INTRODUCTION TO DEEP LEARNING

Course number: 00240332

Lecture 1: Introduction

Xiaolin Hu (胡晓林)

Dept. of Computer Science and Technology

Tsinghua University

Outline

I. General concepts
II. History
III. Applications
IV. Summary

MIT Technology Review

10 Breakthrough Technologies 2013

Deep Learning

With massive amounts of computational power, machines can now recognize objects and translate speech in real time. Artificial intelligence is finally getting smart.

Deep learning in industry

Driverless car

Face identification

Speech recognition

Web search

...

Google

阿里巴巴 Alibaba.com

HUAWEI

NVIDIA

...

Microsoft

Baidu 百度

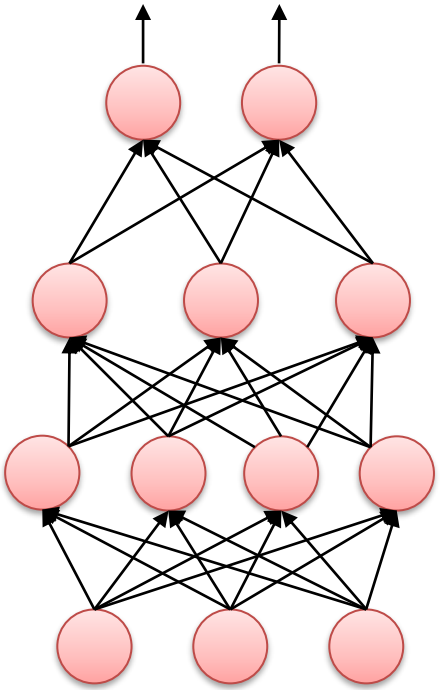
Tencent 腾讯

AMD

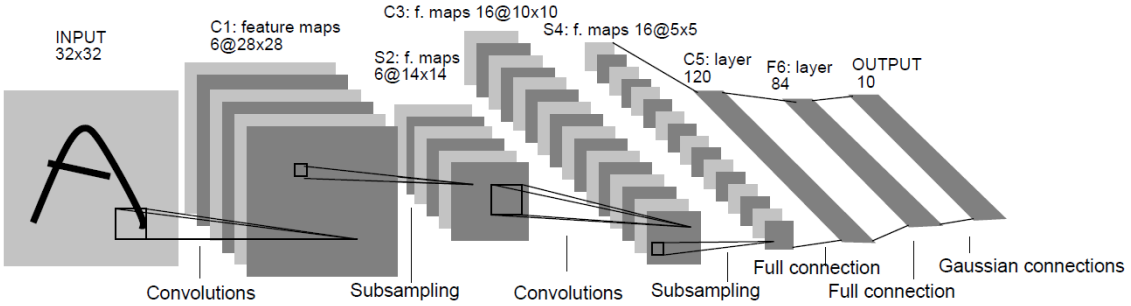
What is deep learning

- Narrow sense: artificial neural networks

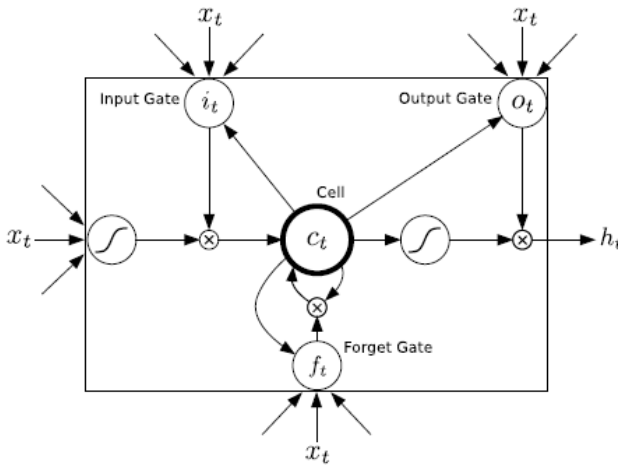
Multilayer Perceptron



Convolutional neural network



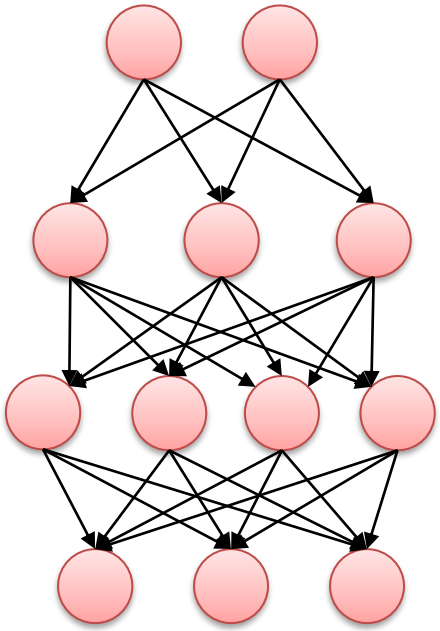
Recurrent neural network



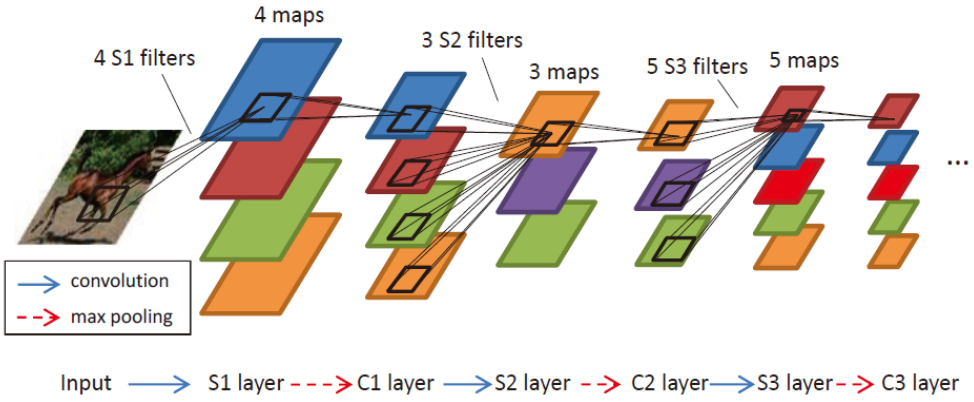
What is deep learning

- Broad sense: hierarchical machine learning models

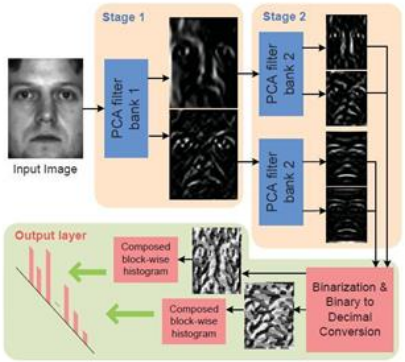
Deep belief network



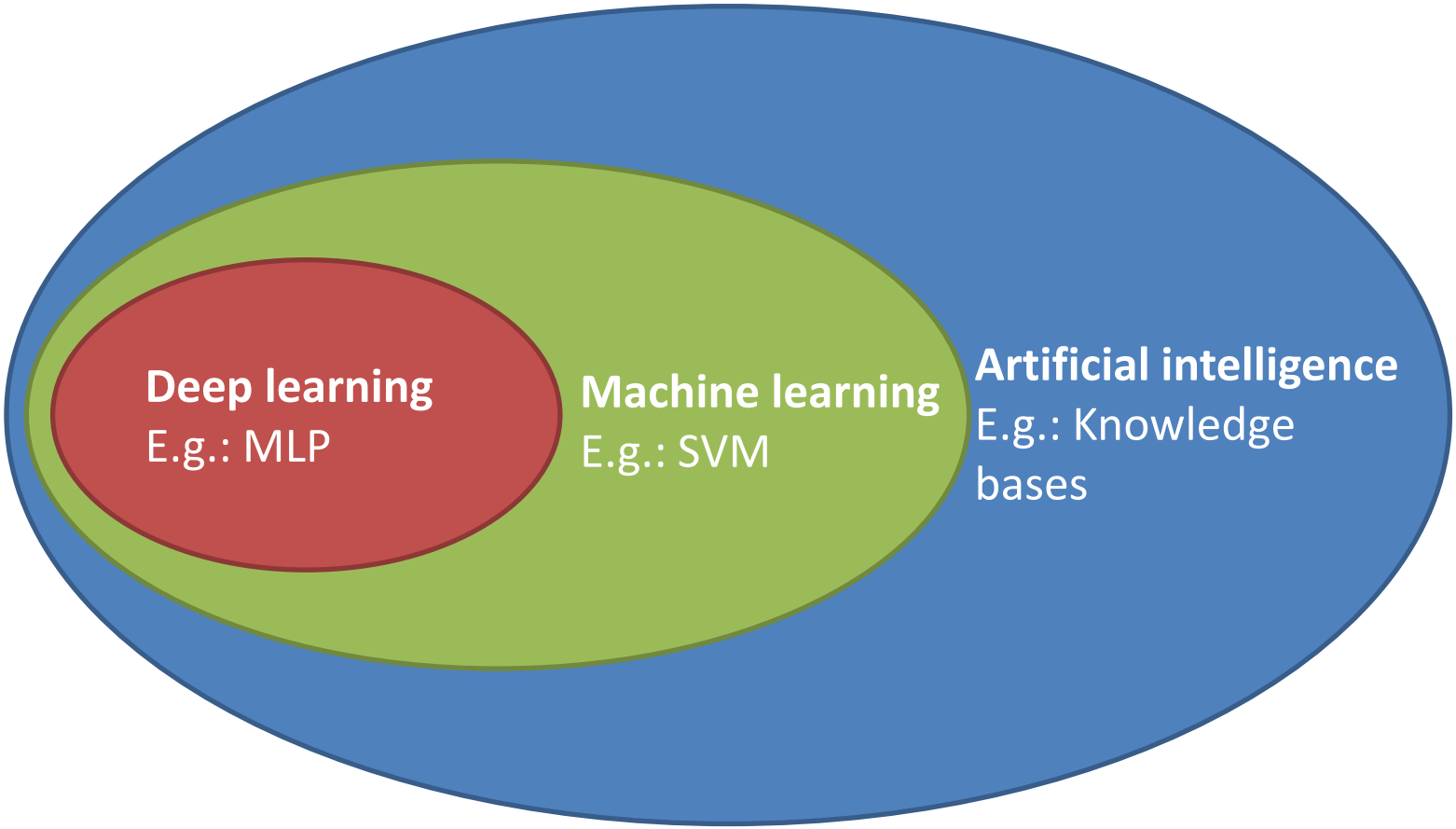
Sparse HMAX (Hu et al., 2014)



PCA net (Chan et al., 2014)



Deep learning, machine learning and AI



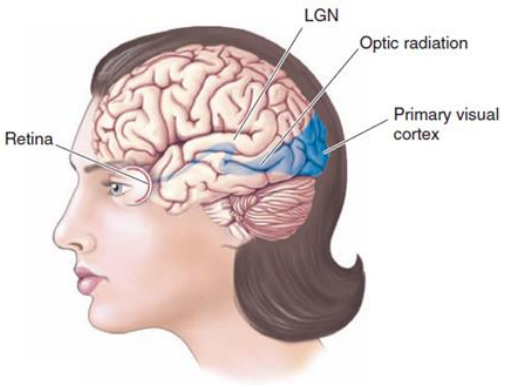
Why go deep?

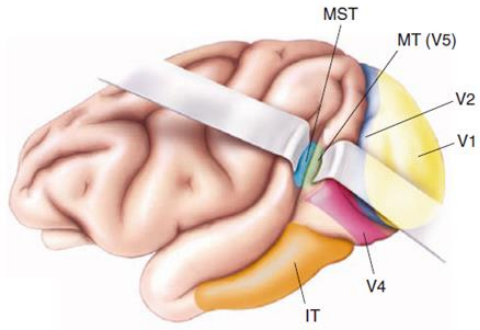
Why go deep?

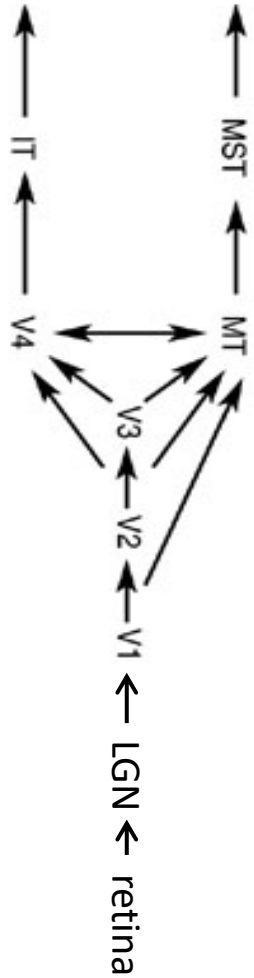
- Data are often high-dimensional.
- There is a huge amount of structure in the data, but the structure is too complicated to be represented by a simple model.
- Insufficient depth can require more computational elements than architectures whose depth matches the task.
- Deep nets provide simpler but more descriptive model of many problems.

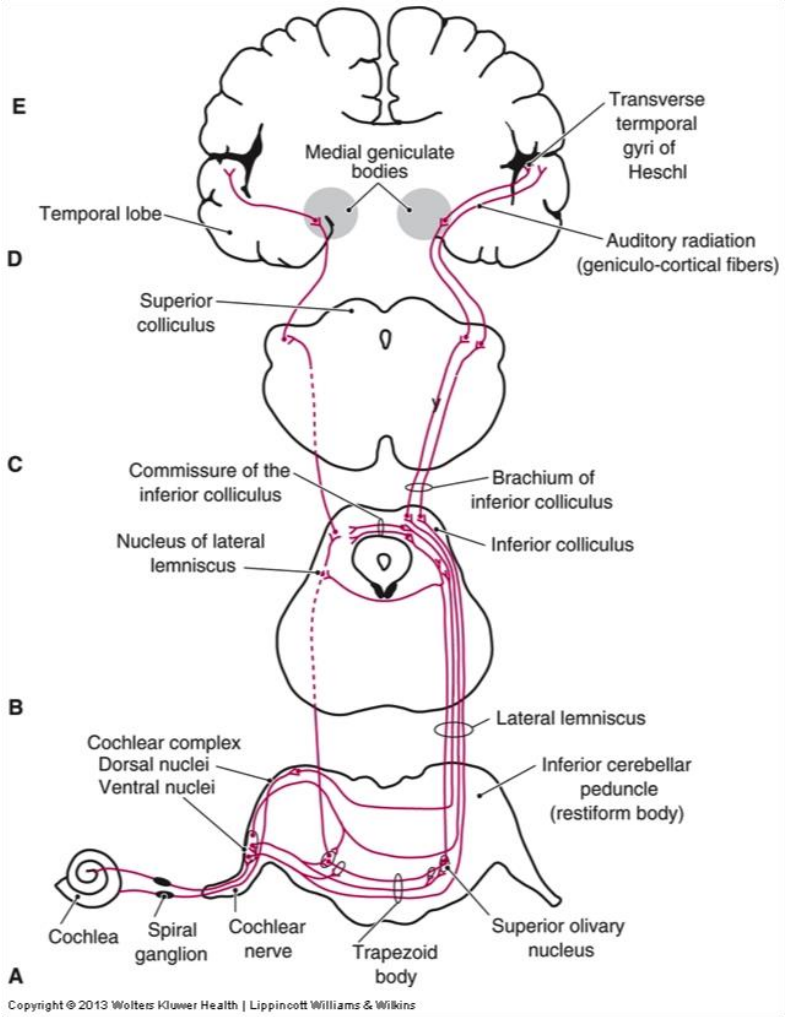
-By Geoffery Hinton

Hierarchical structures in the brain

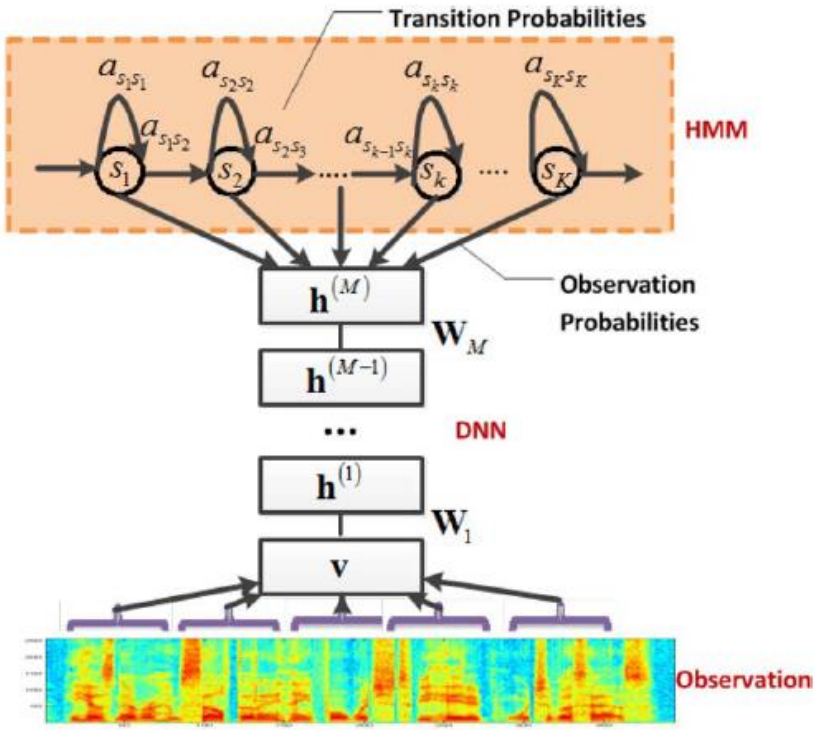




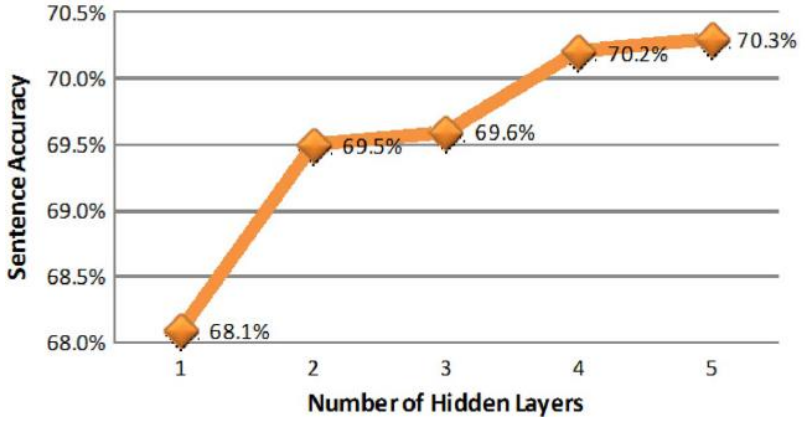




Why is deep learning so popular (1)



Dahl, Yu, Deng, Acero, IEEE TASLP, 2012



http://v.youku.com/v_show/id_XNDc0MDY4ODI0.html

Why is deep learning so popular (2)

ImageNet competition (ILSVRC)

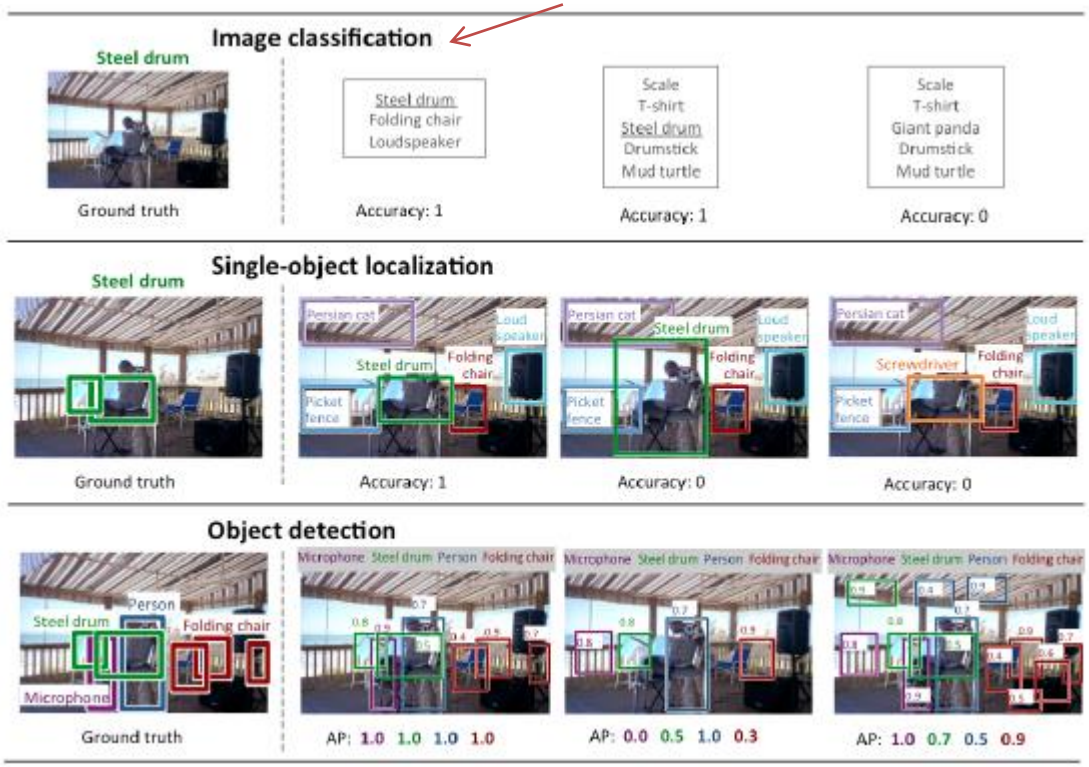
1M training images

Tasks

2010-

2011-

2013-



Top-1

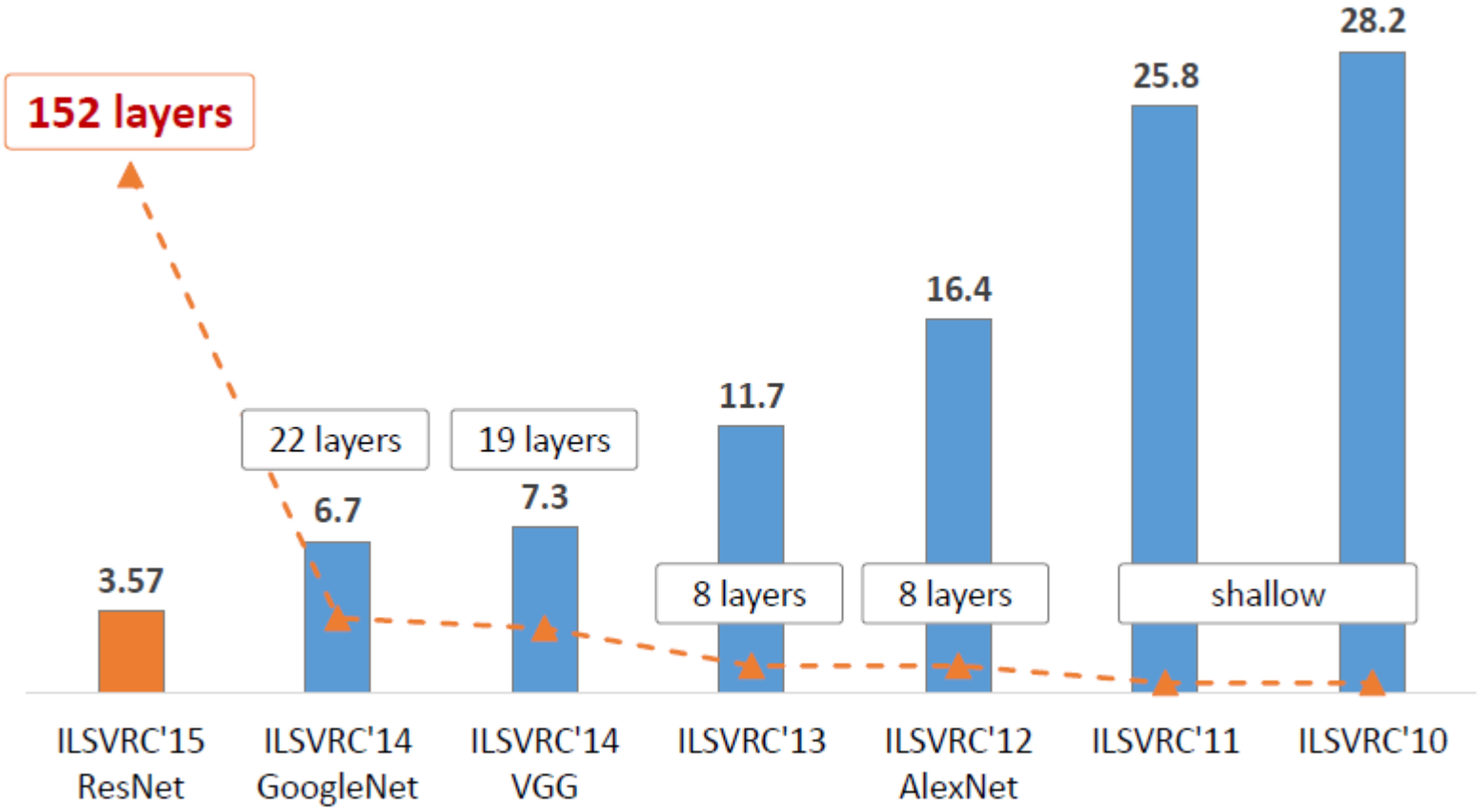
Top-5 (preferred)

Two human expert: 5.1%, 12%

The first column shows the ground truth labeling on an example image, and the next three show three sample outputs with the corresponding evaluation score.

Russakovsky, et al., 2014

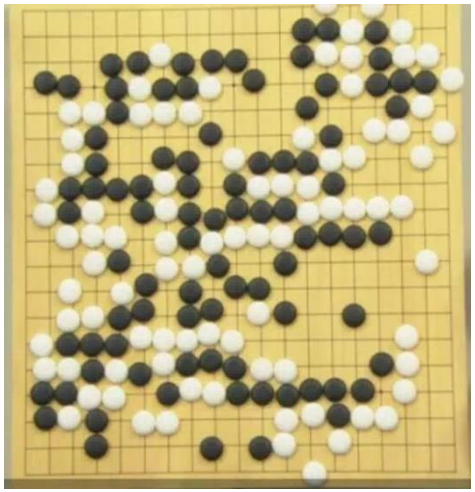
Revolution of depth



Slide credit: Kaiming He

Why is deep learning so popular (3)

28 January 2016



AlphaGo执黑

Outline

I. General concepts
II. History
III. Applications
IV. Summary

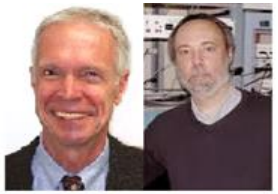
History of deep learning


Pitts McCulloch


Minsky Papert


Rosenblatt


Widrow


Hopfield Tank


Werbos


Rumelhart


Hinton


LeCun

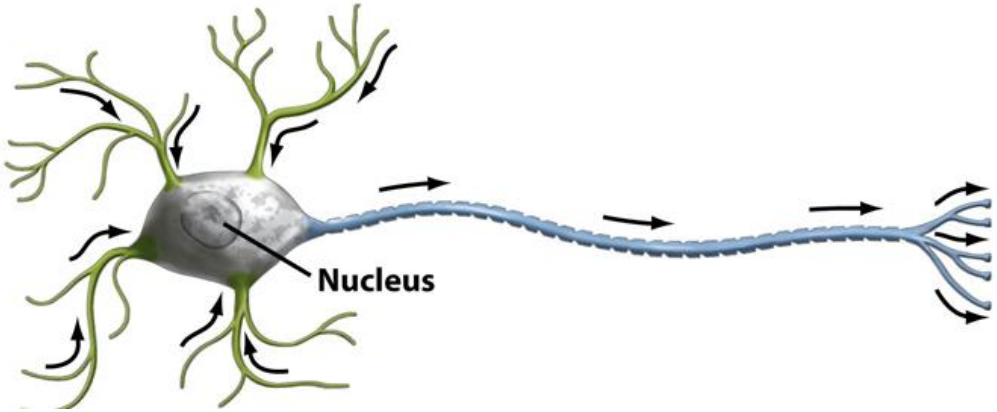

Bengio

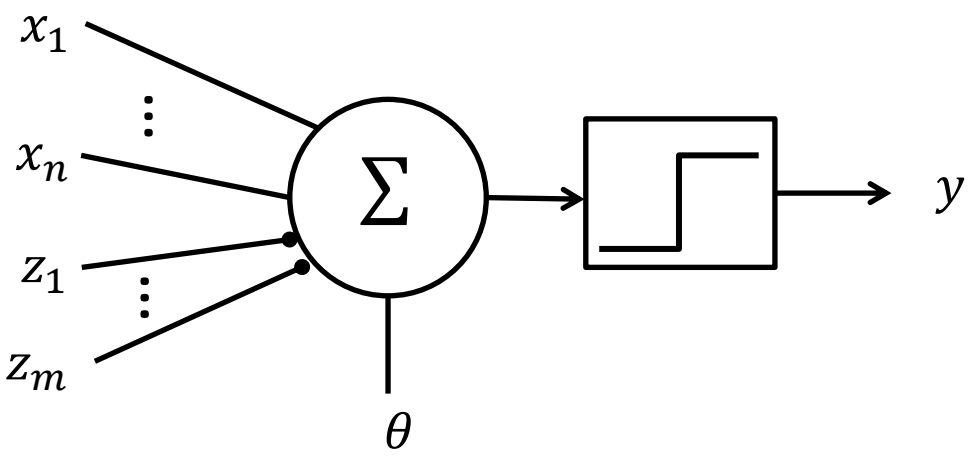
SVM

3 rises and 2 falls

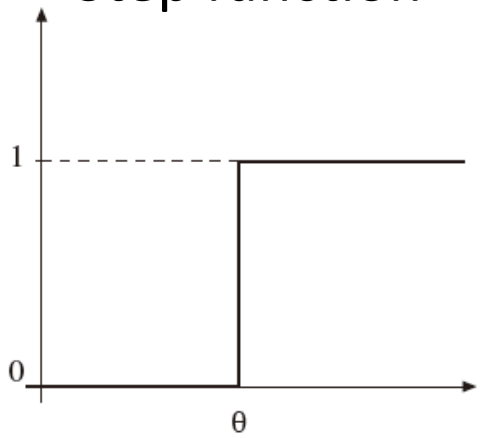
1943 ~1960 1969 1974 1982 1986 1990s 2006

Threshold Logic Unit (TLU)



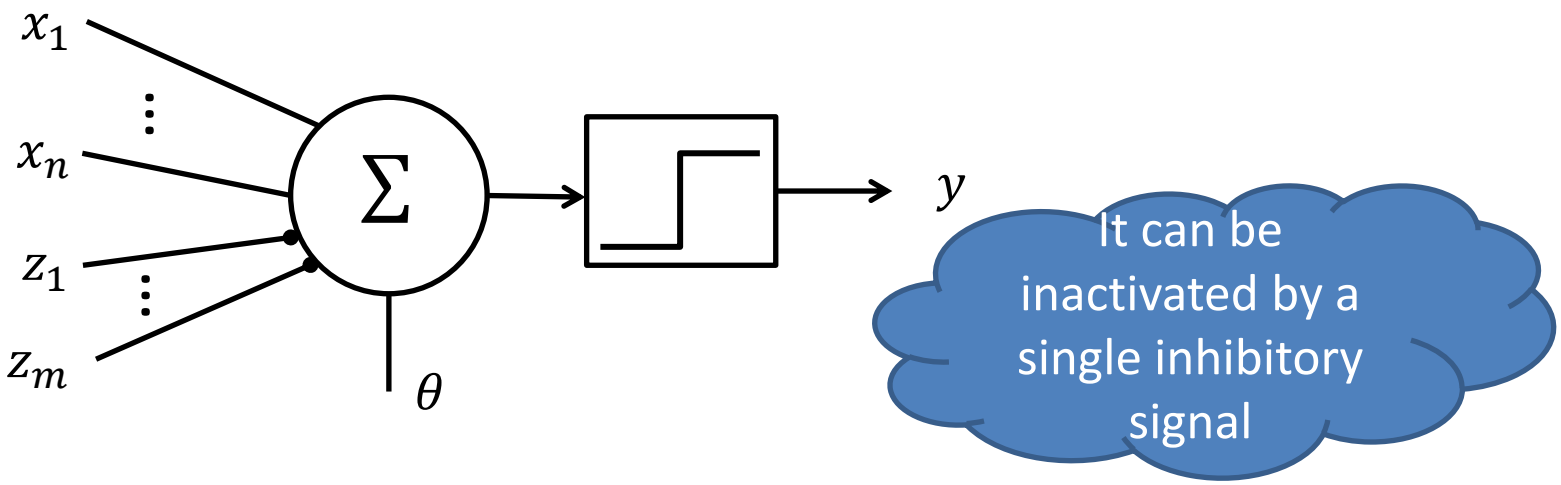


Step function



- Excitatory input x_i
- Inhibitory input z_i
- Binary output y_i
- Threshold θ

McCulloch–Pitts unit (M-P unit)



- If at least one of $z_1, z_2, \ldots, z_m$ is 1, the unit is inhibited and $y = 0$
- Otherwise the total excitation $T = \sum_{i=1}^{n} x_i$ is computed and compared with the threshold θ of the unit (if $n = 0$ then $x = 0$)
  - If $T \geq \theta$ the unit fires a 1
  - If $T < \theta$ the result is 0

Rojas: Neural Networks, Springer-Verlag, Berlin, 1996

Boolean function

- A Boolean function $f: \{0,1\}^n \rightarrow \{0,1\}$
- It can be represented by a table

Input	Output
1	0
0	1

NOT

Input	Output
(0, 1)	0
(1, 0)	0
(1, 1)	1
(0, 0)	0

AND

Input	Output
(0, 1)	1
(1, 0)	1
(1, 1)	1
(0, 0)	0

OR

Boolean function

- A Boolean function $f: \{0,1\}^n \rightarrow \{0,1\}$
- It can be represented by a table

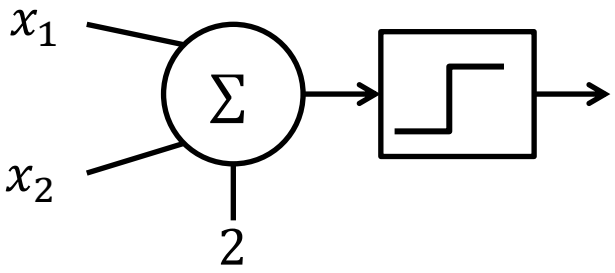
Input	Output
(0, 1, 1, 1)	1
(0, 0, 1, 1)	1
(1, 0, 0, 1)	1
All others	0

Synthesis of Boolean functions using BP units

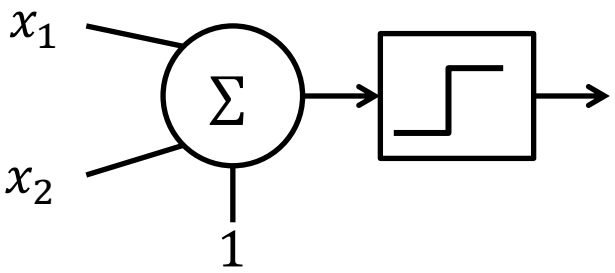
Boolean function: $\{0,1\}^n \rightarrow \{0,1\}$

- Conjunction and disjunction

AND



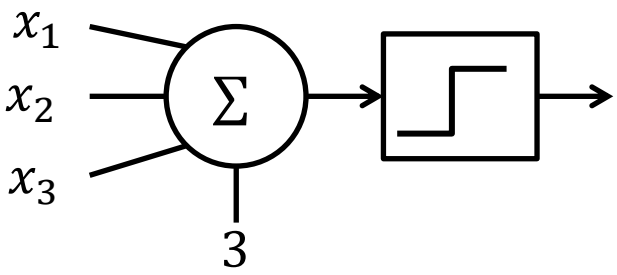
OR



单选题 1分

Setting

What function does this unit implement?



A AND (conjunction)

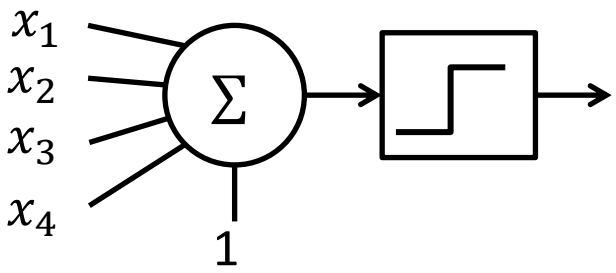
B OR (disjunction)

Submit

单选题 1分

Setting

What function does this unit implement?



A AND (conjunction)

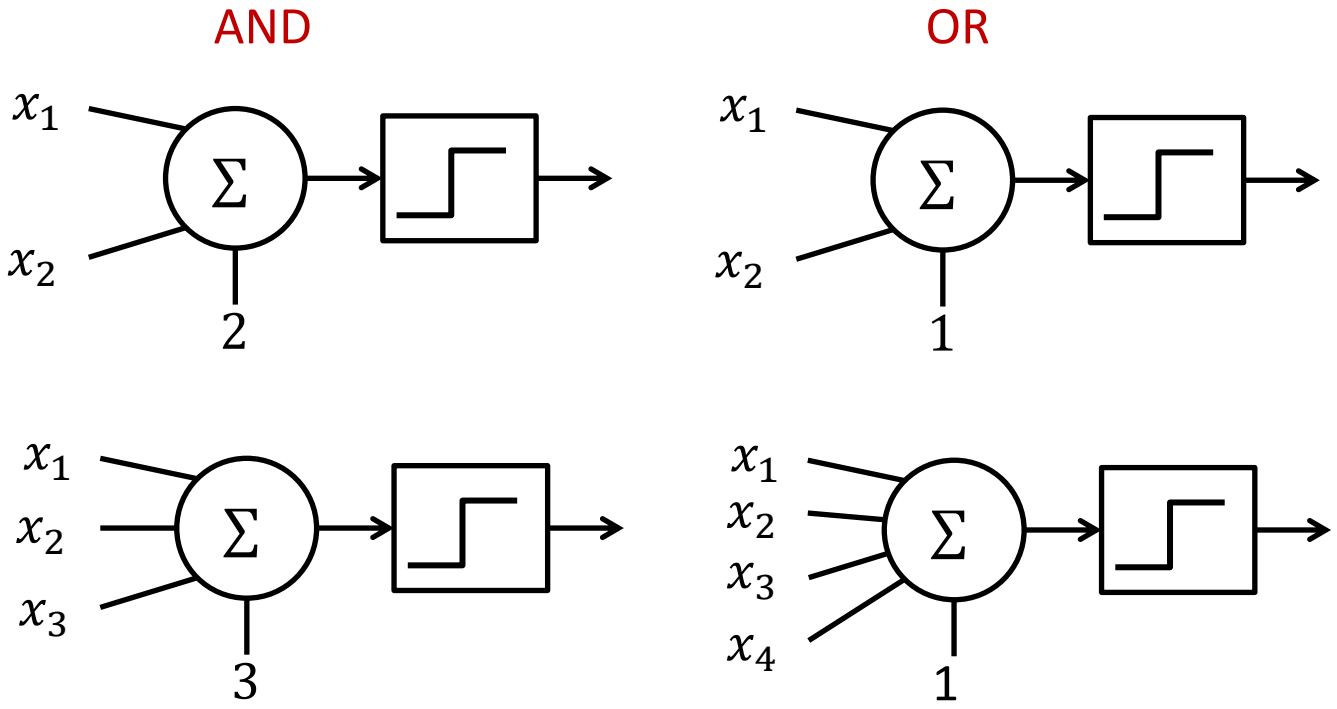
B OR (disjunction)

Submit

Synthesis of Boolean functions

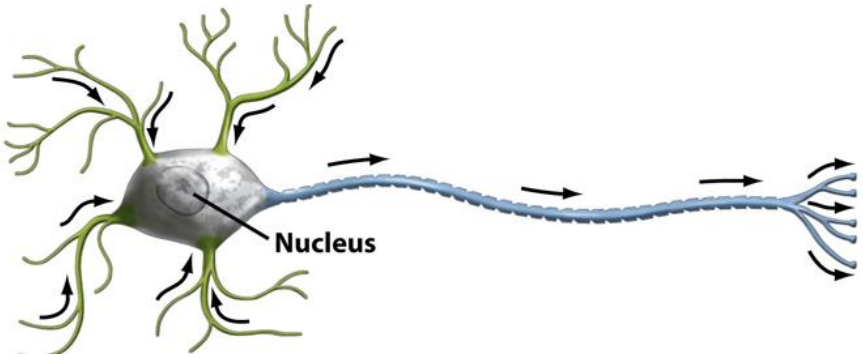
Boolean function: $\{0,1\}^n \rightarrow \{0,1\}$

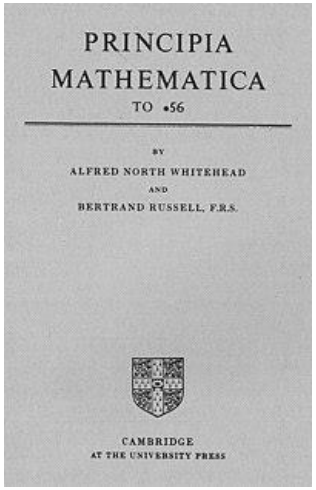
- Conjunction, disjunction, negation



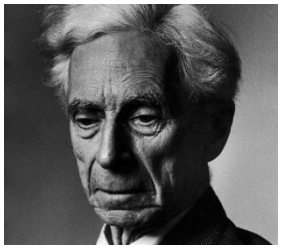
Can you implement negation using the M-P units?

What's the motivation for M-P units?

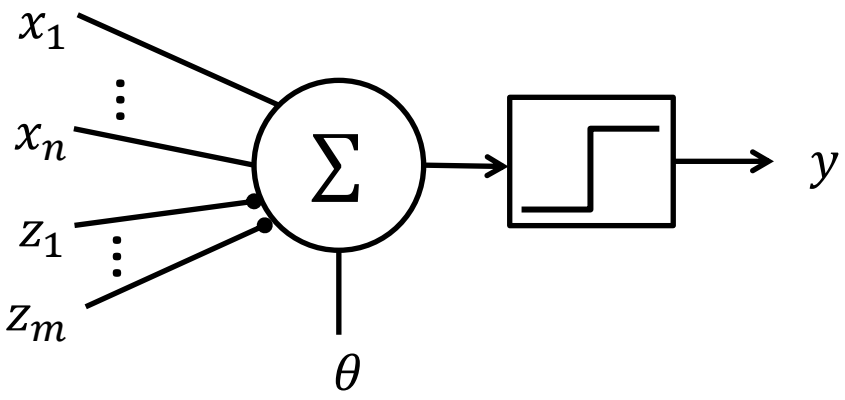


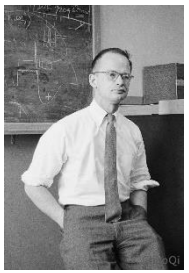


Alfred Whitehead
& Bertrand Russell

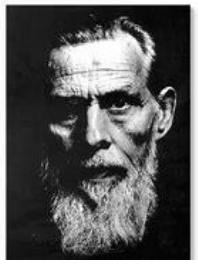


伯特兰·罗素
1872-1970
Nobel Prize in
Literature (1950)





Pitts
1923-1969

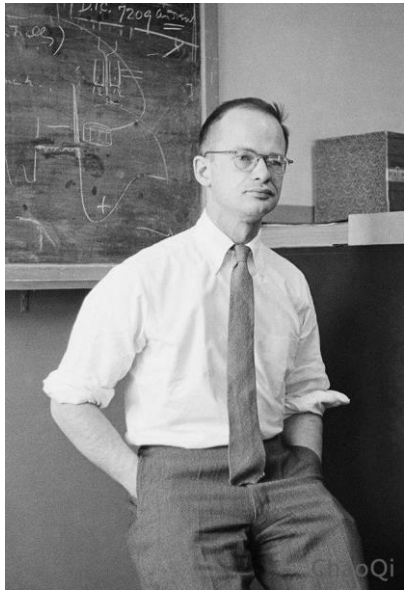


McColloch
1898-1969

Brain=pure
logic

Walter Pitts 1923-1969

- Born in a tough family in Prohibition-era Detroit, where his father, a boiler-maker, had no trouble raising his fists to get his way
- In 1935, he read *Principia Mathematica*, which attempted to reduce all of mathematics to pure logic
- He found several mistakes and wrote to Russel

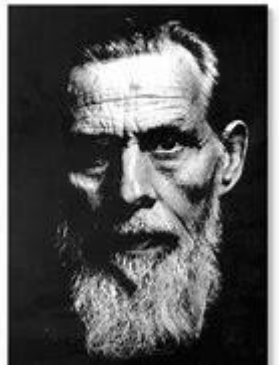


1923-1969

- In 1938, when he heard that Russell would be visiting the University of Chicago, he ran away from home and headed for Illinois. He never saw his family again

Work with Warren McCulloch

- In 1923, the year that Walter Pitts was born, a 25-year-old Warren McCulloch was also digesting the Principia
- McCulloch was born into a well-to-do East Coast family of lawyers, doctors, theologians, and engineers



1898-1969

- They would create the first mechanistic theory of the mind, the first computational approach to neuroscience, the logical design of modern computers, and the pillars of artificial intelligence
- *A Logical Calculus of Ideas Immanent in Nervous Activity*, Bulletin of Mathematical Biophysics, 1943

Work with Norbert Wienner

- In 1943, Pitts became a PhD student of Wienner at MIT
- Wienner realized that it ought to be possible for Pitts' neural networks to be implemented in man-made machines, ushering in his dream of a cybernetic revolution

Norbert Wienner

von Neumann

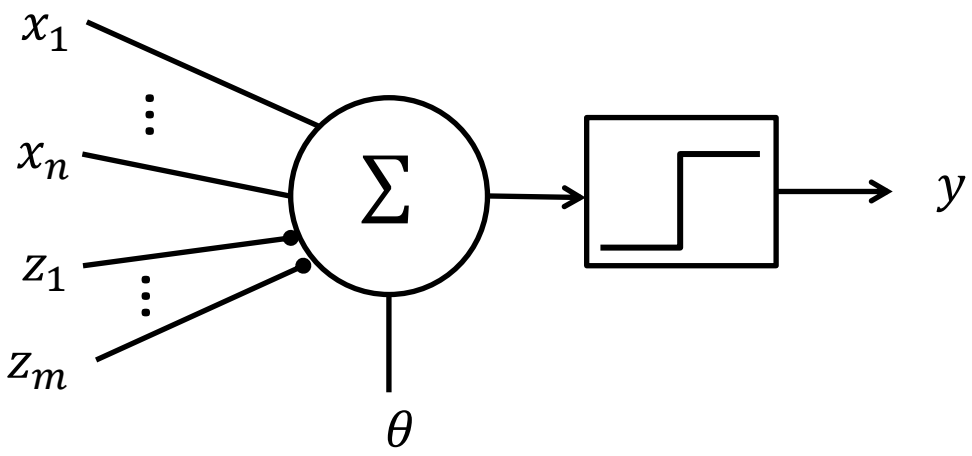
- The beginnings of the group who would become known as the *cybernetic*ians was formed, with Wiener, Pitts, McCulloch, Lettvin, and von Neumann its core.
- von Neumann suggested modeling the computer after Pitts and McCulloch's neural networks

Collapse of logical brain idea

- Wiener's wife invented a story. She said their daughter, Barbara, had been seduced by several of "his boys"
- Wiener never spoke to Pitts again.
- Experiments with frog's eyes. "The eye speaks to the brain in a language already highly organized and interpreted," they reported in the now-seminal paper "What the Frog's Eye Tells the Frog's Brain," published in 1959
- The results shook Pitts' worldview to its core

- In 1969 Pitts died alone in a boarding house in Cambridge. Four months later, McCulloch passed away

Can we modify M-P units to process real valued inputs?

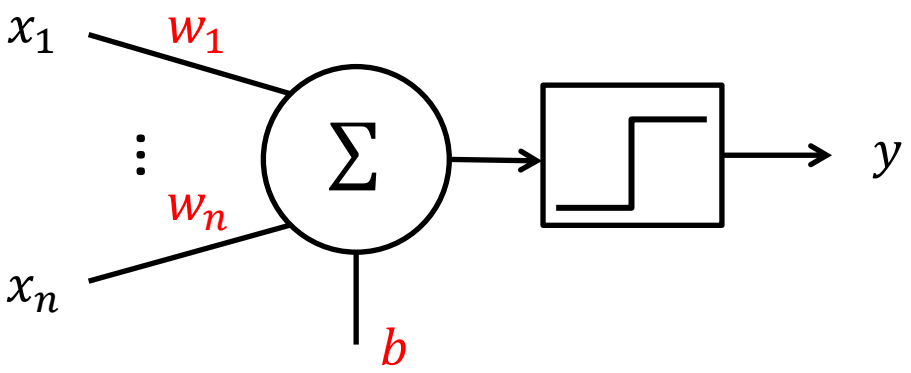


Frank Rosenblatt
1928-1971

How?

The first rise of ANN

Perceptron

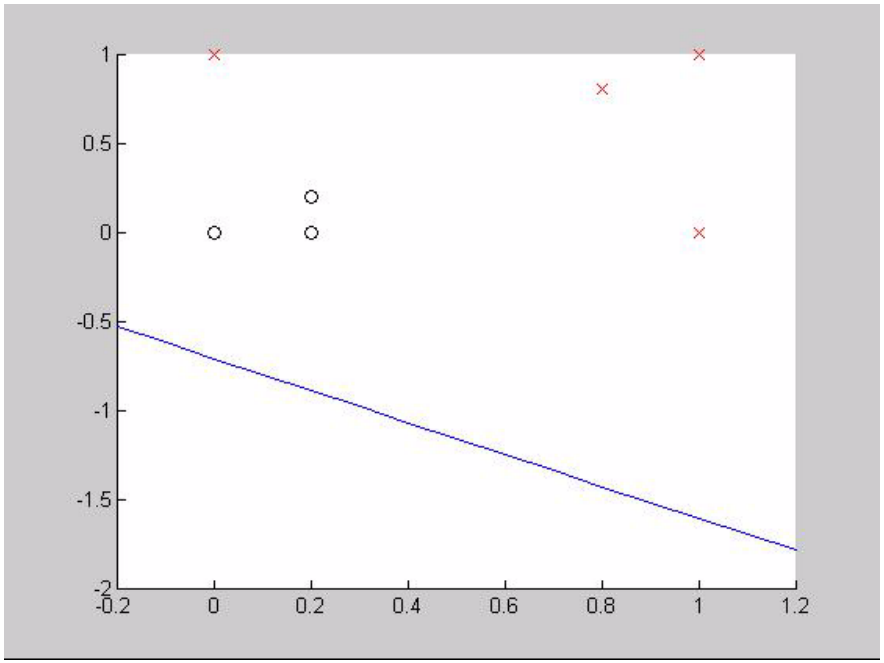


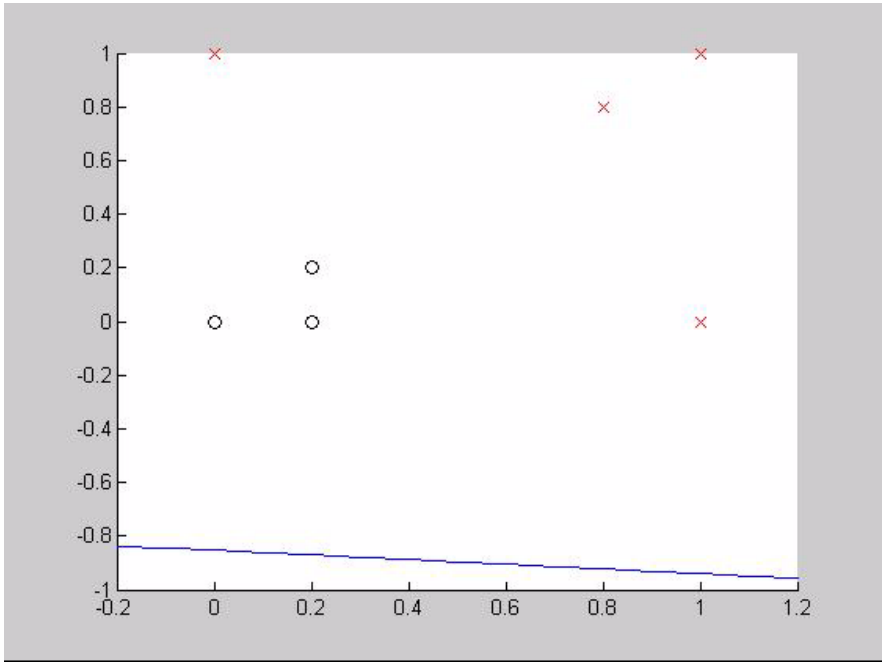
$$y = \begin{cases} 1 & \boldsymbol{w}^\top \boldsymbol{x} + b > 0 \\ 0 & \text{else} \end{cases}$$

Could also be -1

- Add weights to the input connections of the M-P unit
- Propose a supervised learning algorithm: For each data points $\boldsymbol{x}^{(j)} \in R^m$ and the corresponding labels $t^{(j)}$
  - Calculate the actual output $y^{(j)}$
  - Update the weights: $\boldsymbol{w}^{\text{new}} = \boldsymbol{w}^{old} + \eta\left(t^{(j)} - y^{(j)}\right)\boldsymbol{x}^{(j)}$;
    $$b^{\text{new}} = b^{old} + \eta\left(t^{(j)} - y^{(j)}\right)$$
    where $\eta > 0$ is the learning rate

Example



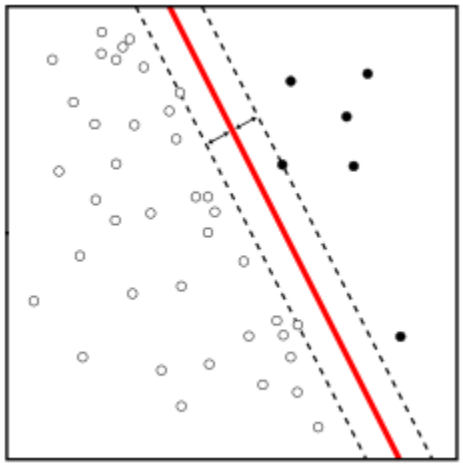


From two different sets of initial weights

Convergence

Proposition. *If the training set is linearly separable, then the perceptron is guaranteed to converge. Furthermore, there is an upper bound on the number of times the perceptron will adjust its weights during the training.*

Proof. See (Novikoff, 1962)



linearly separable

Frank Rosenblatt 1928 -1971

- Bronx High School of Science
- Cornell student (1946 –1956)
- Cornell Aeronautical Laboratories
- Cognitive Systems Research Program
- Neurobiology

1950 Social Psychology

- *political campaigns in NY, NH, VT, CA*
- *music (piano, composition)*
- *astronomy and cosmology*
- *mountain climbing and sailing*

The gravestone of Frank Rosenblatt, Brooktondale, NY.

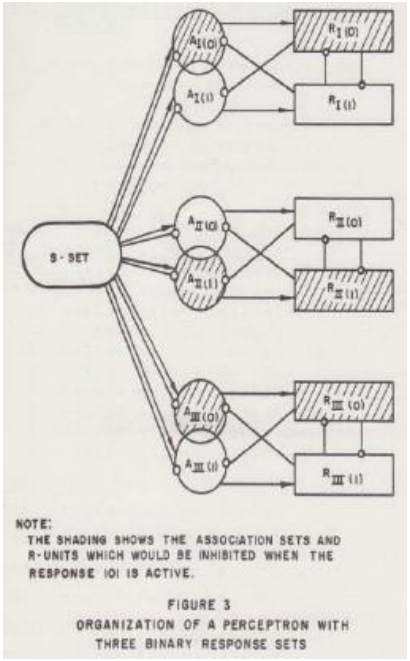
By George Nagy in 2011, Rosenblatt‘s PhD student

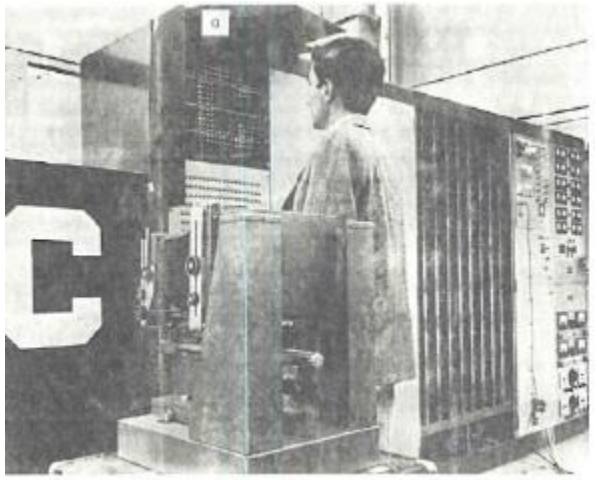
The quest

- How is information about the physical world sensed by the biological system?
- In what form is information stored and retrieved?
- How does remembered information influence recognition and behavior?

By George Nagy in 2011, Rosenblatt‘s PhD student

Perceptron

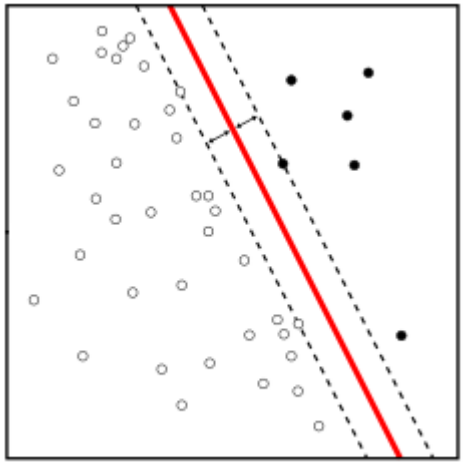




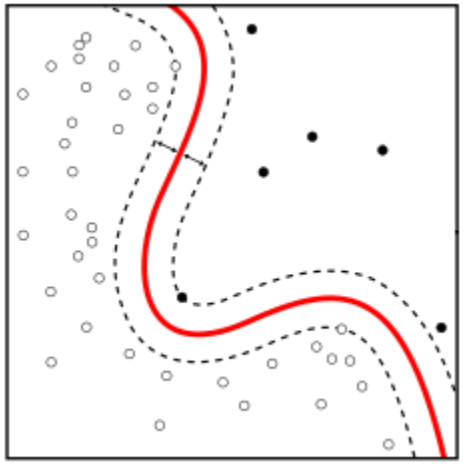
Mark I: 400 S-units, 512 A-units, 8 R-units

By George Nagy in 2011, Rosenblatt‘s PhD student

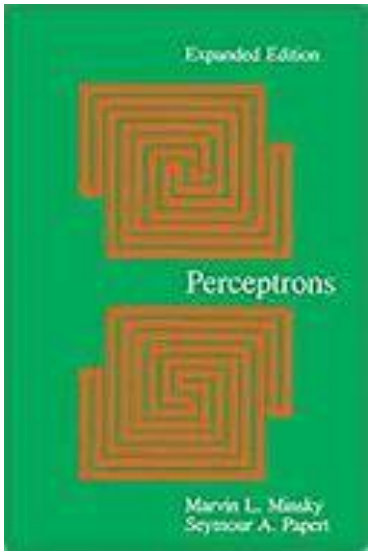
Perceptron's limit



linearly separable



linearly non-separable



Marvin Minsky and
Seymour Papert, 1969

The first fall of ANN

The 2nd rise and fall

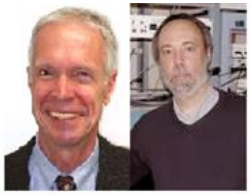

Pitts McCulloch


Minsky Papert


Rosenblatt


Widrow


Werbos


Hopfield Tank


Rumelhart


Hinton


LeCun


Bengio

MLP and Hopfield network

SVM

Introduced in ML-related courses

1943 ~1960 1969 1974 1982 1986 1990s 2006

The 3rd rise

Geoffrey Hinton Yann LeCun Yoshua Bengio Jürgen Schmidhuber

Established in 1982, CIFAR is a Canadian-based, international research institute with nearly 400 fellows, scholars and advisors from 18 countries.

Hinton’s interview by Ng

单选题 1分

设置

Which model has real valued weights?

A MP unit

B Perceptron

提交

单选题 1分

设置

What’ s the most serious problem of the Perceptron?

A It doesn’ t have weights on connections

B The convergence is not guaranteed for linearly separable data

C It cannot separate data that are linearly non-separable

提交

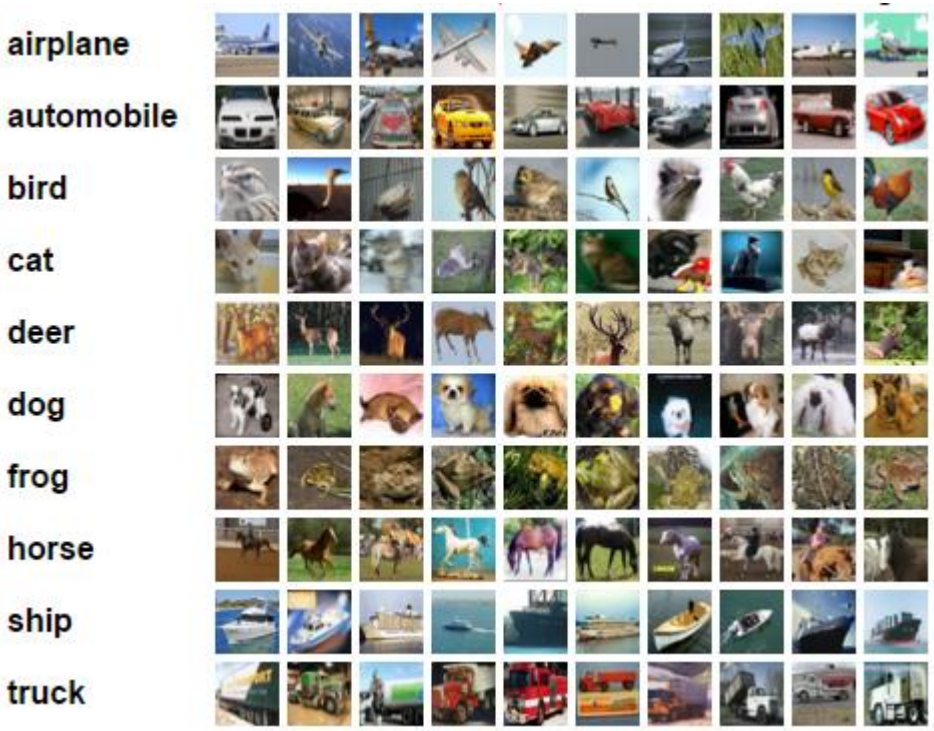
Outline

I. General concepts
II. History
III. Applications
IV. Summary

General object classification

CIFAR-10 & CIFAR100 datasets

- 50,000 training, 10,000 test
- 32x32 RGB imgs



ILSVRC2012 dataset

- ~128M training
- 50,000 validation
- 100,000 test

Specific object classification

Face verification

Coo d'Este Melina Kanakaredes

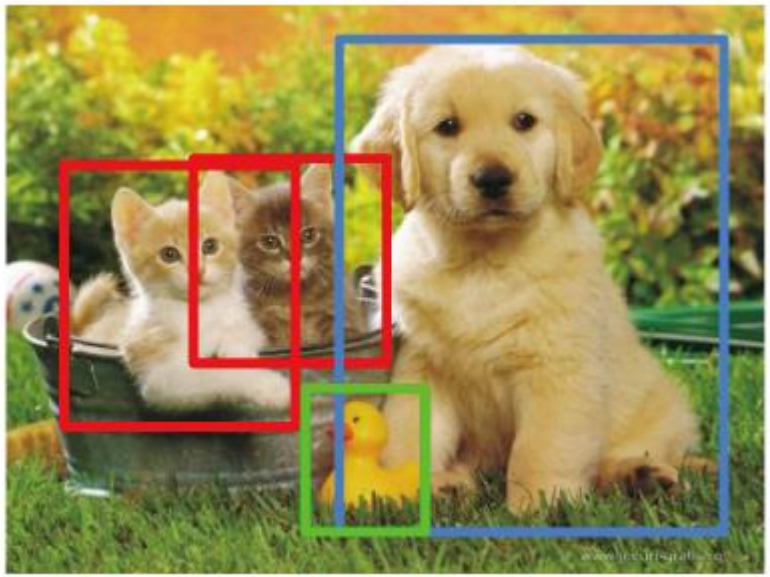
Elijah Wood Stefano Gabbana

Jim O'Brien Jim O'Brien

Model	Accuracy (%)
DeepFace (2014)	97.25
DeepID (2014)	97.45
DeepID2 (2014)	99.15
DeepID2+ (2014)	99.47
DeepID3 (2014)	99.53
FaceNet (2015)	99.63

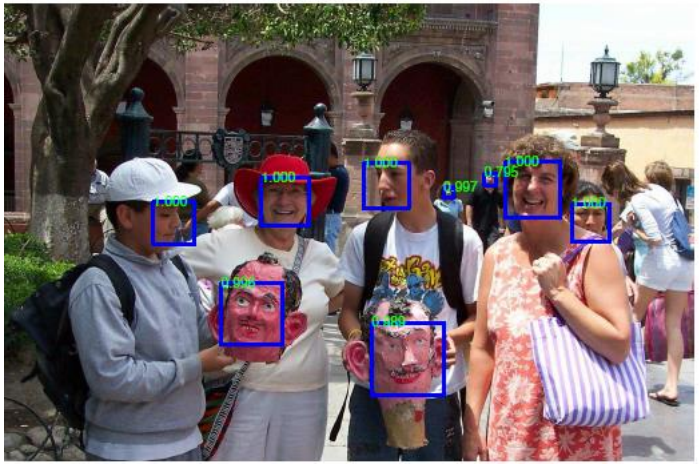
Pig face recognition

General object detection

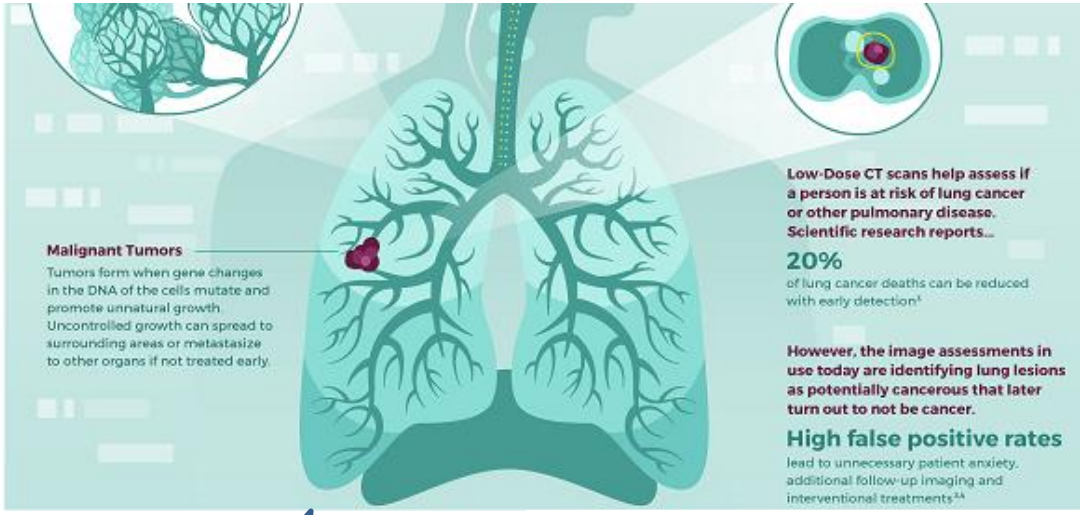


DOG, (x, y, w, h)
CAT, (x, y, w, h)
CAT, (x, y, w, h)
DUCK (x, y, w, h)

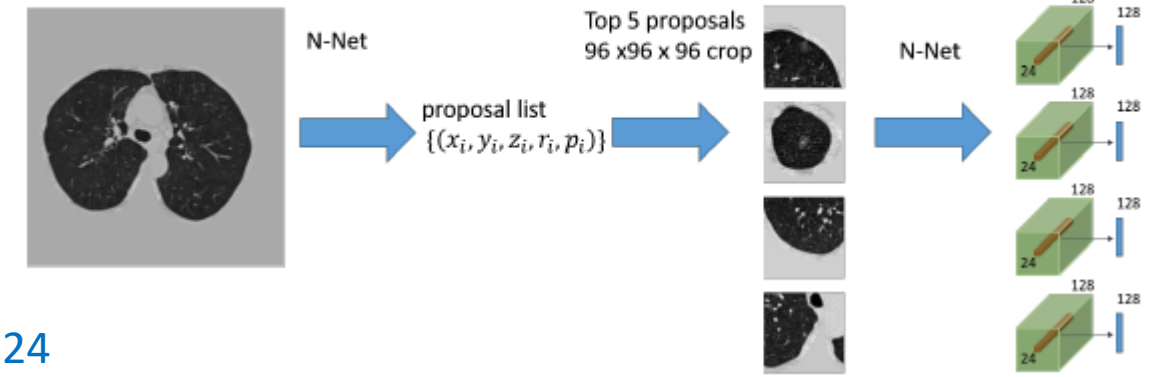
Specific object detection



Medical image analysis



A 500,000 USD solution!



Liao et al., arXiv:1711.08324

Image generation

64*64

Generated bedrooms after five epochs of training of a GAN

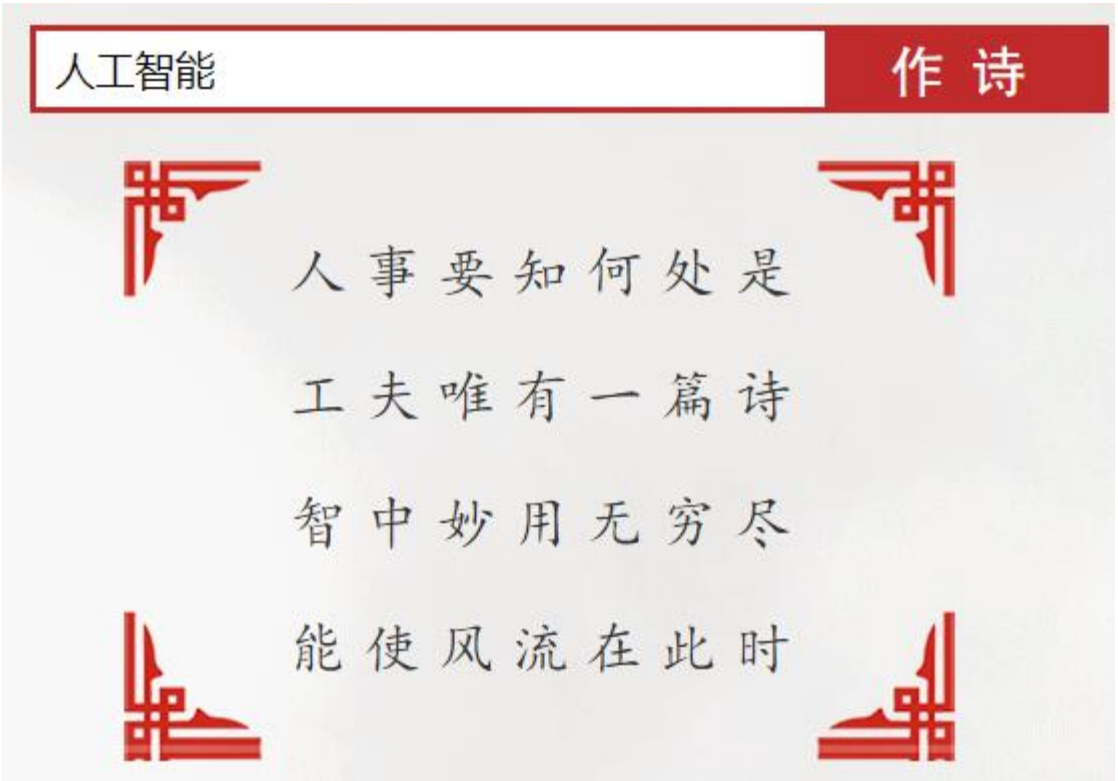
Controllable image generation

’An image of an animal half mouse half octopus’

’A painting of a squirrel eating a burger’

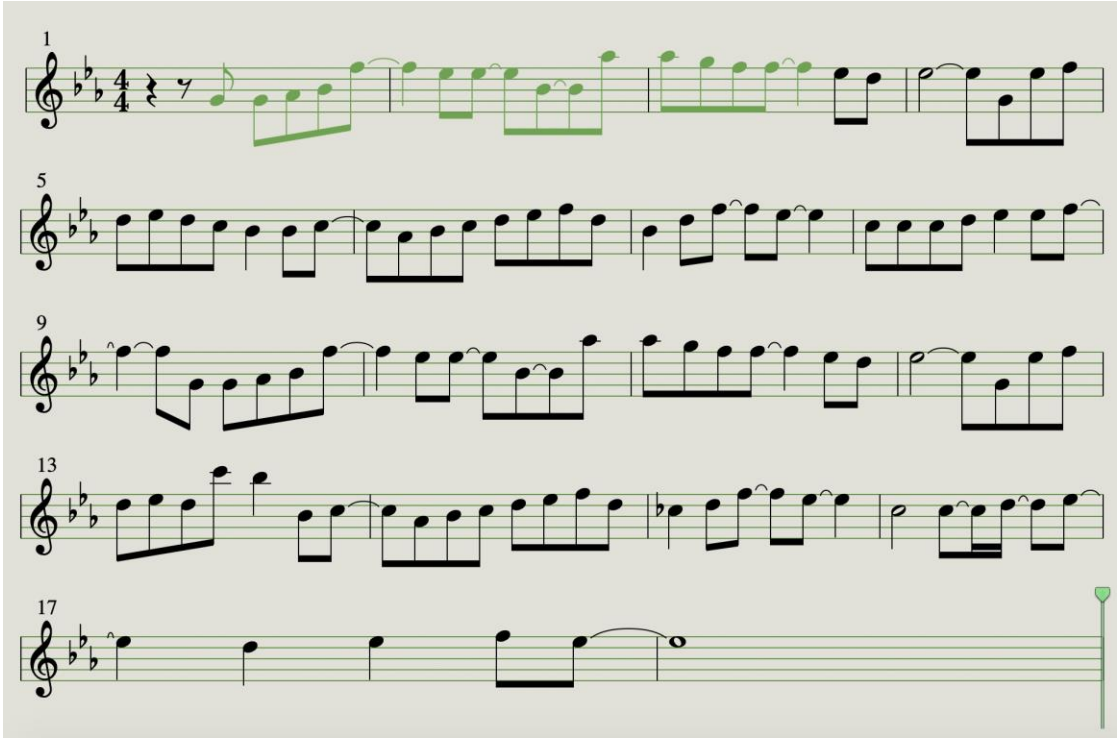
’A shirt with the inscription: “I love generative models!” ’

Chinese poem generation



http://jiuge.thunlp.org/

Music generation



旋律+和弦

旋律+AI钢琴伴奏

https://demo.lazycomposer.com/compose/pianoSoloExpert/

Alphafold2

DEEPMIND'S AI PREDICTS STRUCTURES FOR A VAST TROVE OF PROTEINS

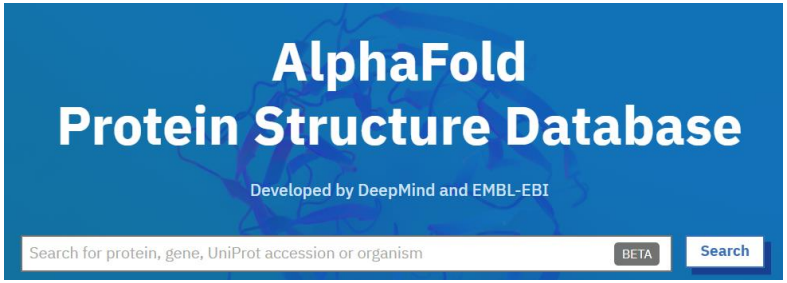
AlphaFold neural network produced 'transformative' database of more than 350,000 structures.

By Ewen Callaway

dump is a beginning, not an end. They will

proteome predictions (K. Tunyasuvunakool *et al. Nature* https://doi.org/gk9kp7; 2021). For the human proteome, 58% of AlphaFold's predictions for the locations of individual amino acids were good enough to be confident in the shape of the protein's folds, Tunyasuvunakool says. A subset of those predictions – 36% of the total – are potentially precise enough to detail atomic features useful for drug design, such as the active site of an enzyme.

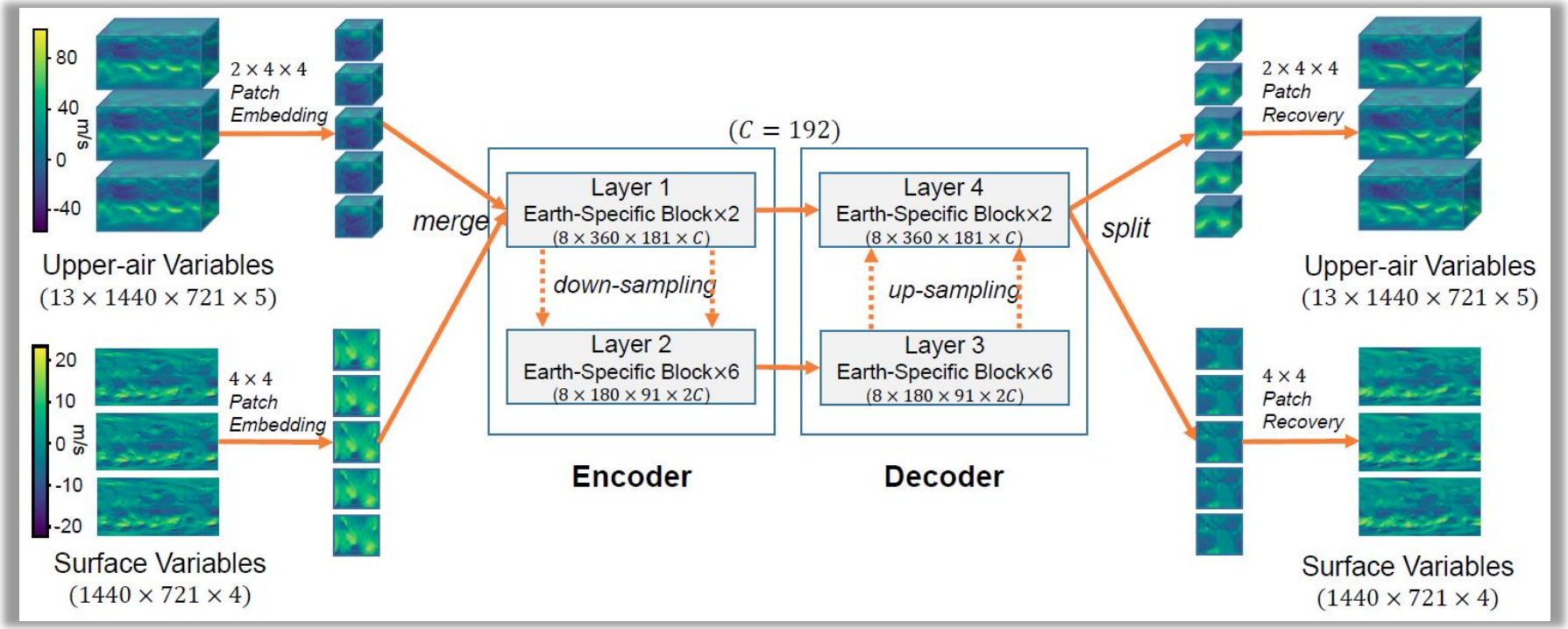
Even the less-accurate predictions might offer insights. Biologists think that a large proportion of human proteins and those of other



Nature, July 2021

Weather forcasting

盘古气象大模型



中长期气象预报精度首次超过传统数值方法，速度提升10000倍以上

What interesting applications do you know?

AI Spots Mysterious Signals Coming from Deep in Space

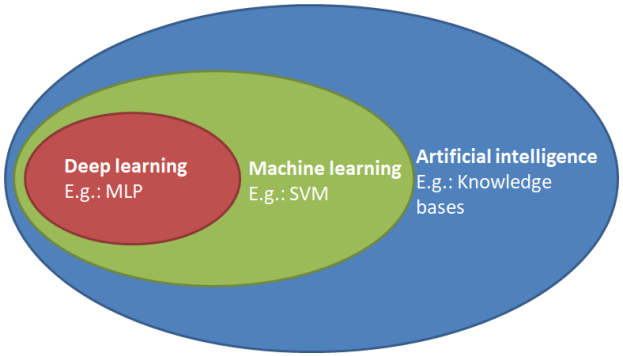
Outline

I. General concepts
II. History
III. Applications
IV. Summary

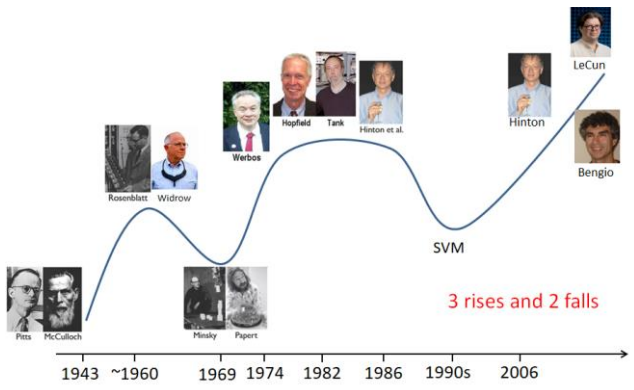
Summary of this lecture

Knowledge

I. General concepts



II. History



III. Applications

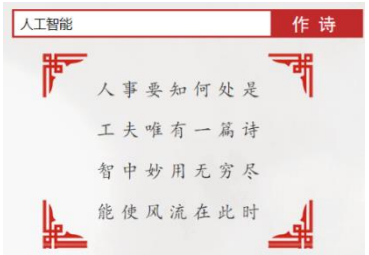
Computer vision

Biology

Natural language processing

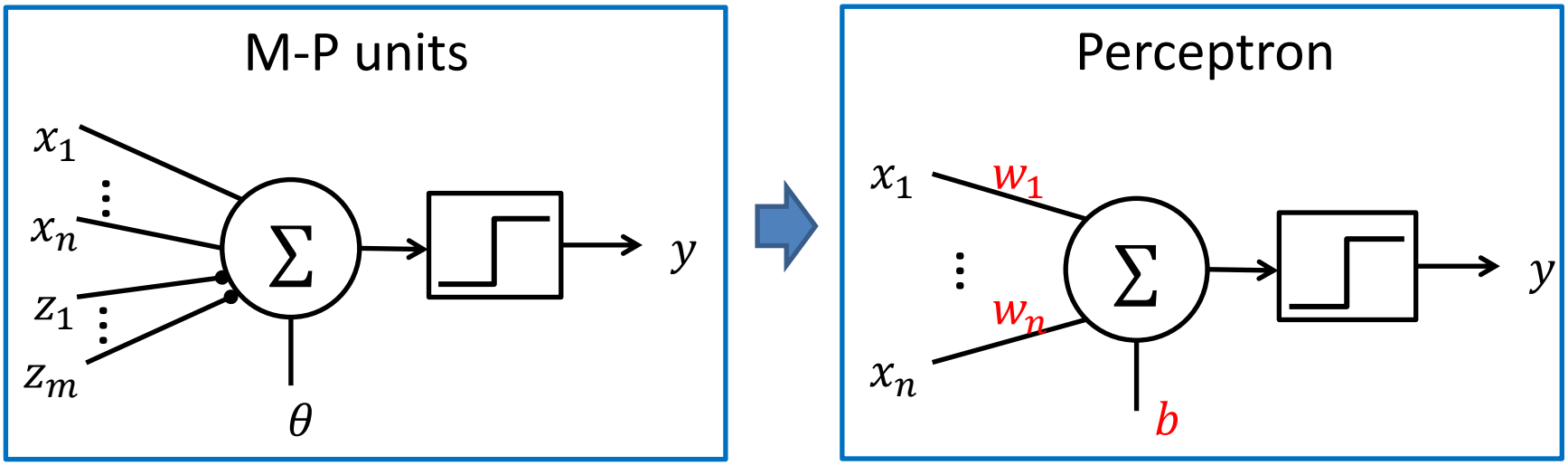
Speech recognition

Music



Summary of this lecture

Capability and value



- 对科学的执着信念
- 尊重事实，转换思路：既然大脑不是一个逻辑运算器，那就改模型

Recommended reading

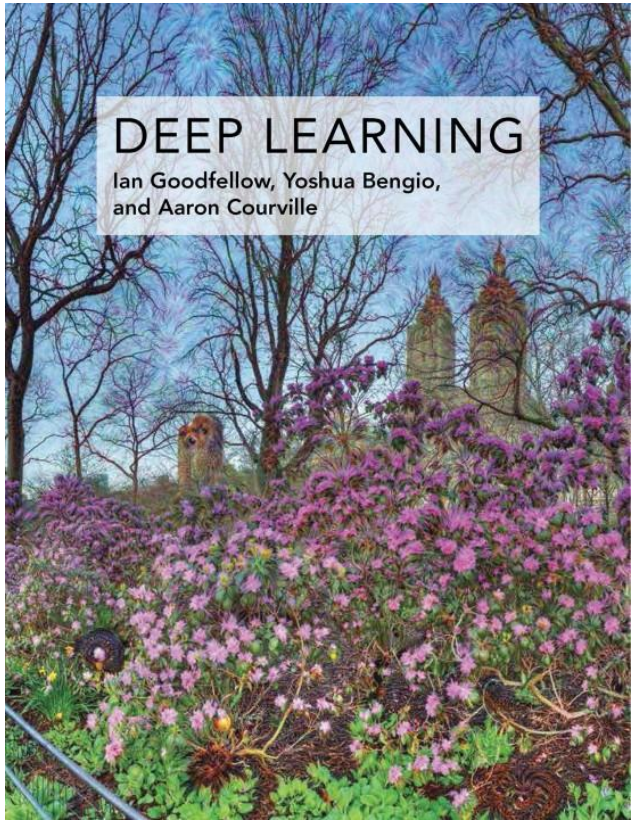
- Walter Pitts: The Man Who Tried to Redeem the World with Logic

  http://nautil.us/issue/21/information/the-man-who-tried-to-redeem-the-world-with-logic
- 杨晓凡 AI科技评论2018-04-08

  吴恩达专访LeCun：即便在神经网络的寒冬，我也坚信它终会重回公众视野

  https://www.leiphone.com/news/201804/fGJ32alQQVnBneWJ.html

Prepare for the next lecture



Deep Learning
Ian Goodfellow, Yoshua Bengio and Aaron Courville
The MIT Press, 2018

https://github.com/janishar/mit-deep-learning-book-pdf

Chapters 2-5