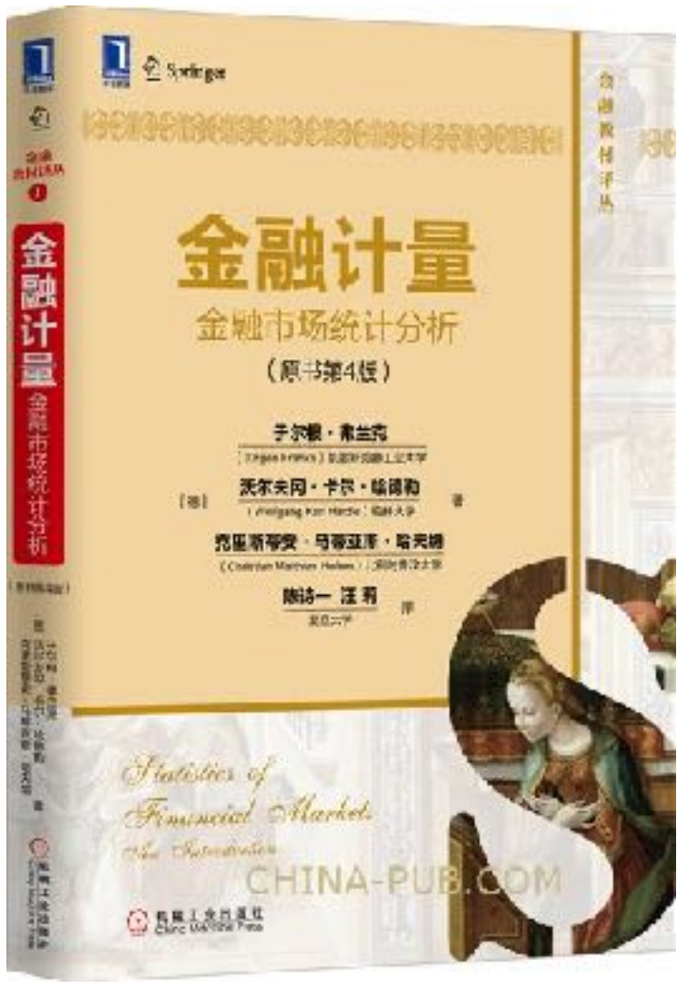


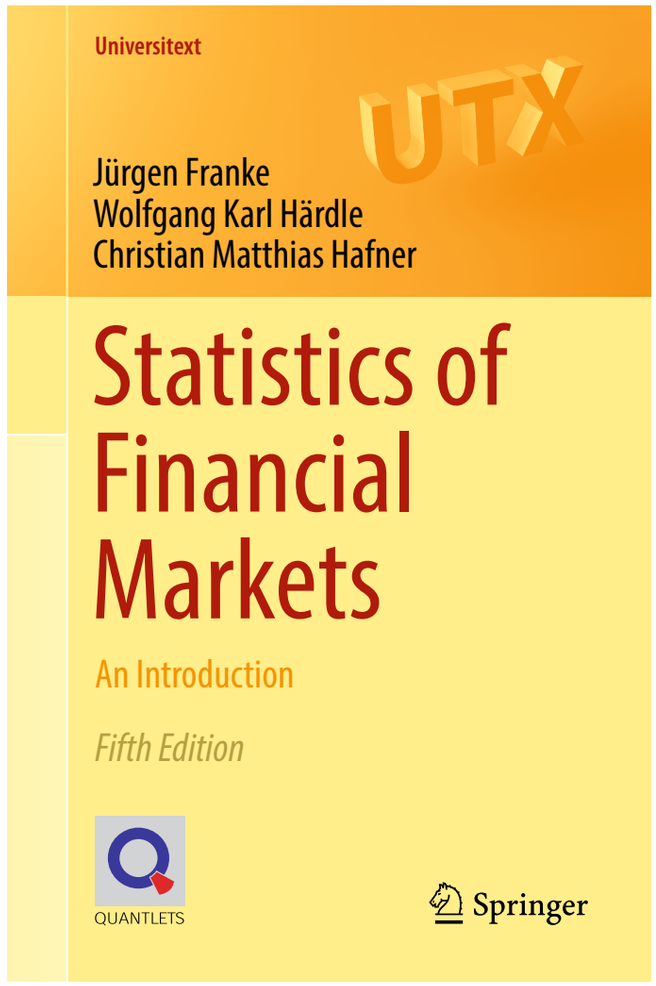
$$ReLU(x) = \max(0,x)$$

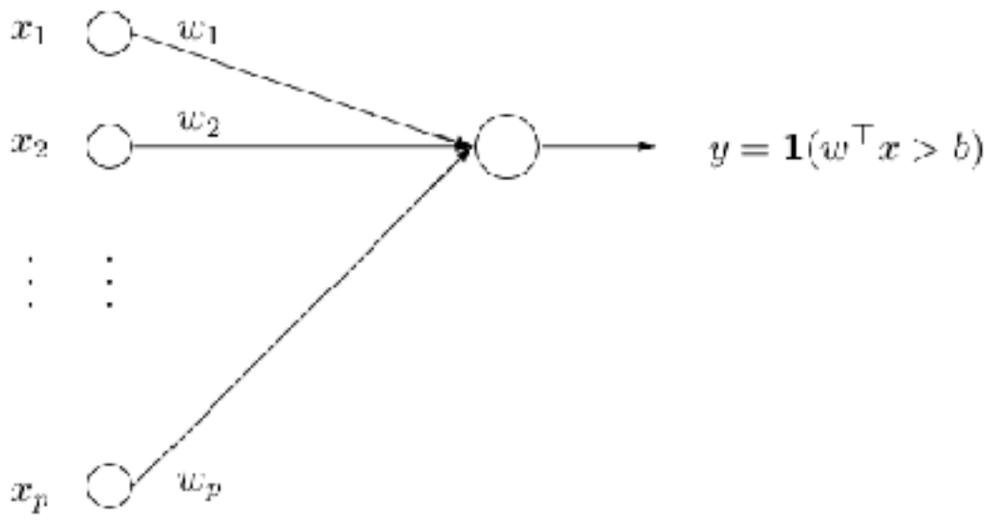


The perceptron

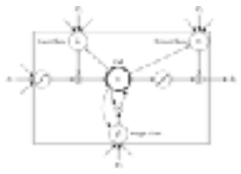
- Mimic human biological neutrons
- Take some variables as input
- Apply linear threshold operations
- The output is the perceptron
- Cannot represent XOR function



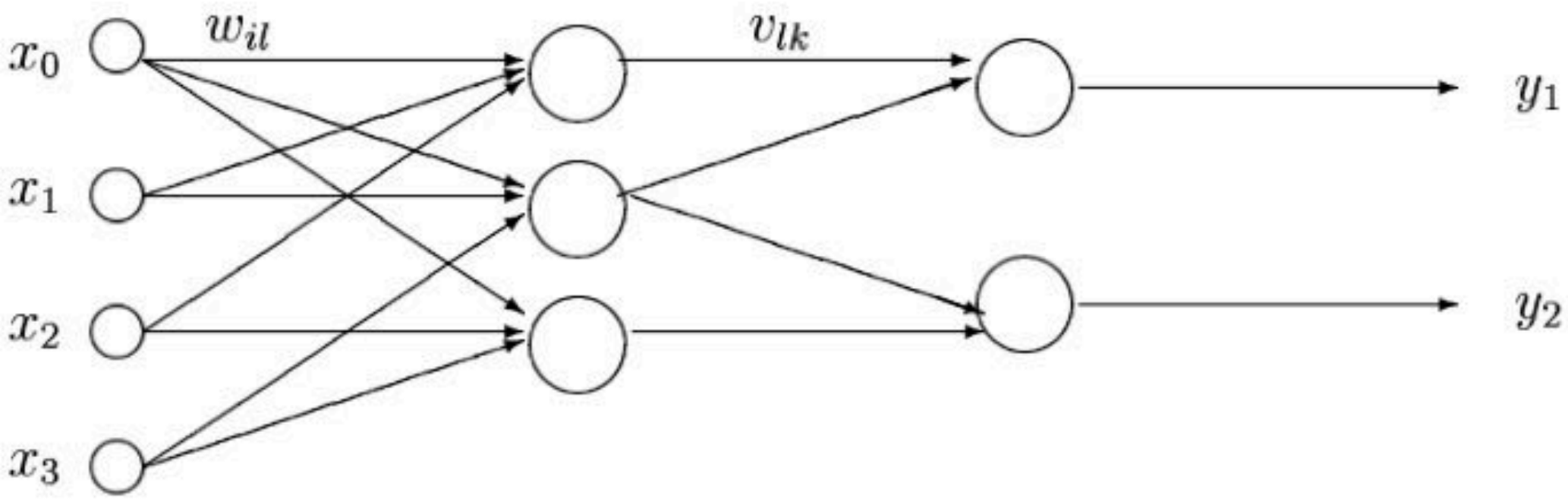




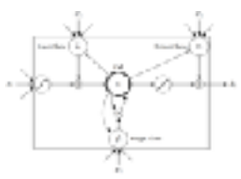
$$where \quad x = (x_1, \ldots, x_p),\ p \in \mathbb{N}^* \quad and \quad w = (w_1, \ldots, w_p),\ p \in \mathbb{N}^*$$



Multilayer perceptron



- One input layer, multiple hidden layers and one output layer
- Non-linear activation function (*sigmoid*, *tanh*, *relu*)
- Feedforward flow of information: no cyclic connection

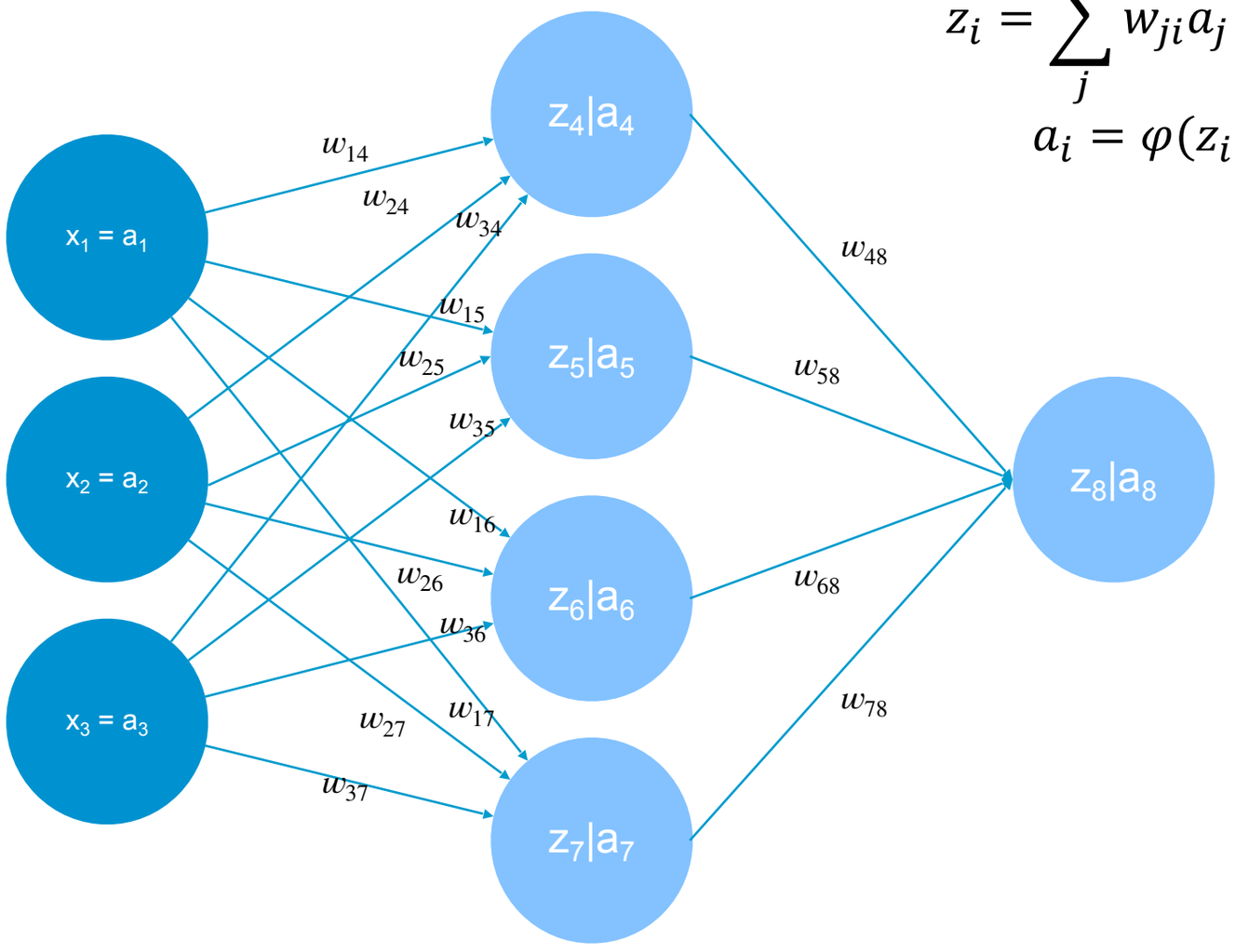


Neural Network

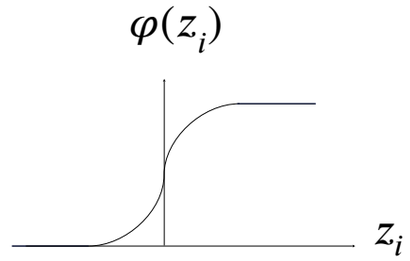
Example: 2-Layer Feed-Forward NN for Binary Classification

$$z_i = \sum_j w_{ji} a_j + b_i$$

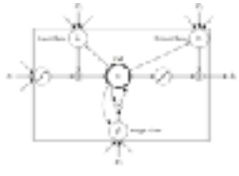
$$a_i = \varphi(z_i)$$



Activation function : e.g. sigmoid



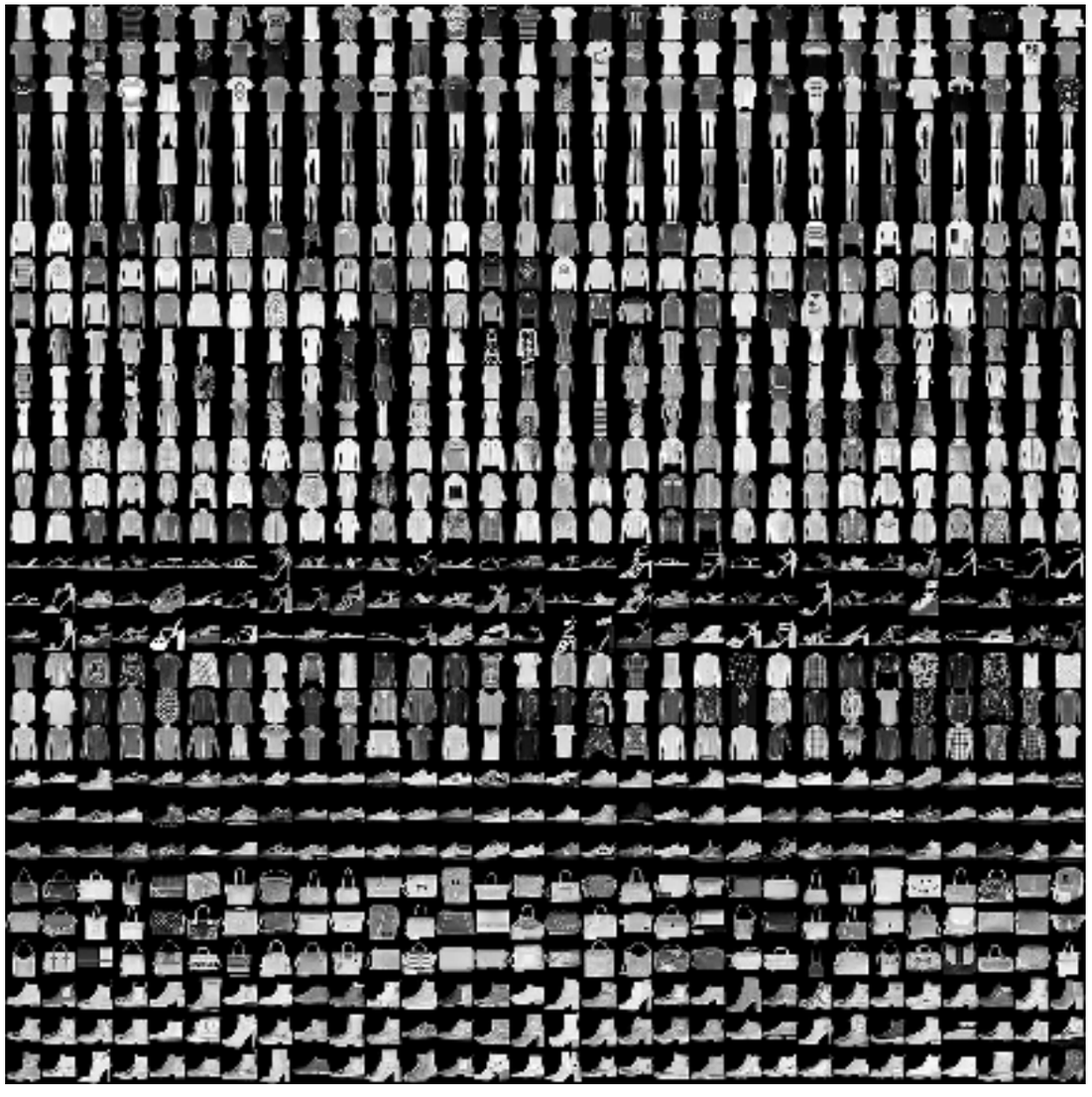
Internal DB Rating	Assigned PD in %
iAAA	> 0.00, ≤ 0.01
iAA+	> 0.01, ≤ 0.02
iAA	> 0.02, ≤ 0.03
iAA-	> 0.03, ≤ 0.04
iA+	> 0.04, ≤ 0.05
iA	> 0.05, ≤ 0.07
iA-	> 0.07, ≤ 0.11
iBBB+	> 0.11, ≤ 0.18
iBBB	> 0.18, ≤ 0.30
iBBB-	> 0.30, ≤ 0.50
iBB+	> 0.50, ≤ 0.83
iBB	> 0.83, ≤ 1.37
iBB-	> 1.37, ≤ 2.27
iB+	> 2.27, ≤ 3.75
iB	> 3.75, ≤ 6.19
iB-	> 6.19, ≤ 10.22
iCCC+	> 10.22, ≤ 16.87
iCCC	> 16.87, ≤ 27.84
iCCC-	> 27.84, ≤ 99.99



Implementation: one classical example

- MNIST fashion dataset: 60000 training samples, 10000 test samples
- Input: 28x28 grayscale images
- Output: 10 classes (T-shirt, trouser, pullover, dress, coat, sandal, shirt, sneaker, bag, ankle boot)

SIRI + „Schuhe“= Viola Härdle, 16 Y

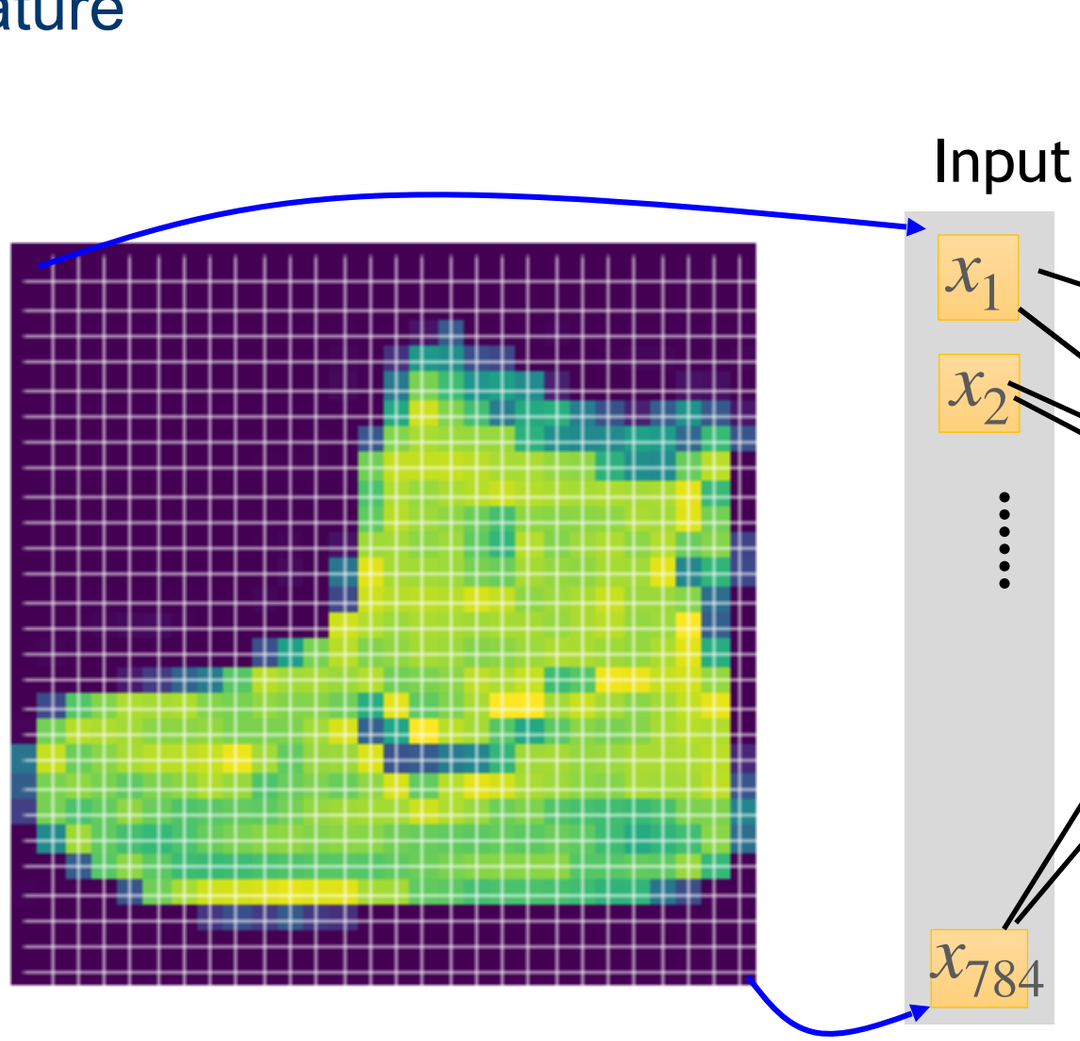


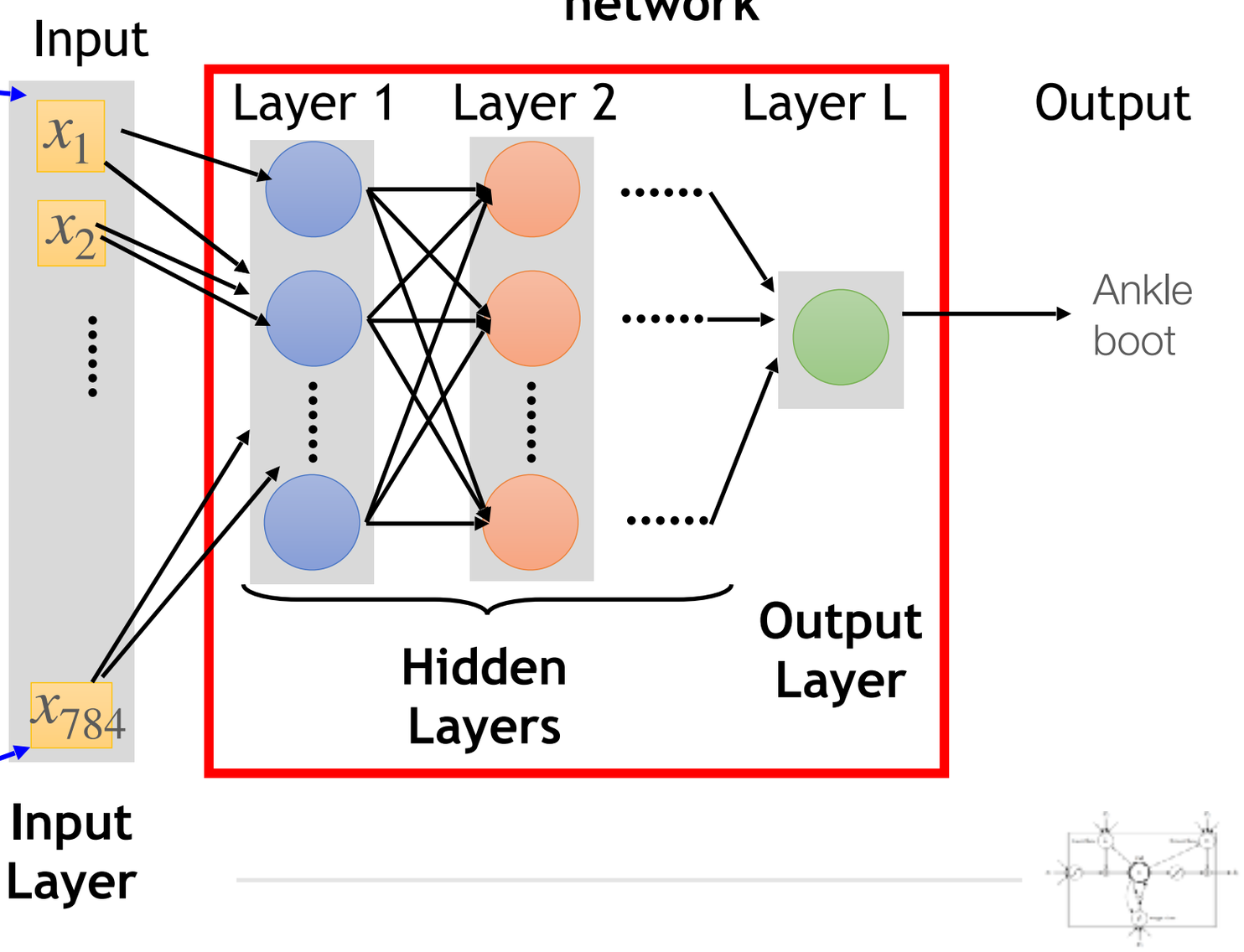
Implementation: dense layer

Input transformation into a vector of length 28x28 = 784

$$\begin{pmatrix} x_{1,1} & \cdots & x_{1,784} \\ & \cdots & \\ x_{n,1} & \cdots & x_{n,784} \end{pmatrix}$$

Each new pixel is treated as an independent feature





Implementation: sequence learning

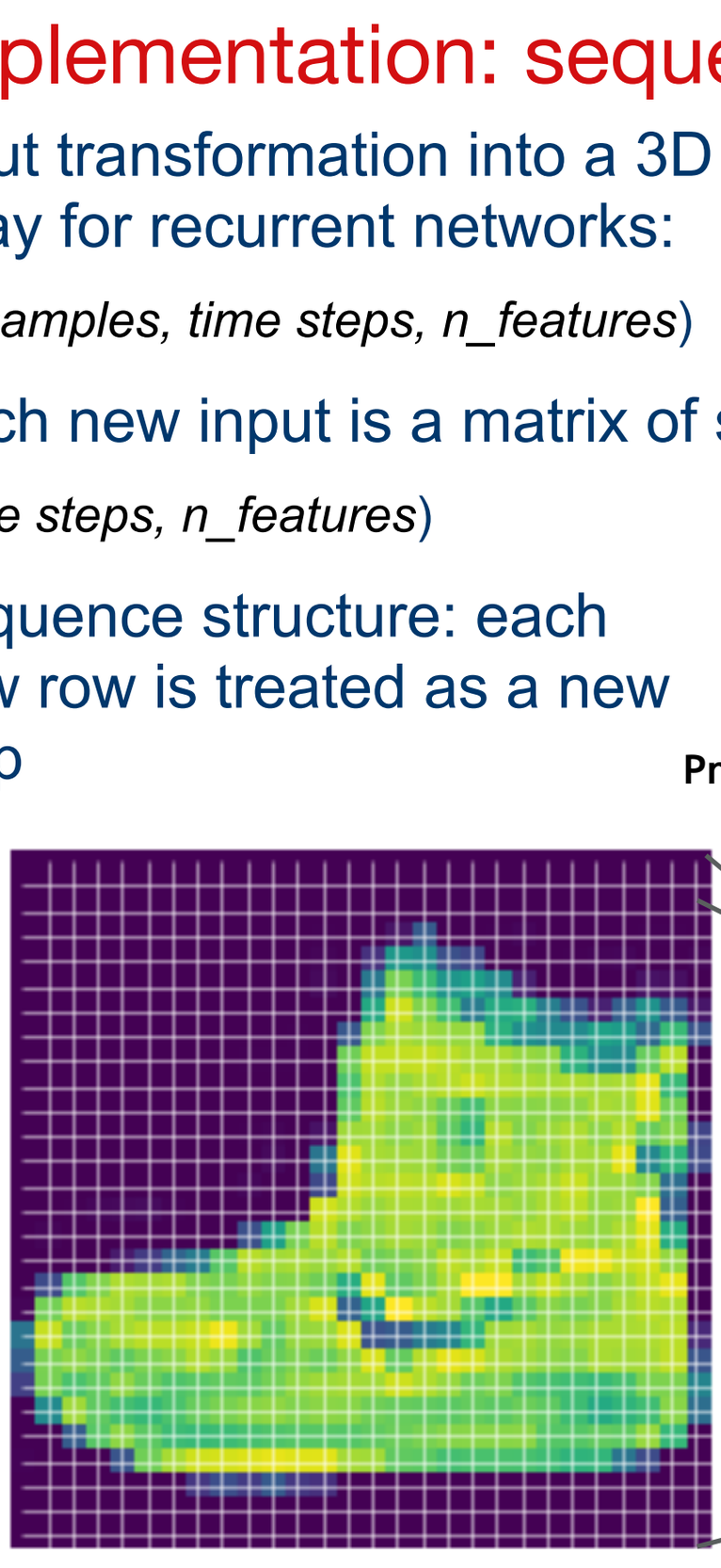
Input transformation into a 3D array for recurrent networks:

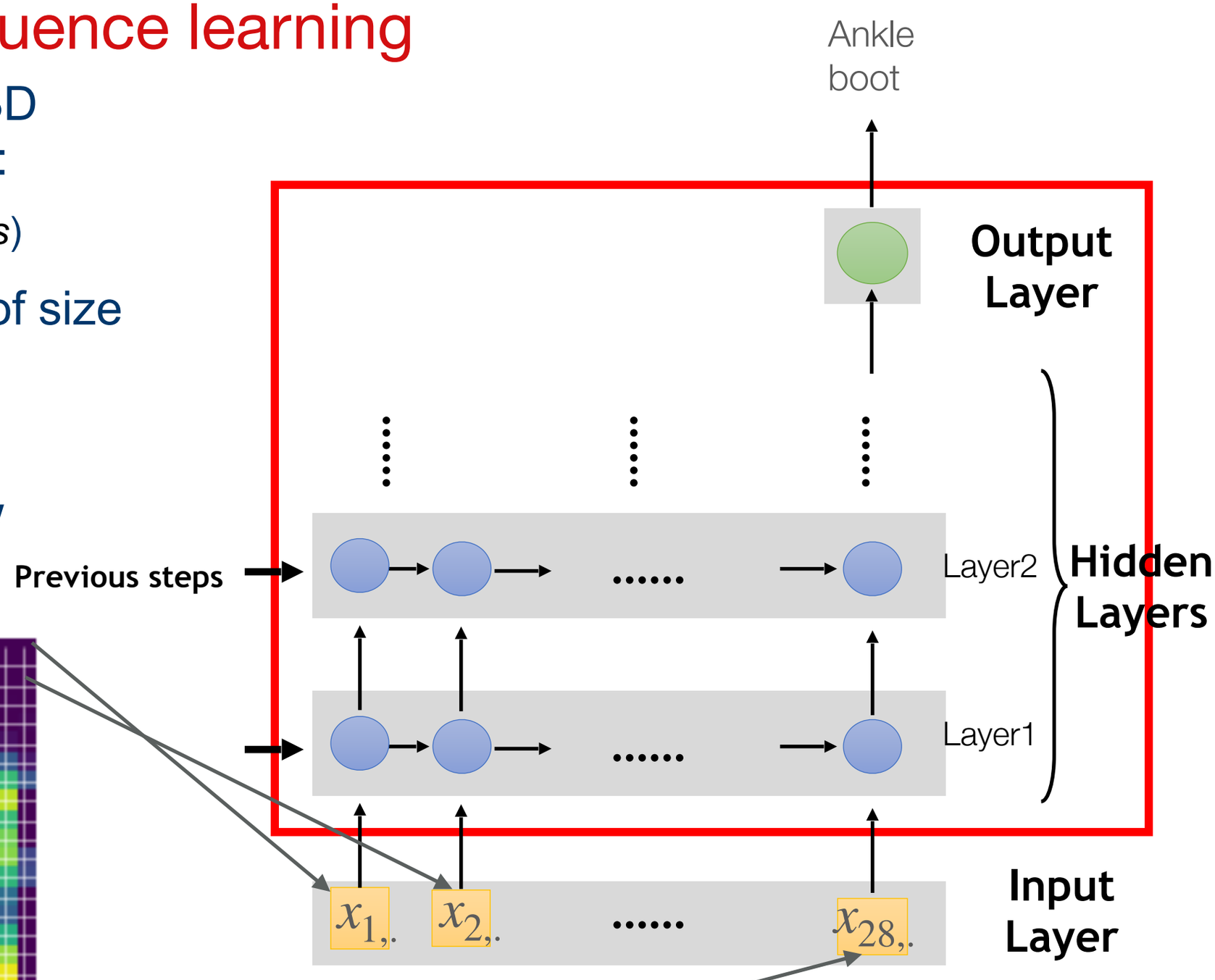
(*n_samples, time steps, n_features*)

Each new input is a matrix of size

(*time steps, n_features*)

Sequence structure: each new row is treated as a new step





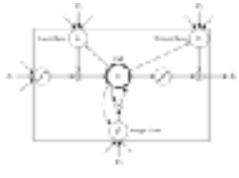
Rise of the machines





- 1981 - $300,000
- 1987 - $50,000
- 1990 - $10,000
- 1994 - $1,000
- 1997 - $100
- 2000 - $10
- 2004 - $1
- 2010 - $0.10
- 2017 - $0.01
- 2018 - $0.004*

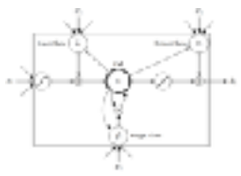
Cost per GB storage p/m



Blockchain

- Computers
- Digital Economy
- New Models.DE

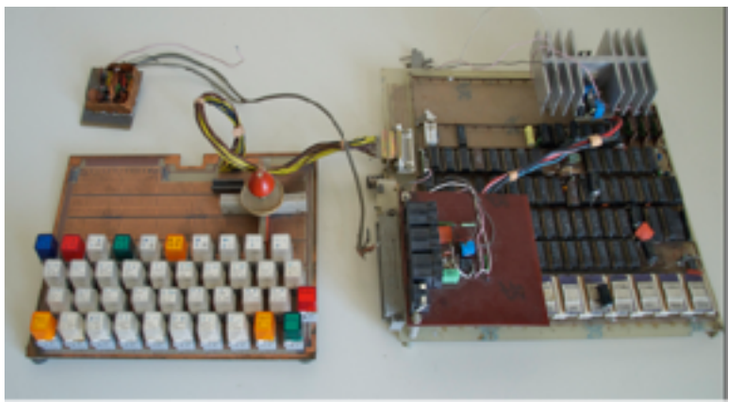
```
> library(digest)
> digest("blockchain", "sha256")
[1] "b960c3a766e37e8eaa76574f7bfd031b5d1fa63bfba8dc5004fa4d0a9eccbee1"
> digest("digital economy", "sha256")
[1] "7204ecd0b6c2d9f463822e081f54d9773ff53fbdc1752841f5d68d7e2a301127"
> digest("newmodels.DE", "sha256")
[1] "ae46ede37d5c0148f0327e67015ee63b797eb1d84258bff636ff4347793c62f9"
>
```



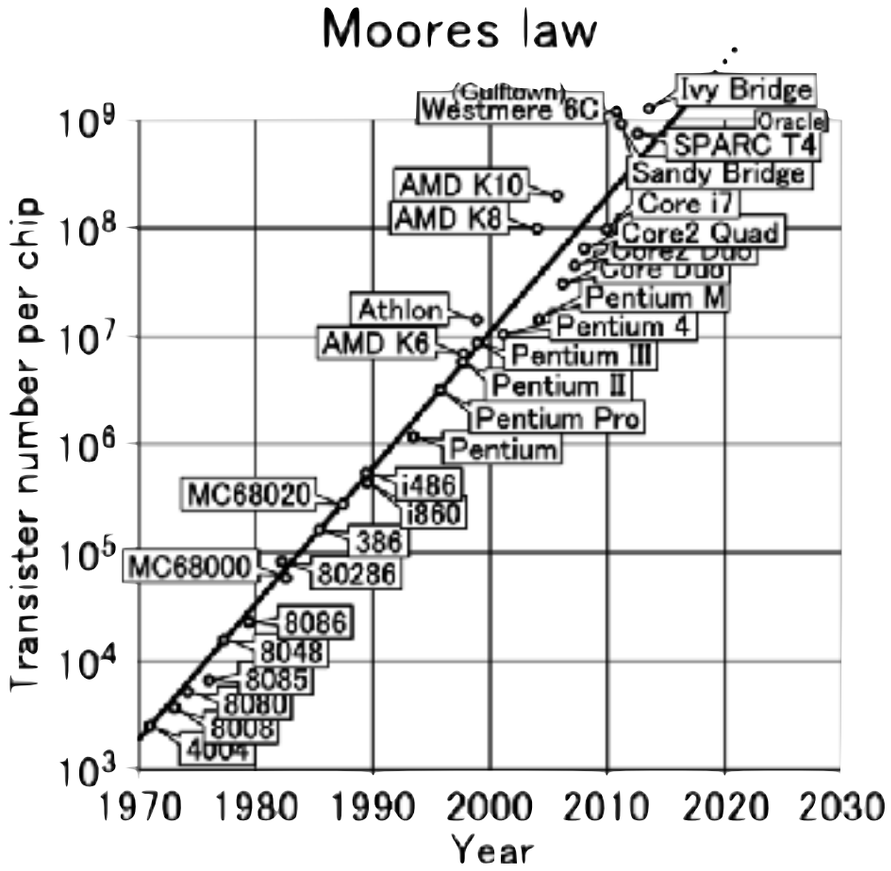
Rise of the machines

Mechanical calculator MADAS
H.W. Egli S.A., Zurich, 1965

VEB Röhrenwerk Mühlhausen, 1975
8 digits

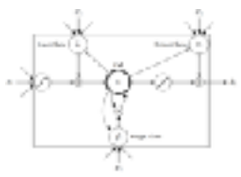


ZX Spectrum East / West 1982



https://phys.org/news/2015-08-silicon-limits-power-electronics-revolution.html

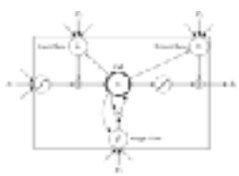
- http://computermuseum.wiwi.hu-berlin.de



Rise of the machines

- Cray 2 is the world's fastest supercomputer: 1985-1990
- 1.9 GFLOPs*
- 5,500 pounds
- $32 million (current $)
- 舞 锡市 太湖之光, (wu3xi1shi4tai4hu2zhi1guang1) 95 PFLOPS**

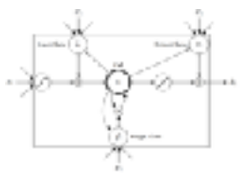
*10^9 FLoating point OPerations per Second **10^15 FLOPS



Rise of the machines

- 2016 iPhone 7* , 178 GFLOPs
- 2017 iPhone X* , 600 GFLOPs
- A11 Bionic chip, 256 GB storage
- 174 g
- 1.3 K EUR

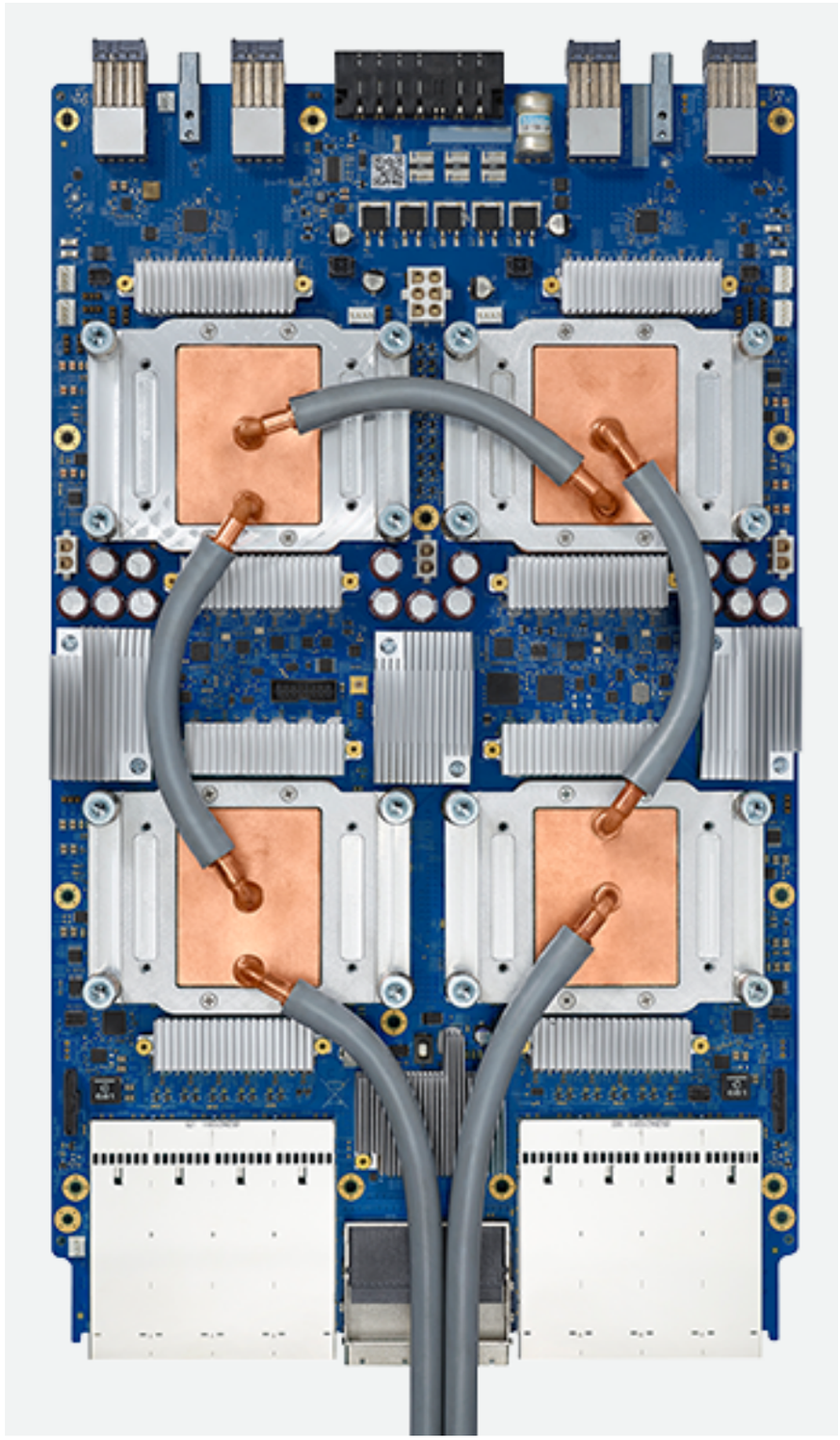
*A10 Fusion. The Apollo guidance system had only 4K of RAM.



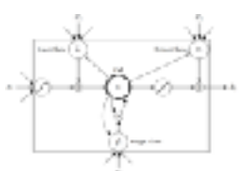
Rise of the machines

- 2017 Nvidia Titan Xp
- 125 TFLOPs, 16 GB
- 1.1 K USD
- ASICS, Antminer

- Google’s Tensor Processing Unit (TPUv2)
- 180 TFLOPs, 64GB per TPU
- TPUs important for MLE*

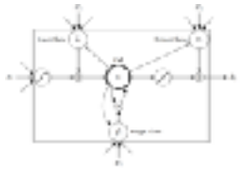


*MLE = „Maximum Likelihood Estimation“ (age>=45)
MLE = „Machine Learning in Economics“ (age<=45)



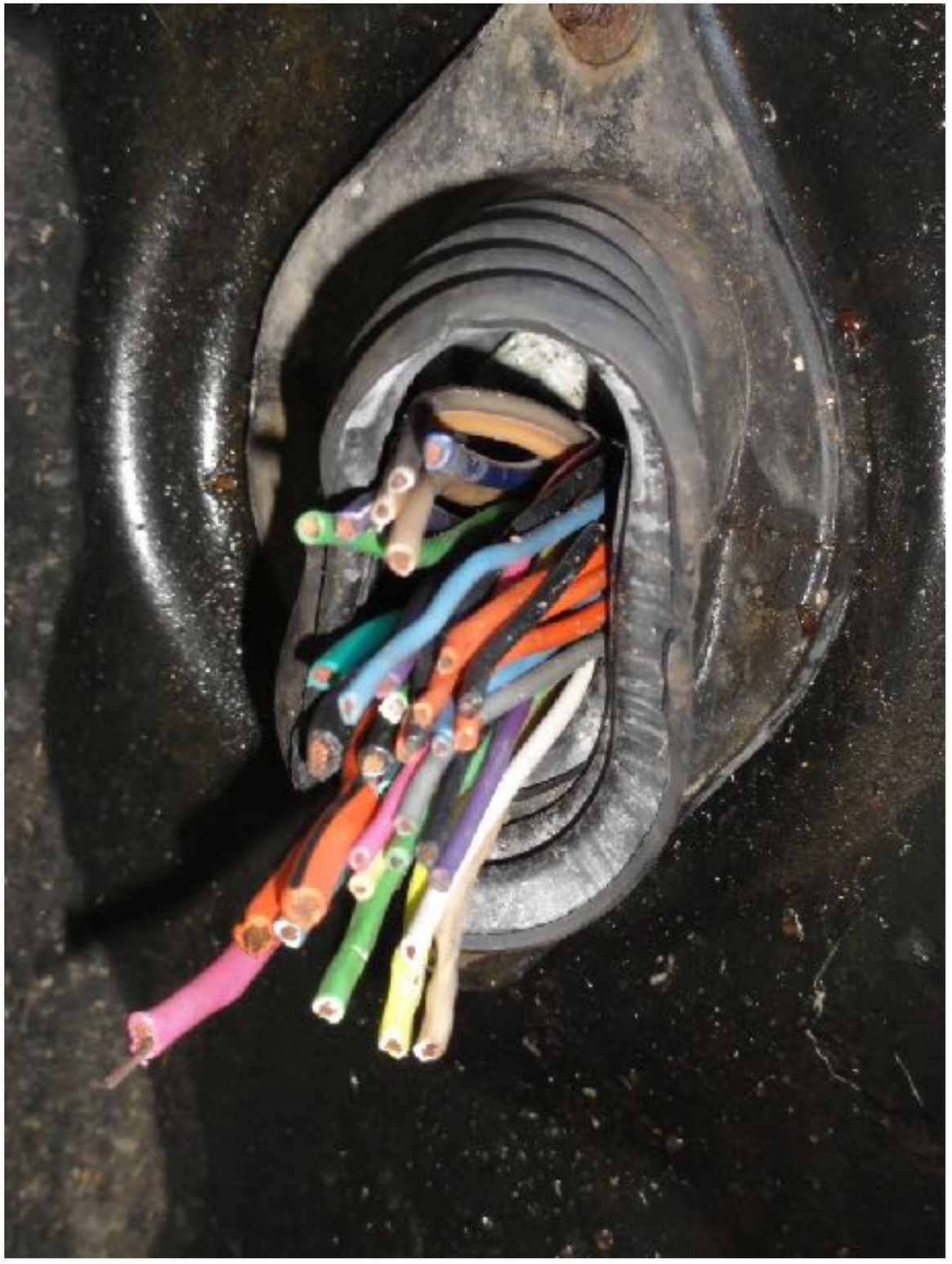
Infrastructure

- badly maintained switch
- corrupted node sequence
- biased training sample



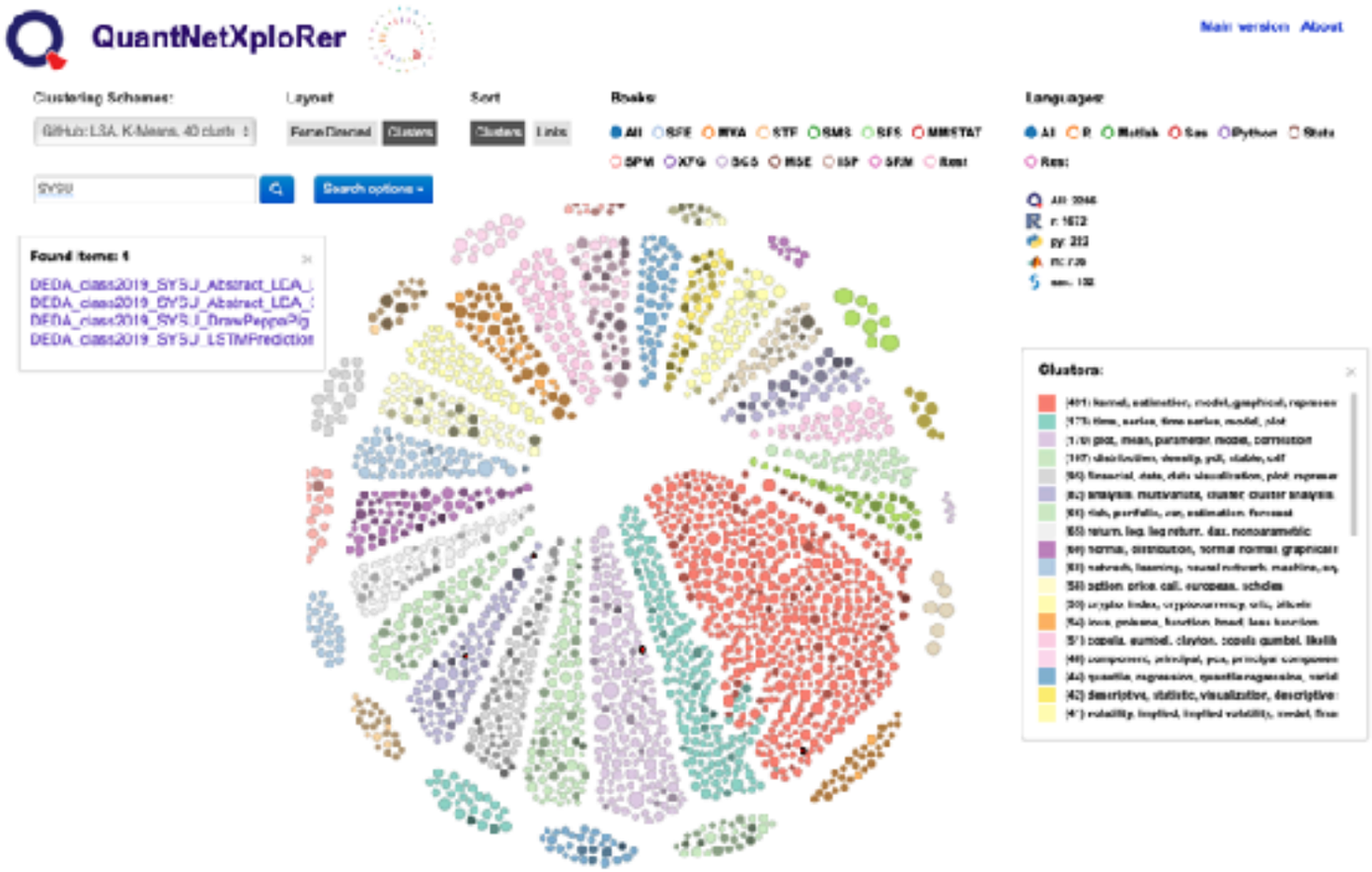
Maintenance

- do not overweigh your switch
- keep the electricity running
- invest in human brains



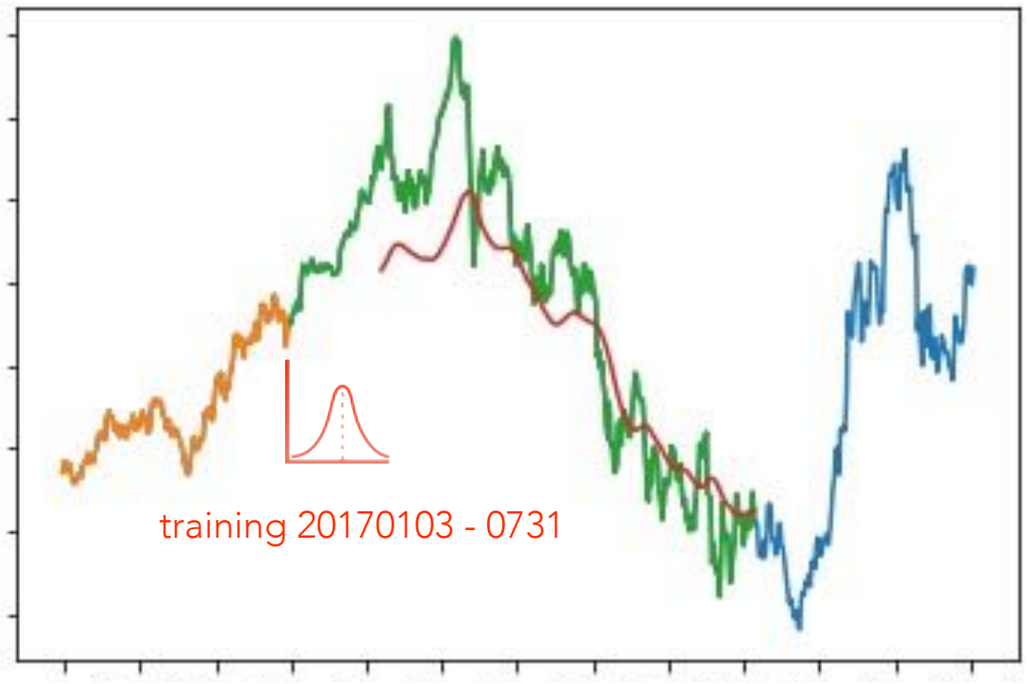
ML on Time Series

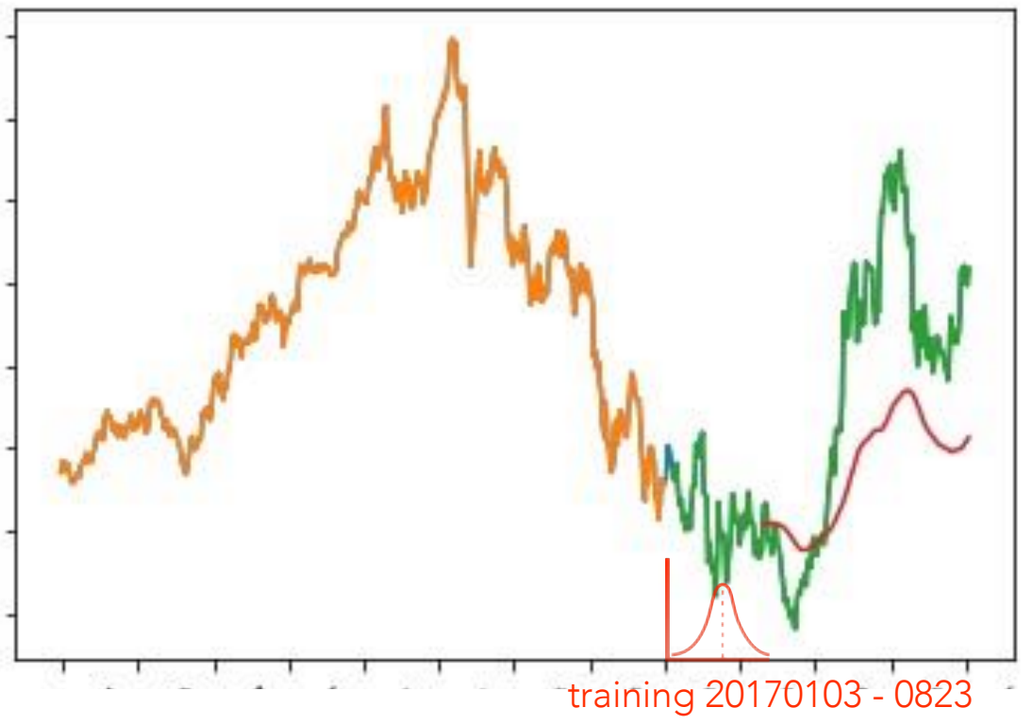
- quantlet.de
- *SYSU, LSTM prediction*
- train on different time periods

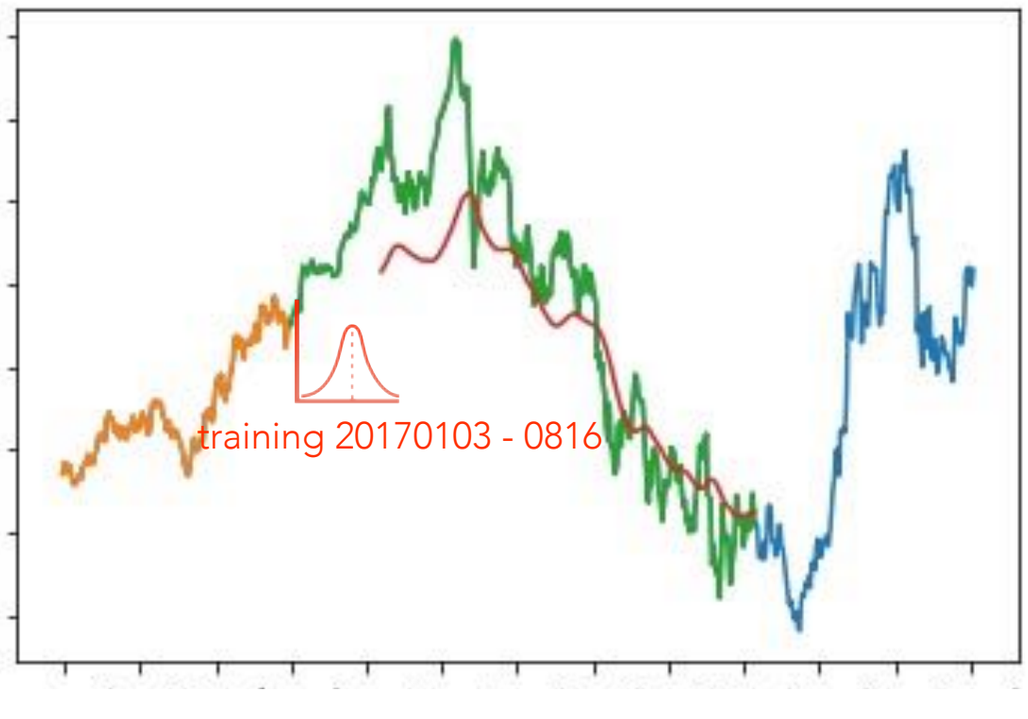


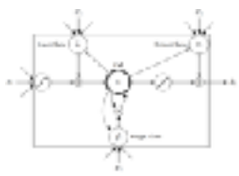
ML on Time Series

- quantlet.de
- *SVSU, LSTM prediction*
- train on different time periods





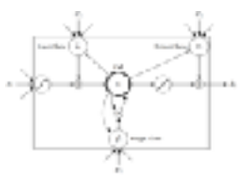




Will AI classify this correctly?

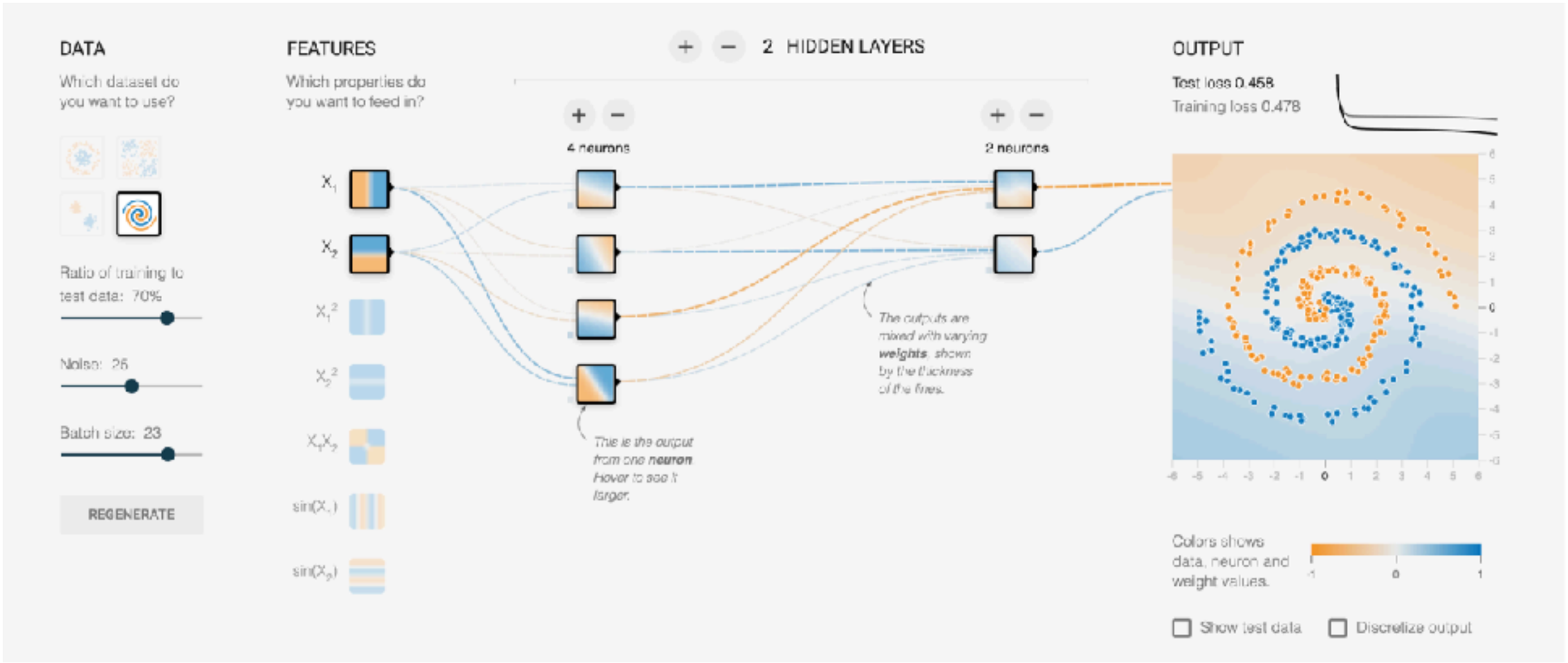
Danish FSA muffin

Training machines to "see" - to recognise and differentiate between objects and faces - is notoriously difficult.



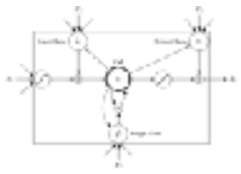
The difficulty of getting it right

Neural Network Tensorflow Playground



The 7 AI principles

- AI has to be useful for the society
- AI must be safe
- AI has to avoid bias
- AI should be made castigatable
- AI must respect individual privacy
- AI has to follow highest scientific standards
- AI should only be employed for applications that fulfill the above principles



Future

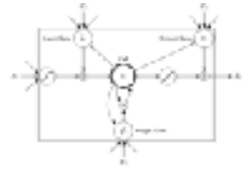
Should we wait: NO, go for it!

Will you take a standard air plane on long distance that is flown by an AI robot pilot?

You do it already.

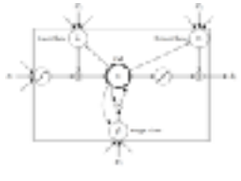
But…

Will you take a standard air plane on long distance that is flown by an AI robot pilot, knowing that there is no additional human pilot?



Questions

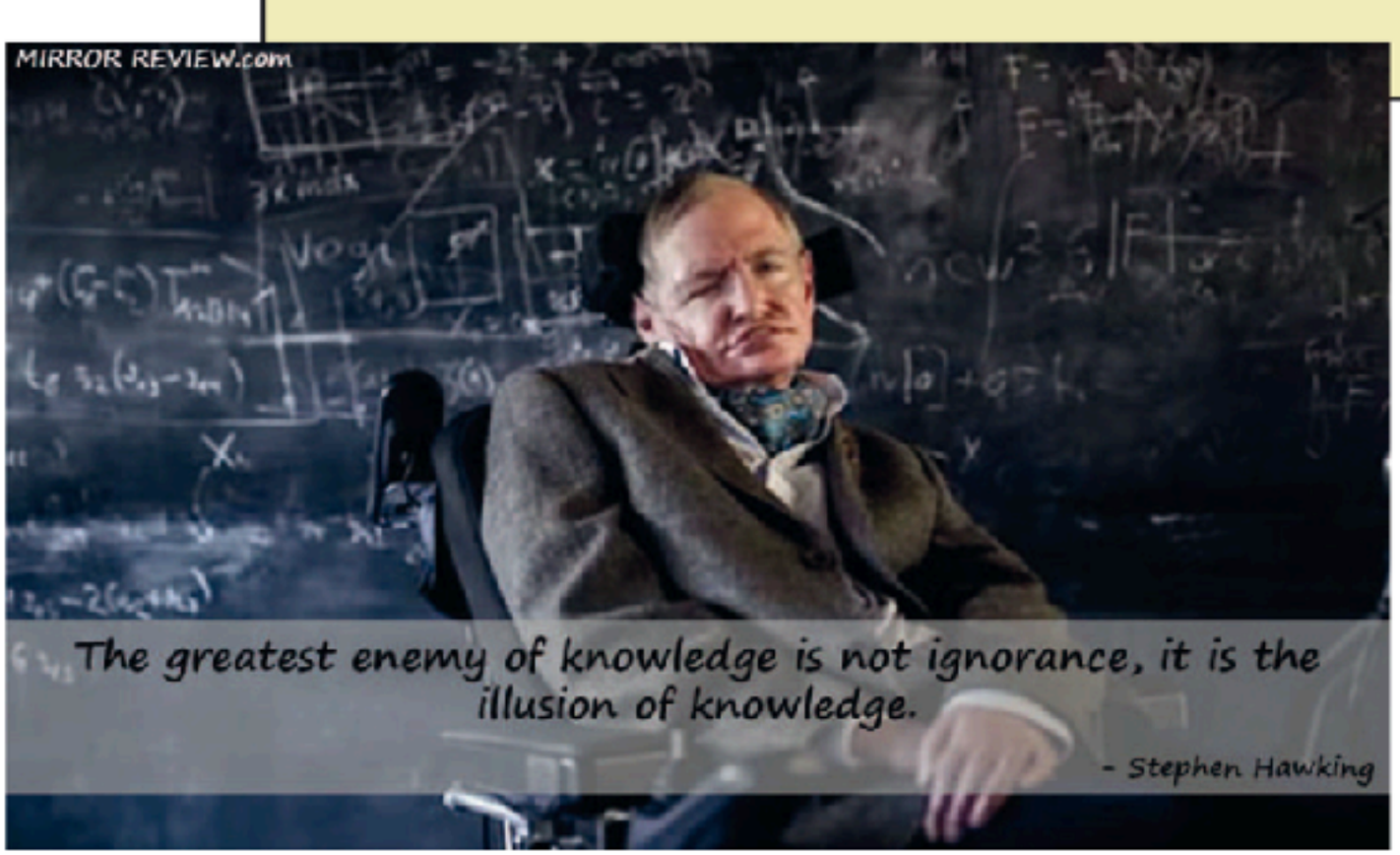
- How do we make AI ethical ?
- How do we react to AI „colleagues“ ?
- Can an AI robot be happy ?
- What must politics and business do ?
- How long is human intelligence superior to AI ?



AI Artificial Intelligence +

„Ali Rahimi, a researcher in artificial intelligence (AI) at Google in San Francisco, California, took a swipe at his field last December—and received a 40-second ovation for it. Speaking at an AI conference, Rahimi charged that machine learning algorithms, in which computers learn through trial and error, have become a form of „alchemy." Researchers, he said, do not know why some algorithms work and others don't, nor do they have rigorous criteria for choosing one AI architecture over another...."

Science, May 2018



AI Artificial Intelligence +

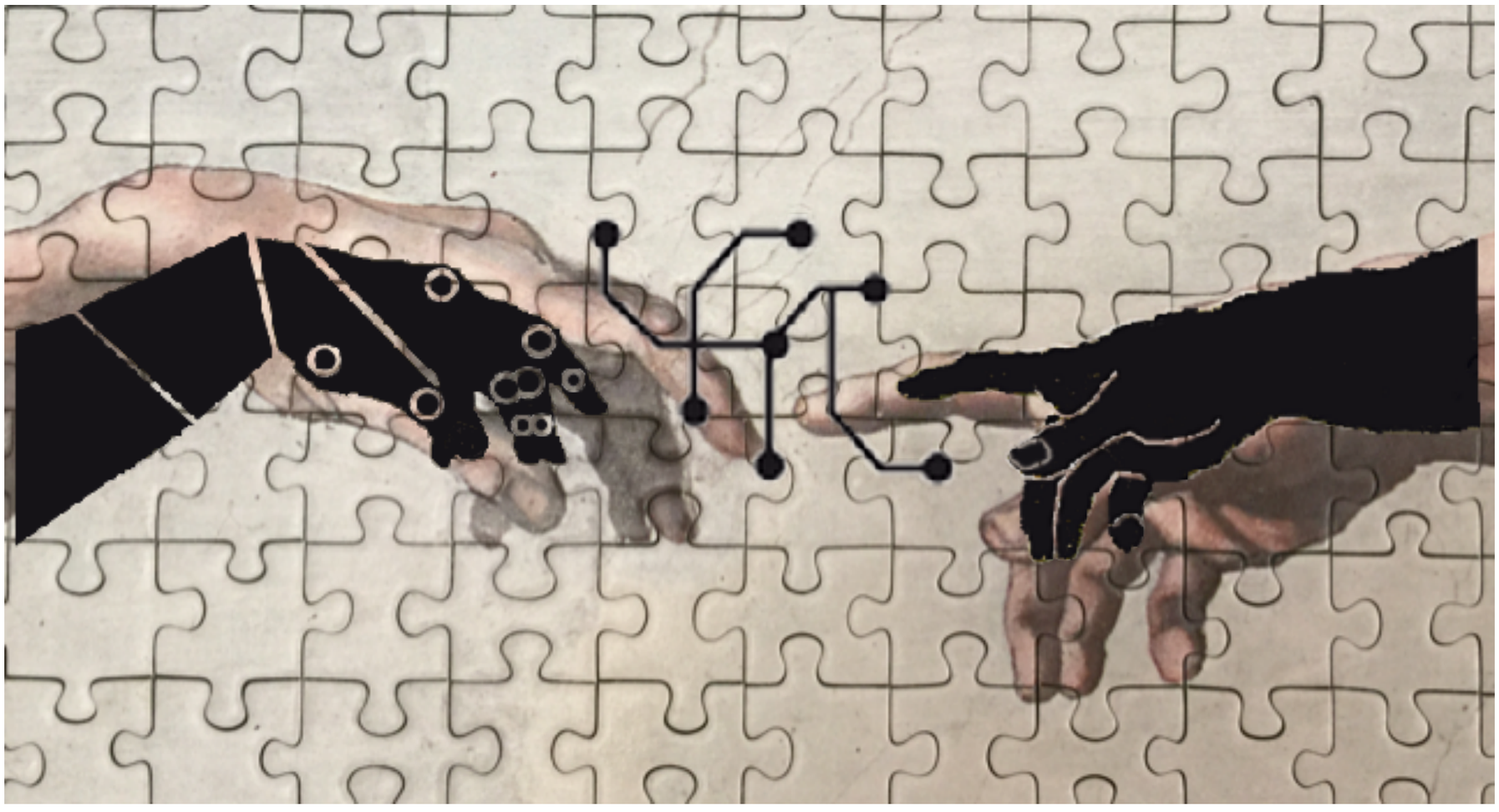
SMART (Specific, Measurable, Achievable, Relevant, Timely) Data Analytics

Wolfgang Karl Härdle hu.berlin/wkh

switches, neurons, networks, nodes,
deep-, machine-, supervised learning, computer, LSTM,
Python, statistics, data mining, analytics, smart data,
cluster analysis, crypto currencies, robo advising, FinTech

AI Artificial Intelligence +

The future of AI?



AI Artificial Intelligence +

Event Stream

- 20181020 HKG Banker's Club WALEX Diamond Class event
- 20190919 HBF Jinan QiLu conference
- 20190921 HBF Beijing
- 20200207 Danish FSA, Copenhagen
- 20200210 Bundesbank Frankfurt