

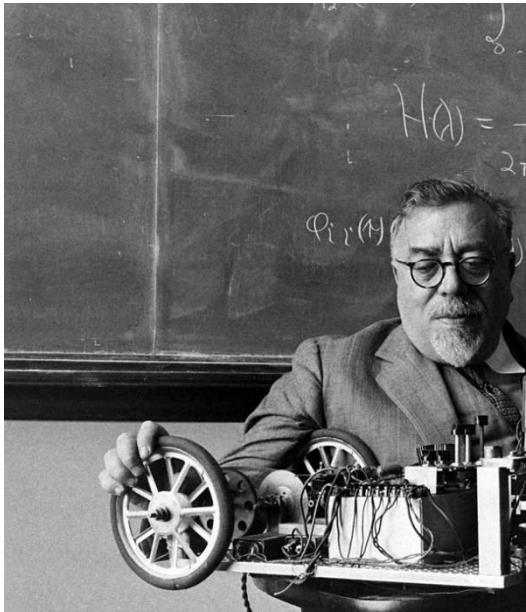
Machine Learning

Anton Selitskiy

Outline

- History and definition
- Uses:
 - Image recognition
 - Voice recognition
 - Video processing

Cybernetics



Norbert
Wiener

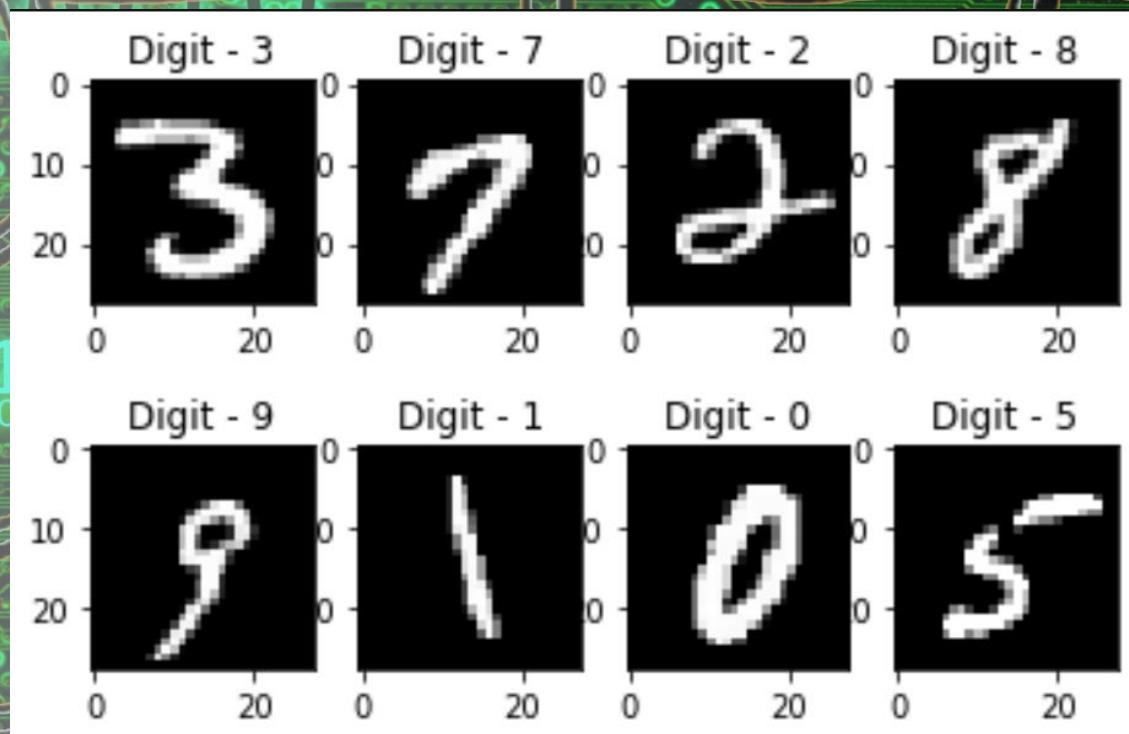


MIT
radio-set



V-weapon

Artificial Intelligence (AI)



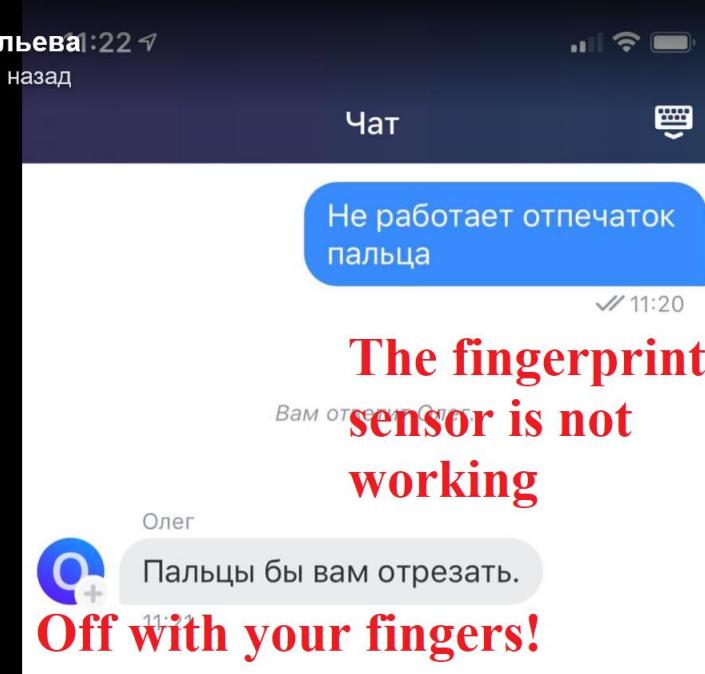
After 1990s
Machine Learning (ML)

Voice recognition and interactive response systems

OFF WITH
THEIR HEAD!



Анна Васильева 11:22 ↗
около недели назад



Video Processing

- Obama's fake speech



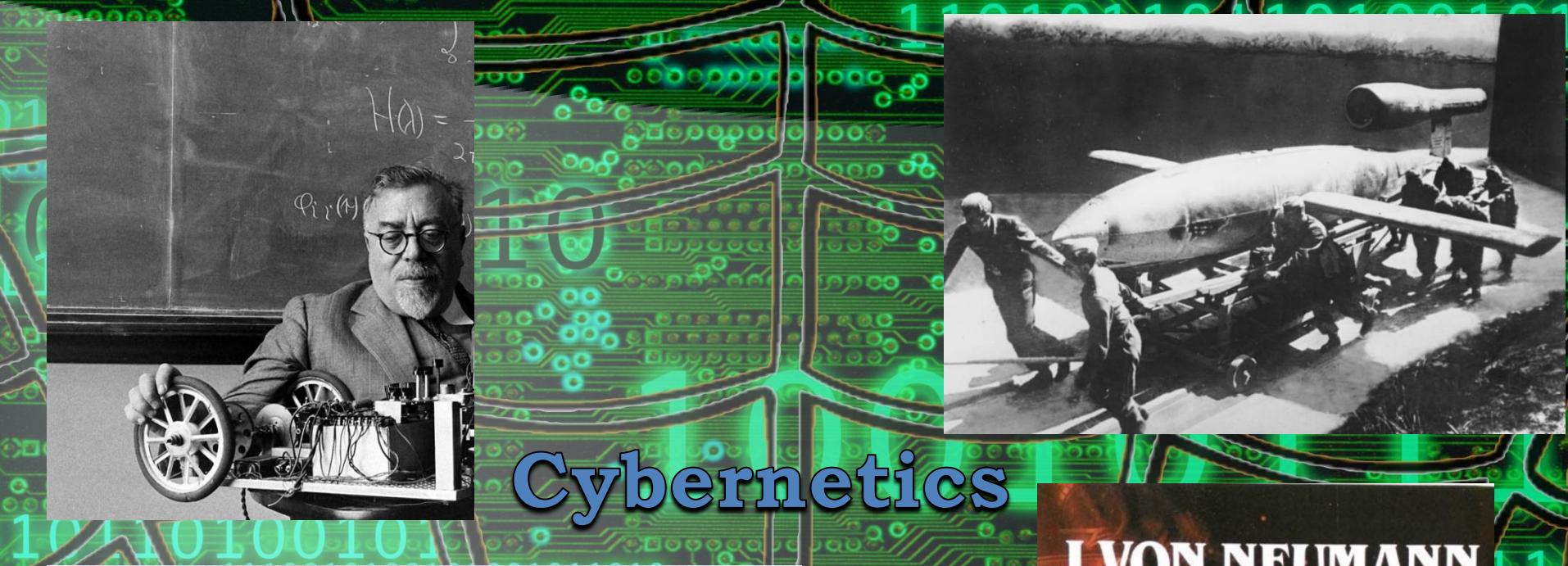
Conclusion

- History and definition
- Uses:
 - Image recognition
 - Voice recognition
 - Video processing

Other uses:

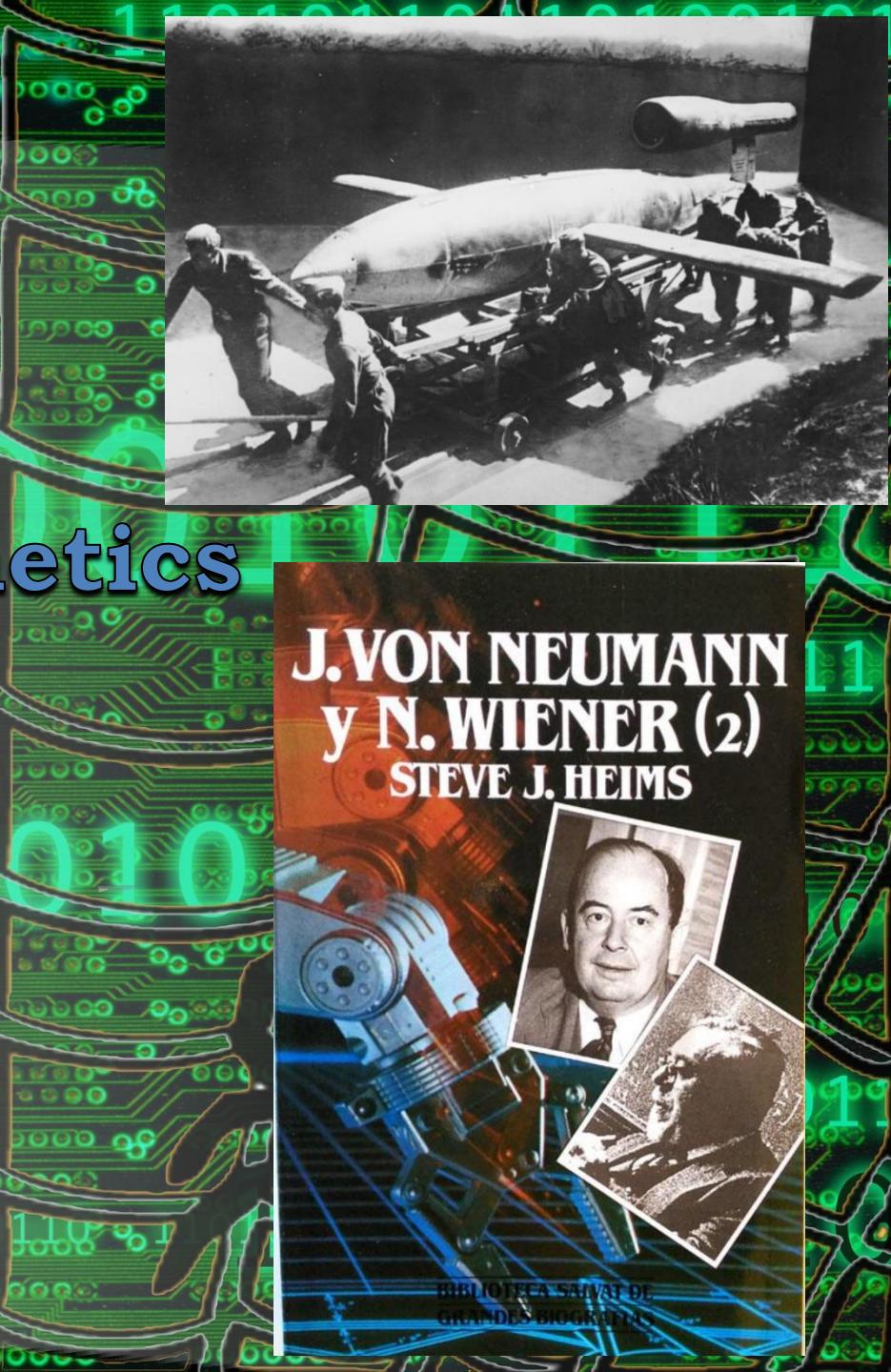
- Character recognition
- Medical diagnostics
- Chat bots
- Games (playing chess)
- Self-driving cars
- Recommendation systems (Google, Netflix, etc)
- Self-service in Kroeger/Walmart
- etc





Cybernetics

J. VON NEUMANN
y N. WIENER (2)
STEVE J. HEIMS



BIBLIOTECA SALVAT DE
GRANDES BIOGRAFÍAS

Convert hours to minutes

- $f(x) = 60x$

Sentiment Analysis

“Well, that movie was just fantastic. I couldn't stop laughing the entire time. The plot was so riveting, and the acting was top-notch. Definitely worth my time. Not.”

- Positive
- Neutral
- Negative

Sentiment Analysis

x – text

$f(x)$ – sentiment of x (1, 0, or -1)

Does $f(x)$ is exact?

Does $f(x)$ is a function?

More Problems

- Not clear dependences
- No exact formula
- We have some examples
- We are OK with an approximate solution
- We can use examples for predictions!

Logistics

My Courses

- Schedule
- Lectures
- Homeworks
- Discussions

Prerequisites

Calculus: derivatives and gradient

Linear Algebra: matrices, norms, and eigenvalues

Probability & Statistics: main distributions, expectation, variance, and correlation.

Python

Notations

x – example/sample of a RV X

\mathbf{X} – all possible examples/sample space

y – response/target (\$ profit)

\mathbf{Y} – target space

Training data: $(x^{(i)}, y^{(i)})$, $i = 1, 2, \dots, N$

Features/factors: $x = (x_1, x_2, \dots, x_d)$

Features

Demographical:

Average age

Average income

Real Estate:

Average price for a squared ft

Traffic:

Number of cars/hour

Prediction/Model/Algorithm

$$a(x) \approx y$$

Linear model:

$$a(x) = w_0 + w_1 x_1 + \cdots + w_d x_d$$

Example:

$$\begin{aligned} a(x) &= 100.000 - 10 * (\textit{distance to closest restaurant}) \\ &\quad + 50 * (\textit{distance to a train station}) \end{aligned}$$

Loss function

$$\text{Loss}(a(x), y)$$

Example 1:

$$\text{Loss}(a(x), y) = 0$$

Example 2:

$$\text{Loss}(a(x), y) = (a(x) - y)^2$$

Example 3:

$$\text{Loss}(a(x), y) = [a(x) \neq y]$$

Expected Loss/Quality Metrics

$$Error = \frac{1}{N} \sum_{i=1}^N Loss(a(x^{(i)}), y^{(i)})$$

Example 1:

$$Error = 0$$

Example 2:

$$MSE = \frac{1}{N} \sum_{i=1}^N (a(x^{(i)}) - y^{(i)})^2$$

Example 3:

$$Counts = \frac{1}{N} \sum_{i=1}^N [a(x^{(i)}) \neq y^{(i)}] = \frac{\# wrong\ predictions}{N}$$

Training/Learning

$$Error = \frac{1}{N} \sum_{i=1}^N Loss(a(x^{(i)}), y^{(i)}) \rightarrow min$$

$$a^* = \operatorname*{argmin}_a Q(a, X)$$

In linear model

$$a(x) = w_0 + w_1 x_1 + \cdots + w_d x_d$$

we find the weights w_k^* .

Other Problem Settings

Supervised

Unsupervised:
clustering

Semi-supervised:
recommendation and ranking systems

Features

Binary/Boolean {0, 1}

Numerical

Categorical: ordinal and nominal

Textual

Datetime

Geospatial

Image

Audio