Seminar HEPD

Artificial Intelligence and
## Rising of Artificial Neural Networks

Andrey Ye. Shevel

Presentation outlook

- Artificial Intelligence

- Machine learning

- Where and who interested for

- Examples

- Risks of use

- Conclusion

Artificial Intelligence

- Artificial Intelligence is not Artificial Intellect.

- Intelligence – can be described as the ability to perceive or infer information, and to retain it as knowledge to be applied towards adaptive behaviors within an environment or context.

- Intellect – human feature.

Turing test (1950)

A.Ye. Shevel
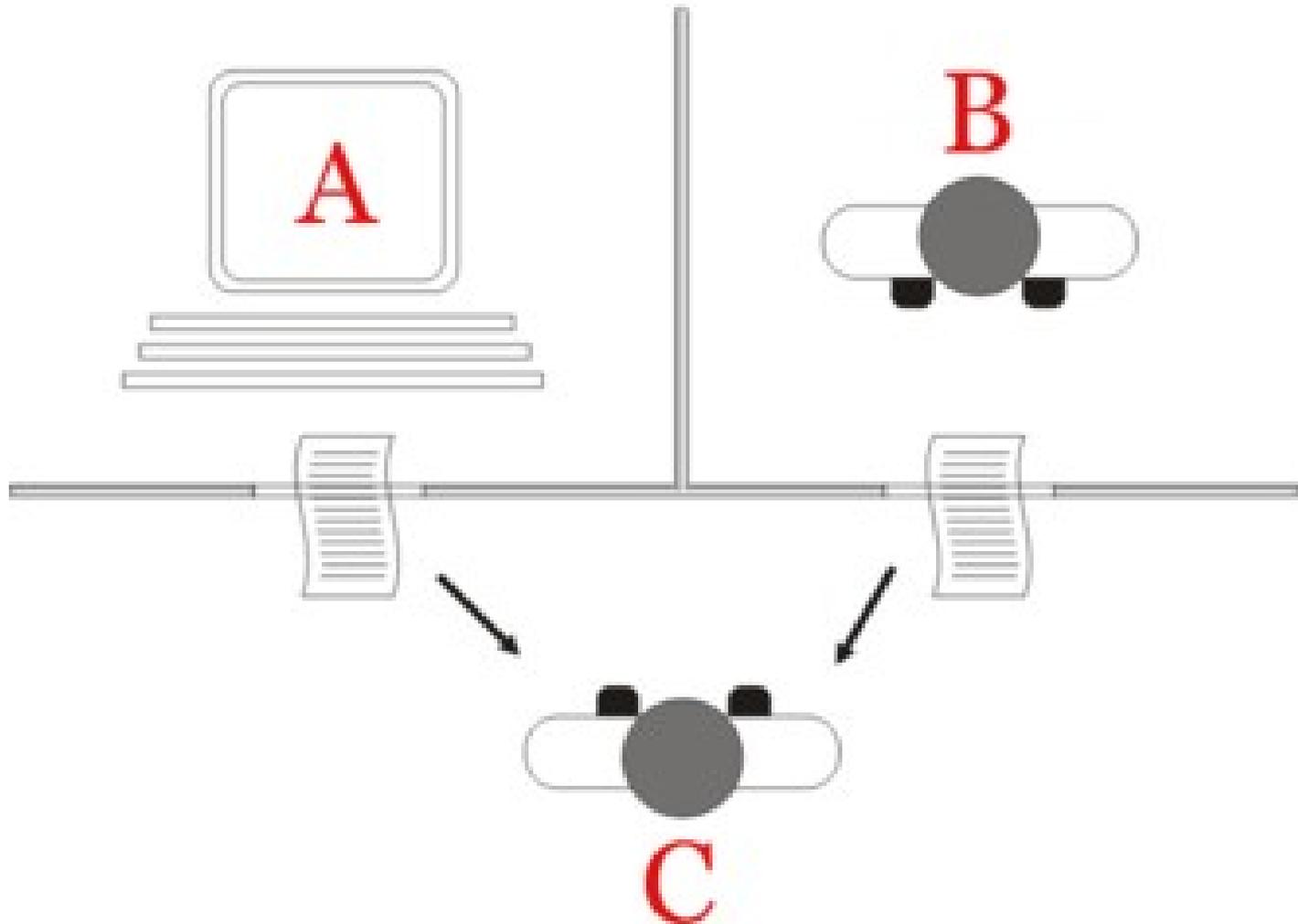
How to teach the machine

- Machine learning (ML) is the study of computer algorithms that can improve automatically through experience and by the use of data. It is seen as a part of artificial intelligence.

    - From: https://en.wikipedia.org/wiki/Machine_learning

- "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."

    - From: Mitchell, T. M. (1997). Machine Learning. McGraw-Hill, New York.

- How to implement machine learning?

Artificial Neural Network & Deep Learning

- Artificial neural networks (ANNs) were inspired by information processing and distributed communication nodes in biological systems. ANNs have various differences from biological brains. Specifically, artificial neural networks tend to be discrete, while the biological brain of most living organisms is analogue.

    – Cite from: https://en.wikipedia.org/wiki/Deep_learning

Simple artificial neural network

A.Ye. Shevel
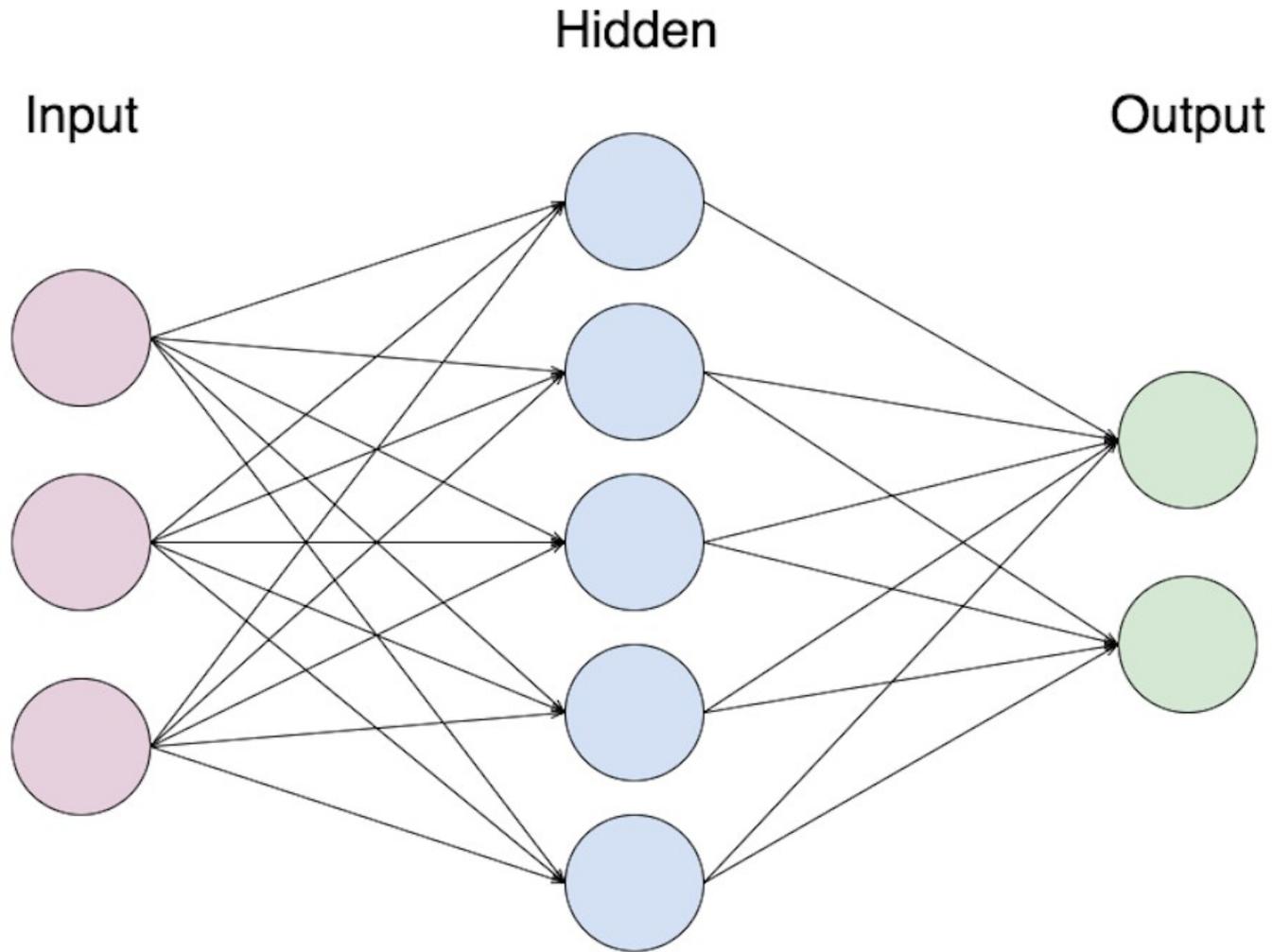
One artificial neuron logic
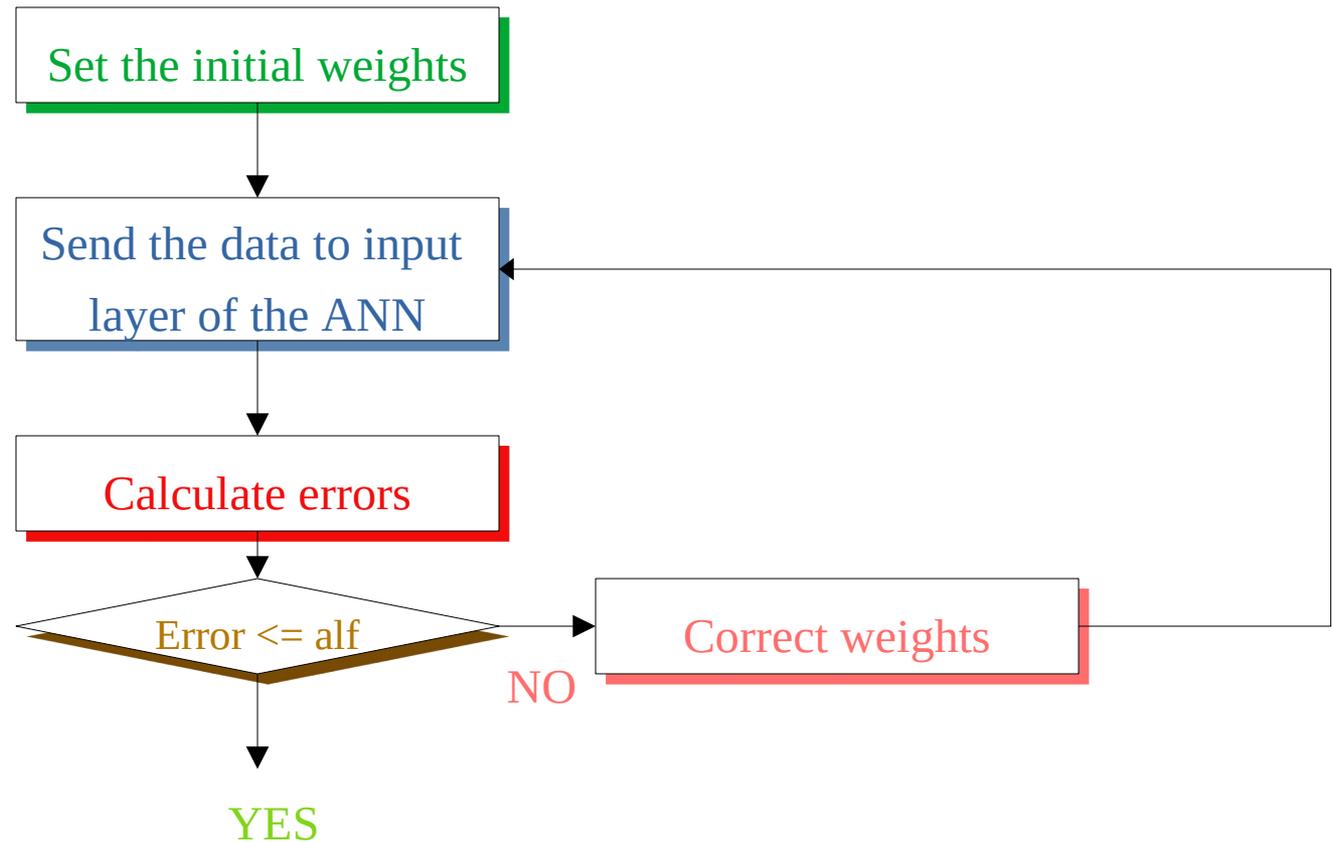
$$f\left(b + \sum_{i=1}^{n} x_i w_i\right)$$
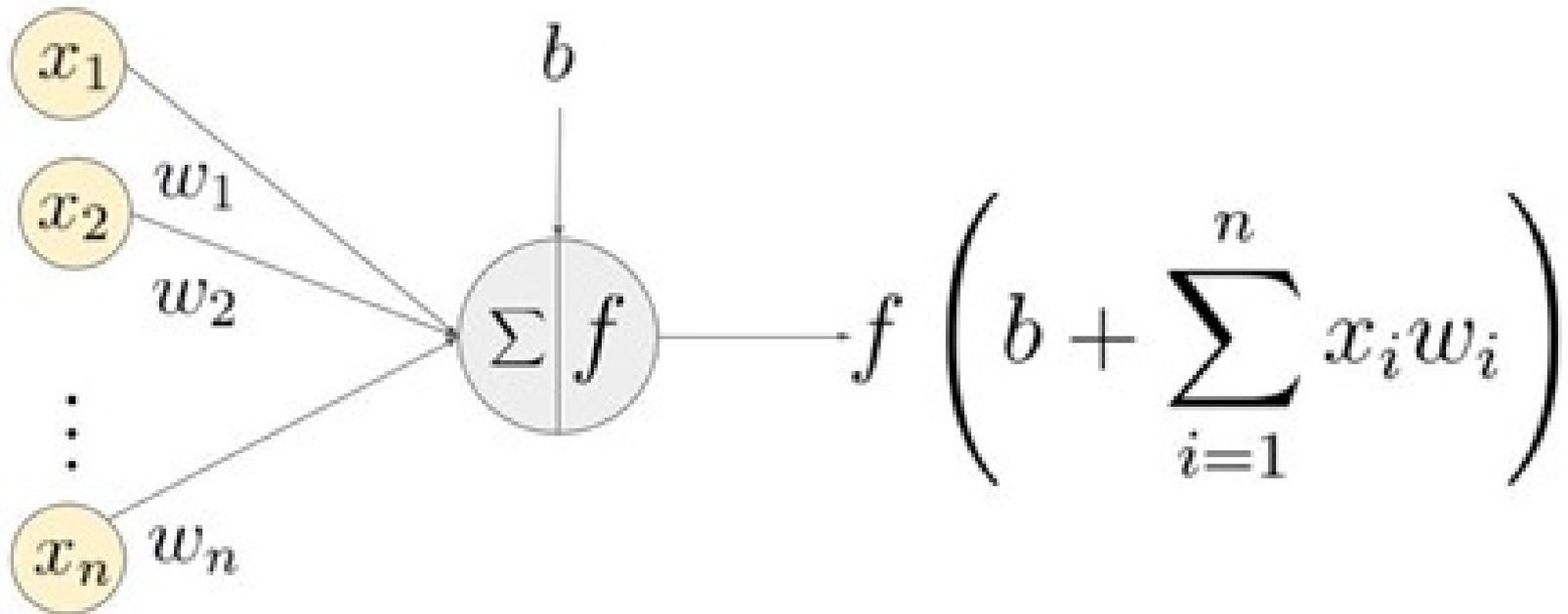
An example of a neuron showing the input (x_1 - x_n), their corresponding weights (w_1 - w_n), a bias (b) and the activation function f applied to the weighted sum of the inputs.

Model Errors

- In general it has to be created loss function which does calculate model error.

- Obviously the lower errors the better.

- Having such the loss function it is possible to use gradient descent to find minimum error.

  - On the convergence gradient descent in ANN https://arxiv.org/pdf/2102.09924.pdf

The cycle of learning
two data sets: training & testing

Set the initial weights

Send the data to input layer of the ANN

Calculate errors

Error <= alf

Correct weights

NO

YES

Main types of machine learning and challenges

- Supervised learning

  - training data in supervised learning consist of input–output pairs: gives good performance but expensive in practice

- Unsupervised learning methods deal with the problems where we can only access the input; learning algorithm should be able to figure out some criteria to group similar inputs together using only the information of all possible inputs: much harder.
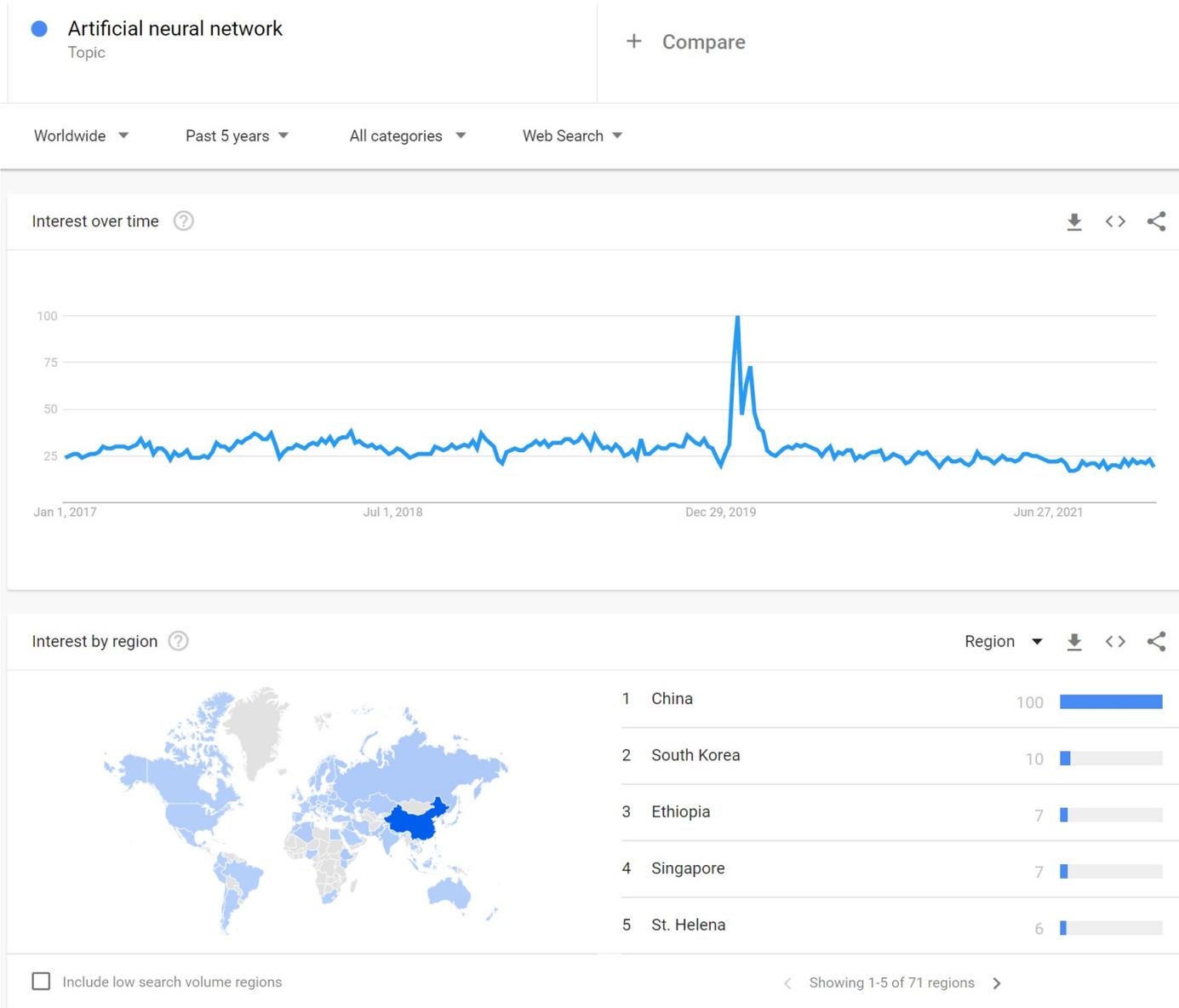
Complexity

A.Ye. Shevel
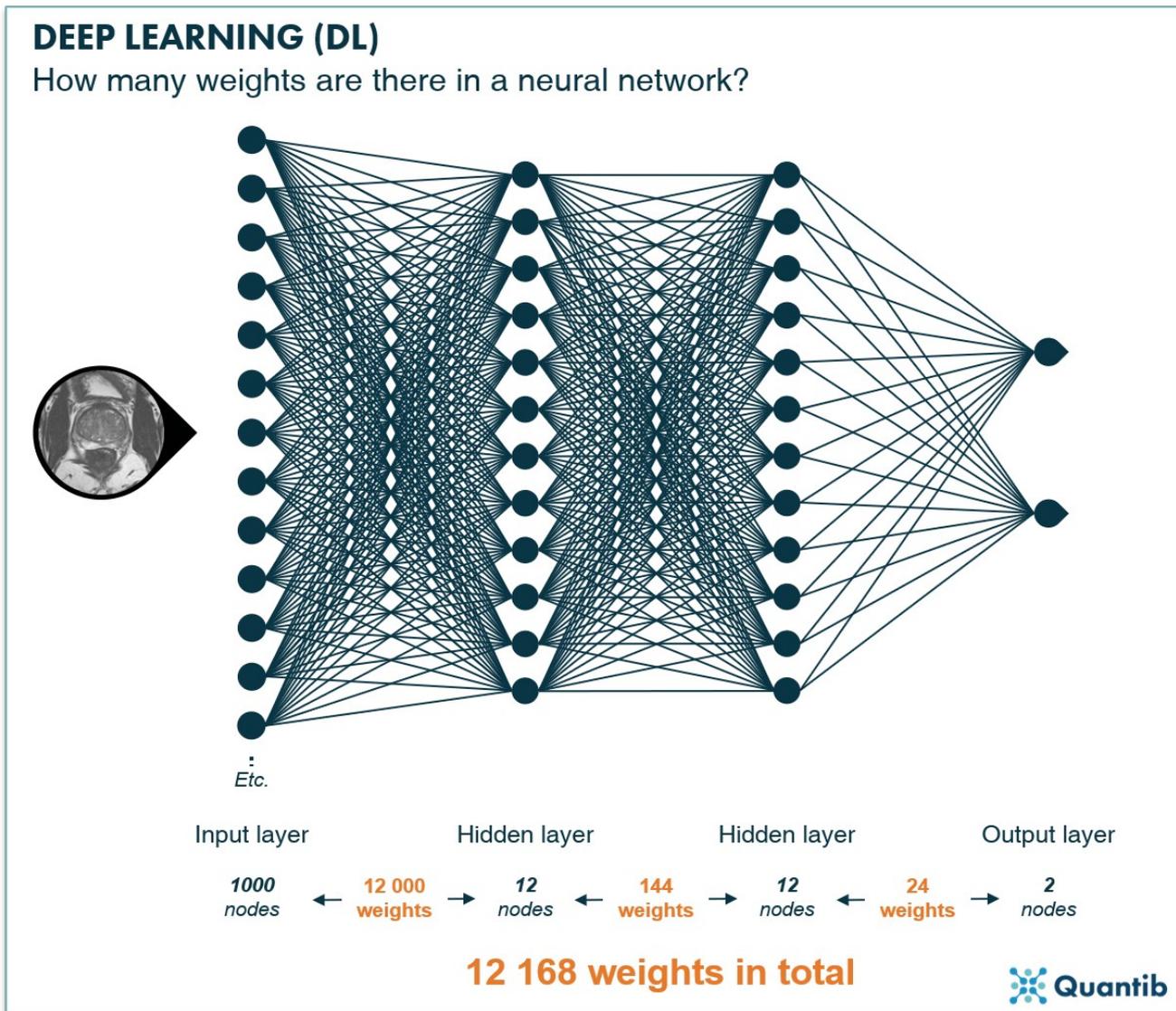
Who is talking about & working on

- Journals: Nature, Science, Cern Courier, *Foreign Affairs*, ...

- Companies: Google, IBM, Microsoft, Facebook, Yandex, Sberbank, ...

- Many other sources and organizations ...
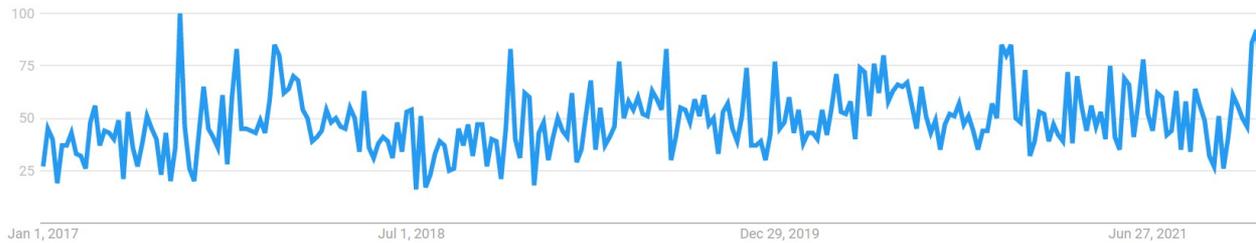
ANN in the World

Artificial neural network — Topic

Worldwide ▼ | Past 5 years ▼ | All categories ▼ | Web Search ▼

Interest over time ⓘ

(Chart from Jan 1, 2017 to Jun 27, 2021, with a peak near Dec 29, 2019)

Interest by region ⓘ — Region ▼

#	Region	Value
1	China	100
2	South Korea	10
3	Ethiopia	7
4	Singapore	7
5	St. Helena	6

☐ Include low search volume regions   ‹ Showing 1-5 of 71 regions ›

ANN in Russia

How many presentations devoted to

In which areas

- Biology

- Physics

- Mathematics

- Linguistics

- Military

- Almost everywhere in science and in real world when you need to get relatively fast informal conclusion about features in huge volume of multidimensional data.

Big projects

- Government projects

  - Human Brain project [EU, started in 2013]

  - Brain Initiative [USA, started in 2013]

  - AI development in Russia [started in 2019]
    President order № 490 10.10.20

- Private

  - Deep Mind [Started in 2010] - AlphaBet

  - OpenAI [Started in 2015] – Ilon Mask?

  - Many others

- Examples

ANN in Math

- Alex Davies et all // Advancing mathematics by guiding human intuition with AI // https://www.nature.com/articles/s41586-021-04086-x

- Davide Castelvecchi // DeepMind's AI helps untangle the mathematics of knots // https://www.nature.com/articles/d41586-021-03593-1

A.Ye. Shevel

ANN in Biology

- John Jumper et al // Highly accurate protein structure prediction with AlphaFold // https://www.nature.com/articles/s41586-021-03819-2

- Robert Service // 2021 BREAKTHROUGH OF THE YEAR // Protein structures for all // https://www.science.org/content/article/breakthrough-2021

ANN in Physics

- Atılım Güneş Baydin et al // Toward Machine Learning Optimization of Experimental Design // https://www.tandfonline.com/doi/full/10.1080/10619127.2021.1881364

- Timo Felser et al // Quantum-inspired machine learning on high-energy physics data // https://www.nature.com/articles/s41534-021-00443-w

- Carlos Bravo-Prieto et al // Style-based quantum generative adversarial networks for Monte Carlo events // CERN-TH-2021-139, TIF-UNIMI-2021-14

- Stephen B Menary and Darren D Price // Learning to discover: expressive Gaussian mixture models for multi-dimensional simulation and parameter inference in the physical sciences // https://iopscience.iop.org/article/10.1088/2632-2153/ac4a3b/pdf

- Theo Heimel, Gregor Kasieczka, Tilman Plehn, Jennifer M Thompson // QCD or what? // https://scipost.org/10.21468/SciPostPhys.6.3.030

- Datasets used to train the Generative Adversarial Networks used in ATLFast3 http://opendata.cern.ch/record/15012

- Inter-experimental Machine Learning (IML) Working Group https://iml.web.cern.ch/homepage

Other examples

- https://yandex.ru/lab/yalm - Russian text generator

- https://deepai.org/machine-learning-model/text-generator - English text generator

- https://www.youtube.com/watch?v=l82PxsKHxYc - video conversion (fake example)

- **Chinese language model with 247.5 billion parameters -** Yuan 1.0 - launched in 2021

Development of ANN

- New architecture: Generative Adversarial Networks

  - is class of ANN - two neural networks contest with each other in a game (in the form of a zero-sum game, where one agent's gain is another agent's loss) developed in 2014.

- Tools: models, libraries

  - PyTorch

  - Transformers

  - TensorFlow

  - Keras

  - Skikit-learn, etc

Fundamental (unanswered?) questions

- Is there fundamental prove of effective use the ANN approach ?

    - Yes: e.g. pls take the textbook by Jiang, Hui // Machine Learning Fundamentals. Cambridge University Press. // York University, Toronto. https://doi.org/10.1017/9781108938051

- What is cost for training ANN ?

    - Improvements by factor K will rise cost computation in between K^2 and K^4 https://spectrum.ieee.org/deep-learning-computational-cost

Chips for ANN

- Saptadeep Pal et al // Designing a 2048-Chiplet, 14336-Core Waferscale Processor https://ieeexplore.ieee.org/document/9586194

- Loihi 2 packs 1 million neurons in a chip // https://spectrum.ieee.org/neuromorphic-computing-with-lohi2

- Wafer Scale Engine chip // With 2.6 trillion transistors and 850,000 cores // https://www.hpcwire.com/2021/04/20/cerebras-doubles-ai-performance-with-second-gen-7nm-wafer-scale-engine/

    - More than 25 leading semiconductor companies are producing ANN chips today (80% companies started do that after 2014 // pls see John VerWey // The Other Artificial Intelligence Hardware Problem // The computer Jan 2022 // pp 34-42).

- Also see CPU fo AI NVIDIA A100, RISC-V

ANN as a Service

- **SambaNova** - https://sambanova.ai (up to 100 trillions of parameters) // https://spectrum.ieee.org/sambanova-ceo-ai-interview

- **Cerebras** - https://cerebras.net (up to 120 trillions of parameters) // https://spectrum.ieee.org/cerebras-ai-computers

- **Meta (Facebook) RSC** (*coming*) - https://spectrum.ieee.org/meta-ai-supercomputer

Risks in nearest future

- Internet attacks

- Military applications

- Human privacy (freedom)

Cyber attacks

- Up to now there are several cases where cyber attacks lead to hundred millions of dollars [https://www.foreignaffairs.com/articles/united-states/2021-12-10/soon-hackers-wont-be-human].

- However until today there is currently limited evidence that hackers have begun making significant use of AI techniques.

- The circumstances might be worse soon.

Military conflicts

- In the realm of global military conflict, **computer networks have become a fifth domain**, in addition to the traditional four of land, sea, air, and space.
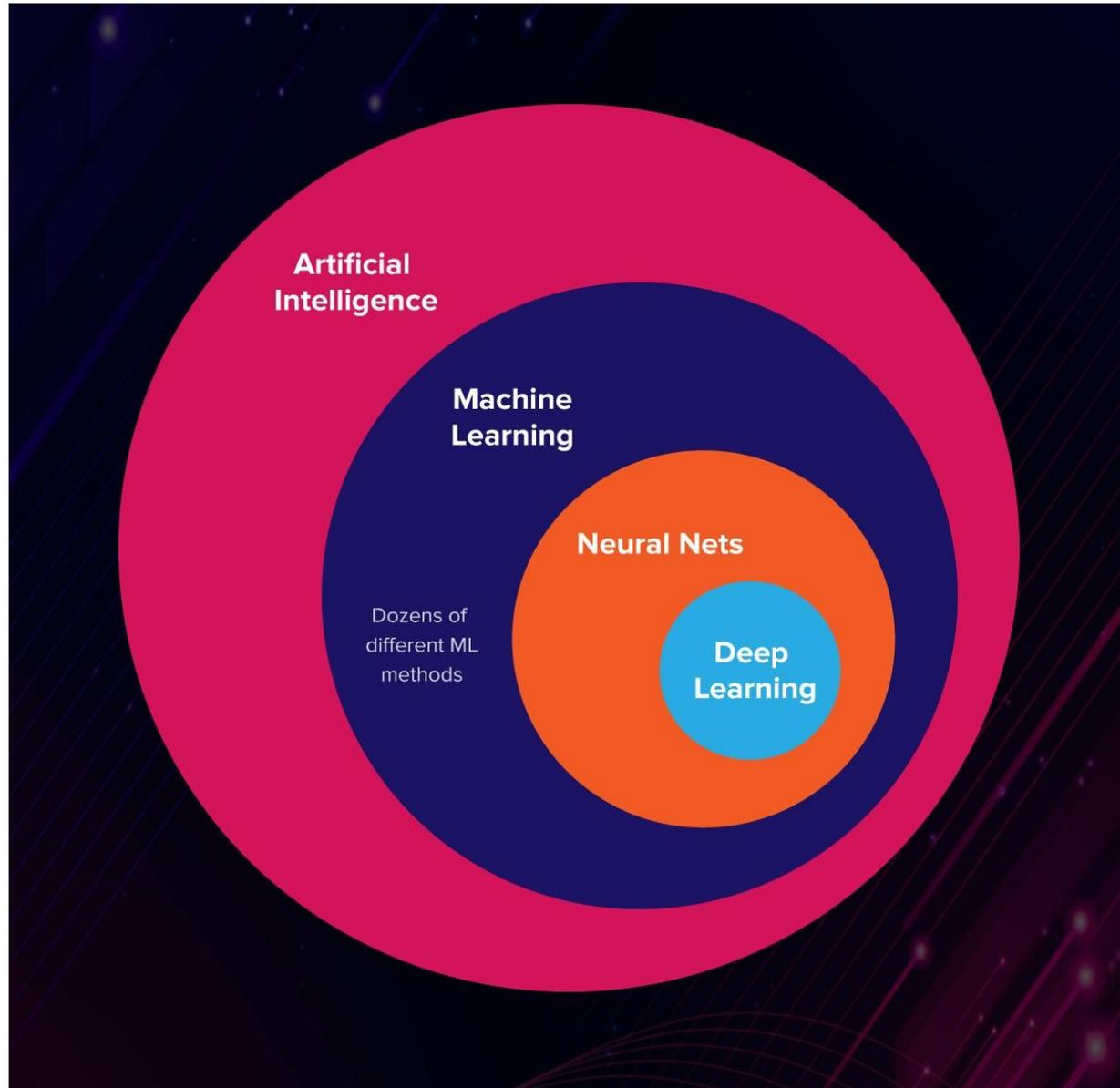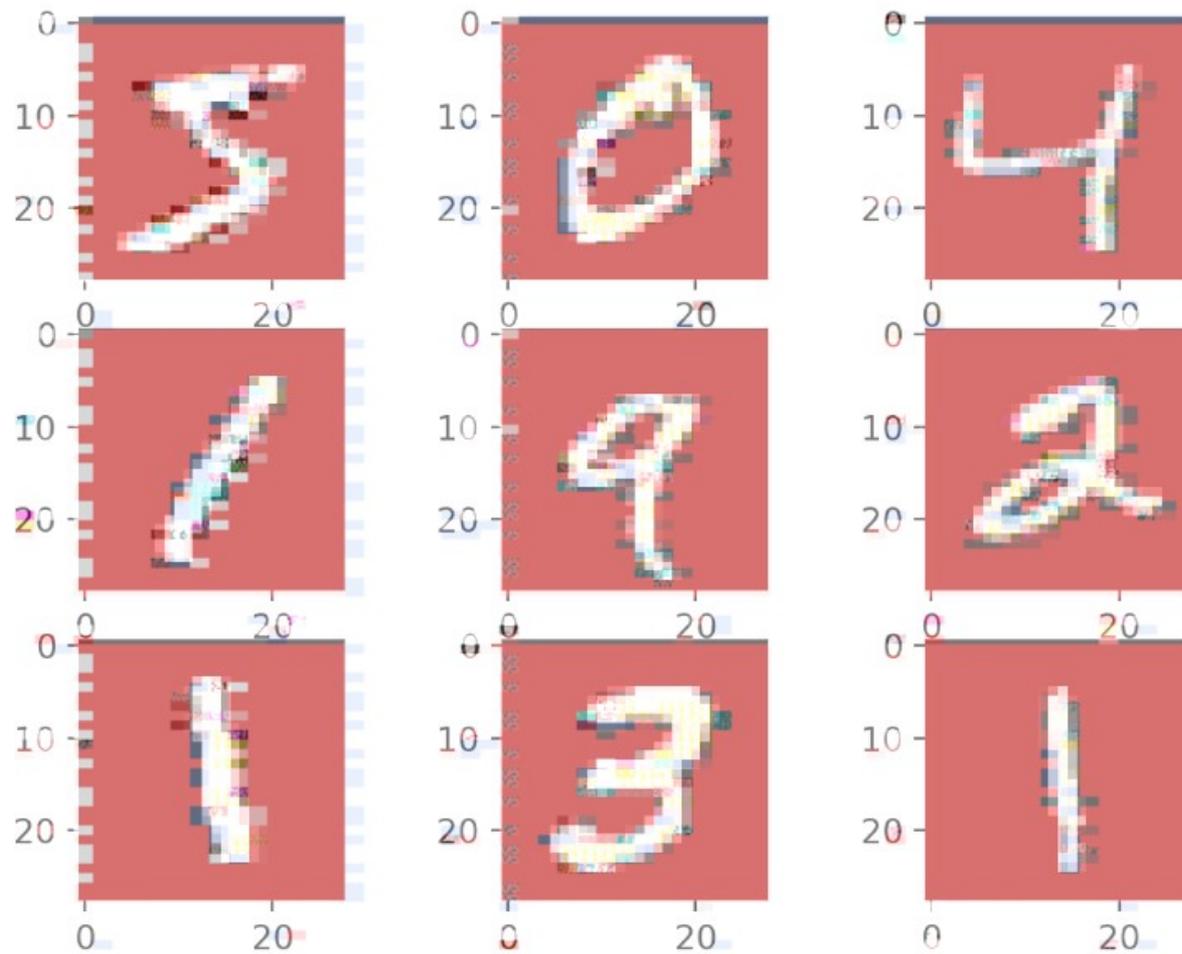
A.Ye. Shevel

Human privacy

- Human privacy is under risk as well (it is just consequence of previous slides).
  - Generation of the photo and video with any person.
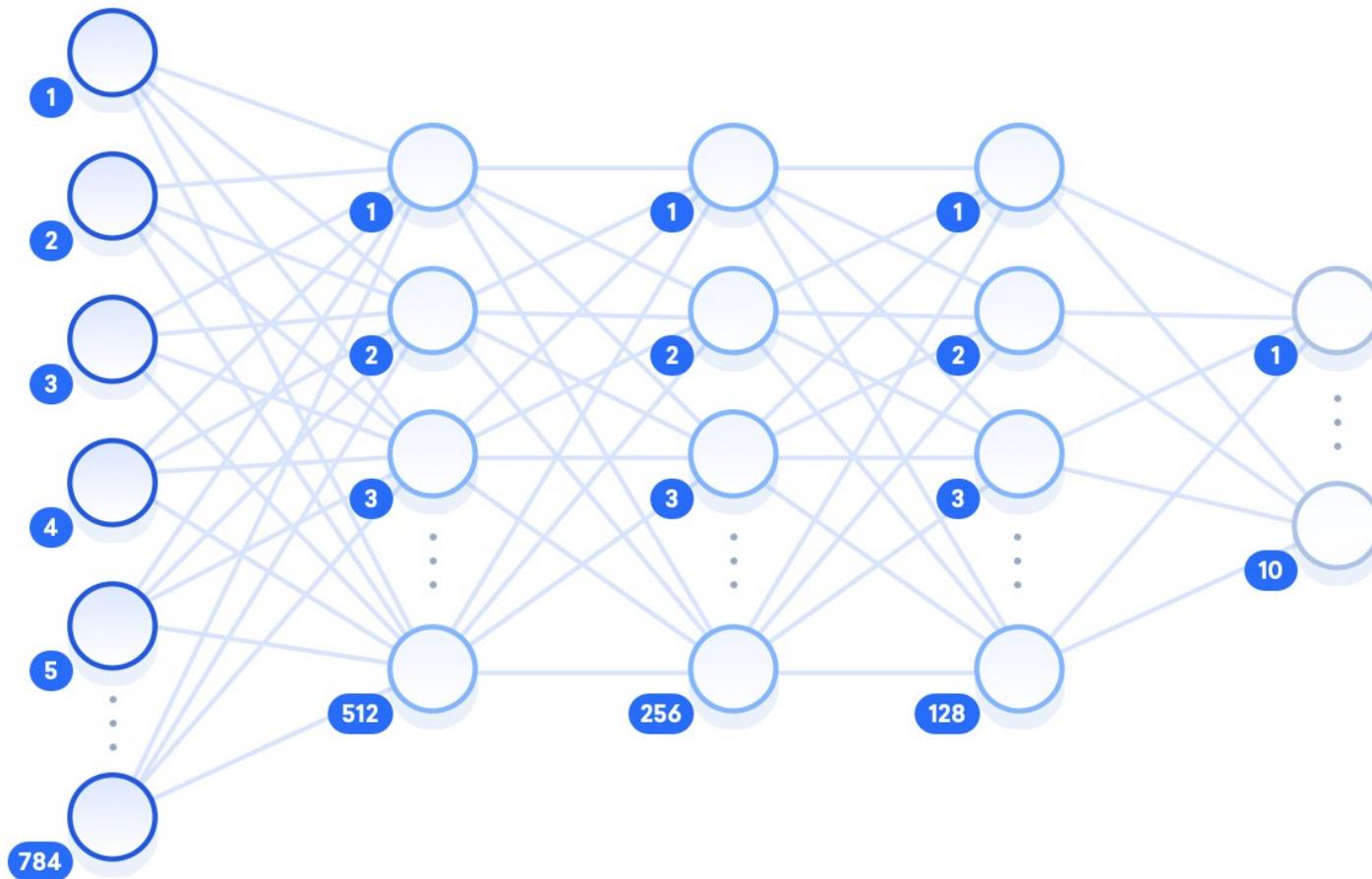 - As the result several laws were approved (e.g. in California in 2019)

Conclusion

- ANN is important topic for all areas where huge volume of multidimensional data takes place (HEP is not exclusion).

  - Main reason to use trained ANN is a shortened time to get the solution or hint.

- ANN is only the high level toolkit, i.e. the person who use it has to have deep understanding of the application area and has full responsibility for the results.

- List of unsorted references on AI and ANN is available at http://hepd.pnpi.spb.ru/CSD/CSD_Docs/References.pdf

Relations in between AI & ANN

A.Ye. Shevel

X

Y

A.Ye. Shevel 34

$$\frac{1}{1 + \exp(-\sum_j w_j x_j - b)}$$

sigmoid function