

Quick Introduction

Neural Networks Theory

Yungcheol BYUN

ycb@jejunu.ac.kr

Machine Learning Lab@CE, JNU

github.com/yungbyun/neuralnetworks

Neural Networks?



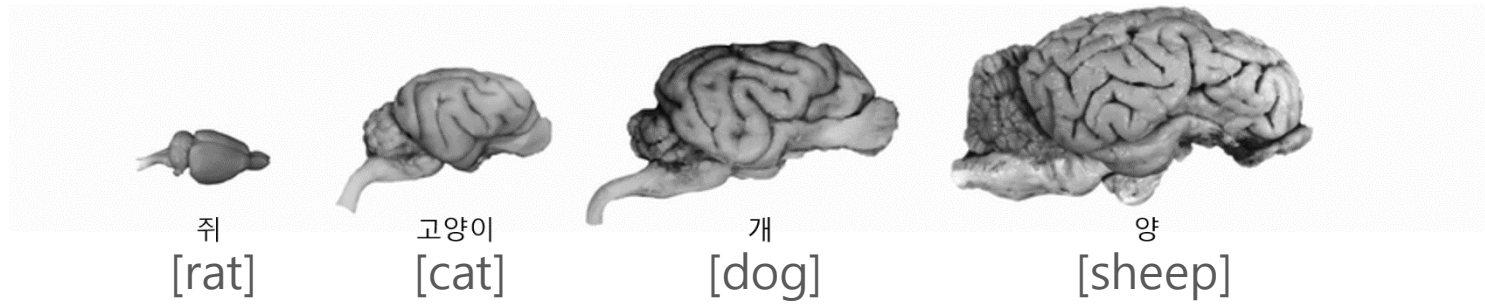
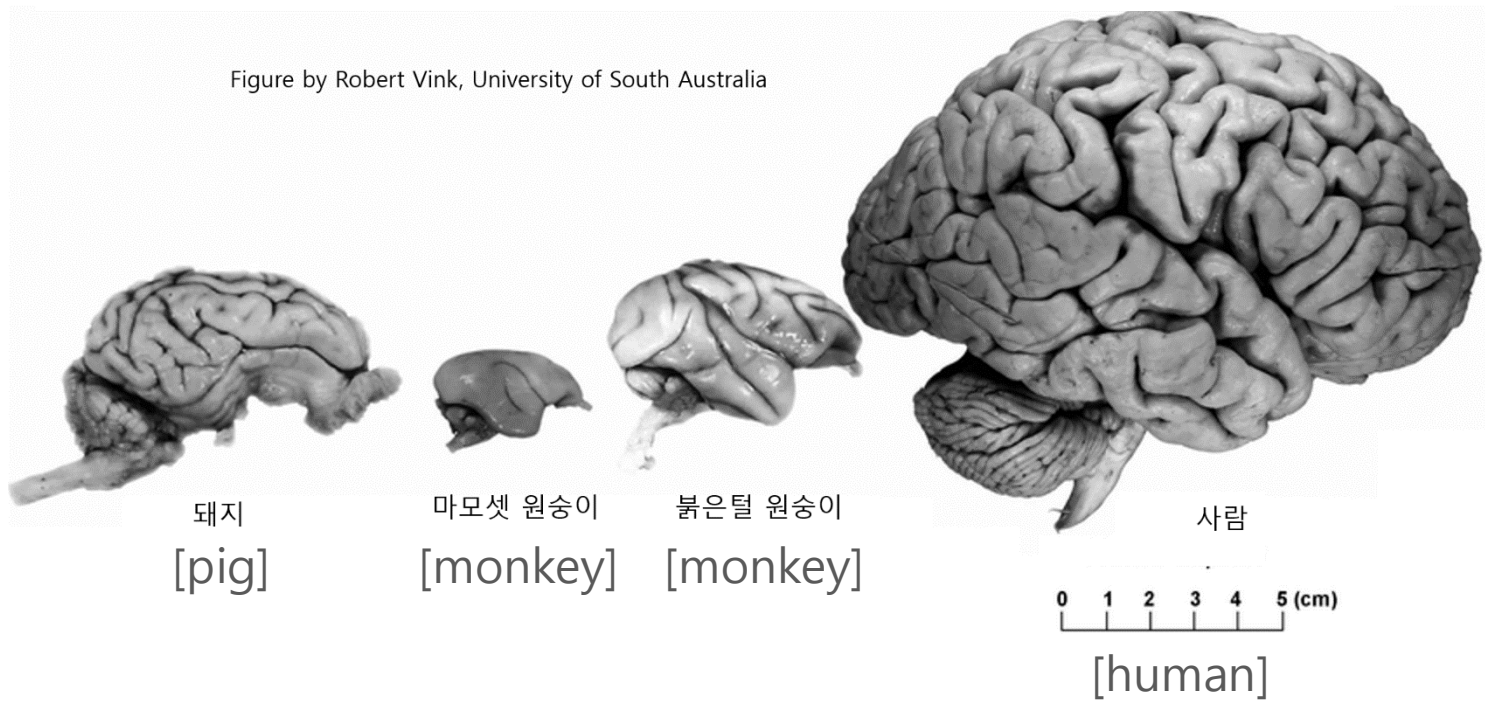


Figure by Robert Vink, University of South Australia





So, neural networks is ...

뇌에 있는 신경세포의 연결

Artificial Neural Networks

man made, 사람이 만든

"...a computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs."

Frank Rosenblatt, Cornell Aeronautical Lab
(1957, *66 years ago*)

Why? What for?

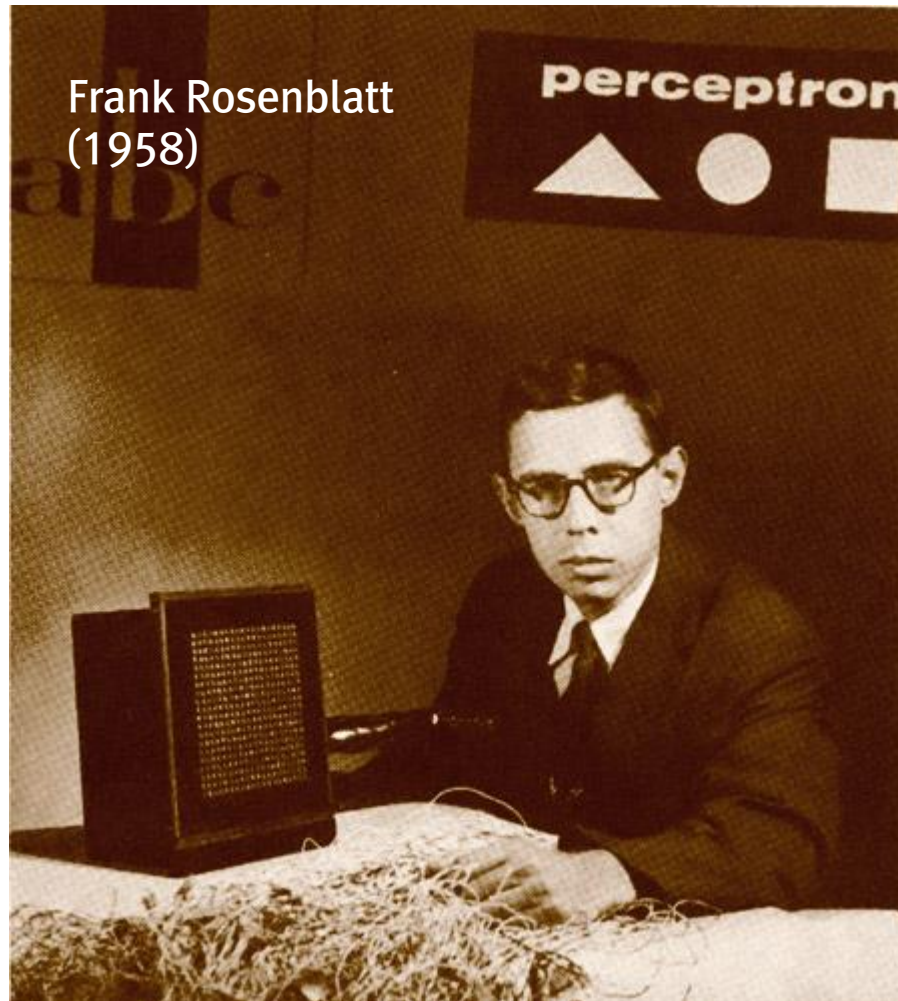
Intelligence

The ability to **perceive** or infer information, and to **retain** it as knowledge to **be applied towards** adaptive behaviors within an environment or context.

Artificial Neural Networks for

Artificial Intelligence

사람이 만든 지능



Rosenblatt (1958). The **perceptron**: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408. <https://doi.org/10.1037/h0042519>



[Journal Information](#)
[Journal TOC](#)

[Search APA PsycNet](#)

APA PsycArticles: Journal Article

The perceptron: A probabilistic model for information storage and organization in the brain.

[© Request Permissions](#)

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408.

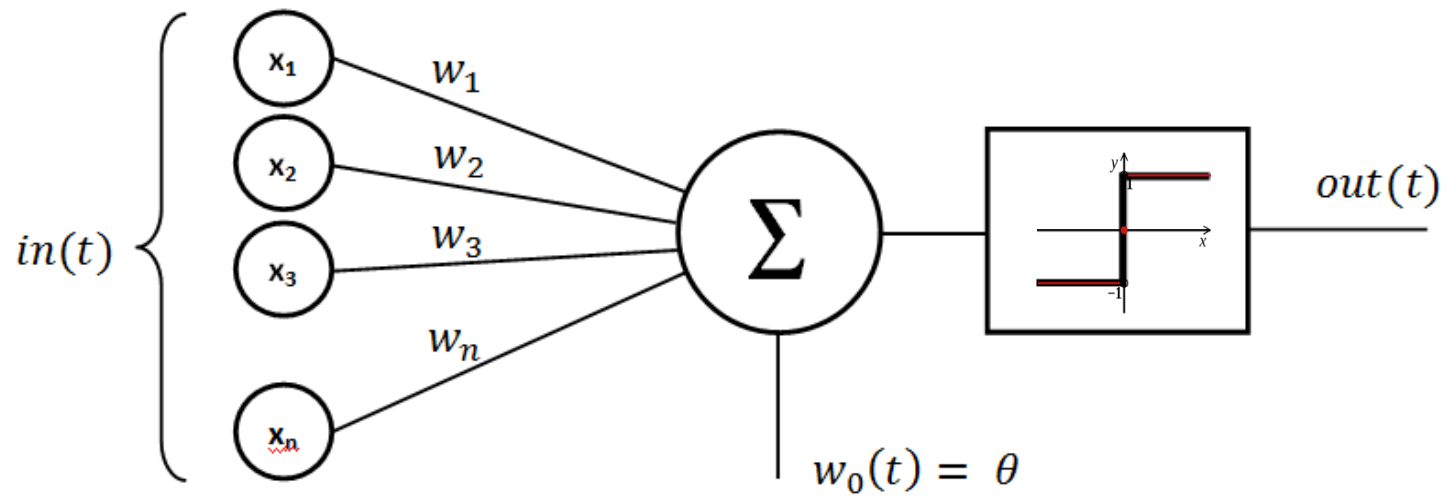
<https://doi.org/10.1037/h0042519>

To answer the questions of how information about the physical world is sensed, in what form is information remembered, and how does information retained in memory influence recognition and behavior, a theory is developed for a hypothetical nervous system called a perceptron. The theory serves as a bridge between biophysics and psychology. It is possible to predict learning curves from neurological variables and vice versa. The quantitative statistical approach is fruitful in the understanding of the organization of cognitive systems. 18 references. (APA PsycInfo Database Record (c) 2016 APA, all rights reserved)



[Get Access](#)

Perceptron



ANN and **rosy** prospects in 1950s

ANN에 대한 장미빛 전망

NEW NAVY DEVICE LEARNS BY DOING; Psychologist Shows Embryo of Computer Designed to Read and Grow Wiser

July 8, 1958



See the article in its original context from
July 8, 1958, Page 25 | [Buy Reprints](#)

New York Times subscribers* enjoy full access to
TimesMachine—view over 150 years of New
York Times journalism, as it originally appeared.

SUBSCRIBE

*Does not include Crossword-only or
Cooking-only subscribers.

The New York Times Archives

Rosy Period → 1차 침체기

1st Period of Depression in 1960s

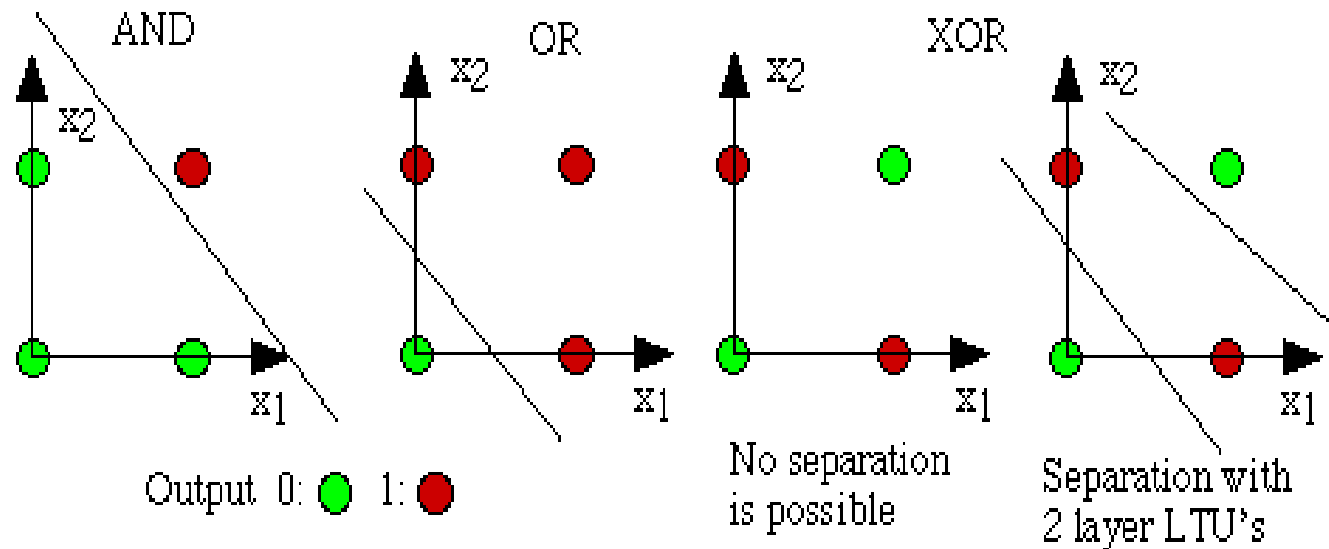


by **Marvin Minsky**, MIT AI Lab.

in 1969

with Perceptron

No solution for XOR problem

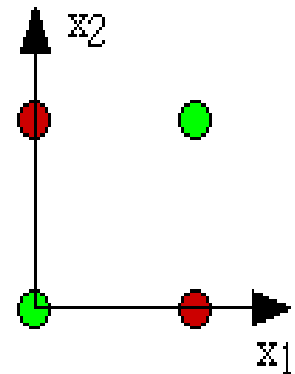


“17 years later”

Multi-Layer Perceptron

by Rumelhart, Hinton, Williams in 1986

“Rosy Period again!”



2nd Rosy Period → 2차 침체기

2nd Period of Depression in 1990s

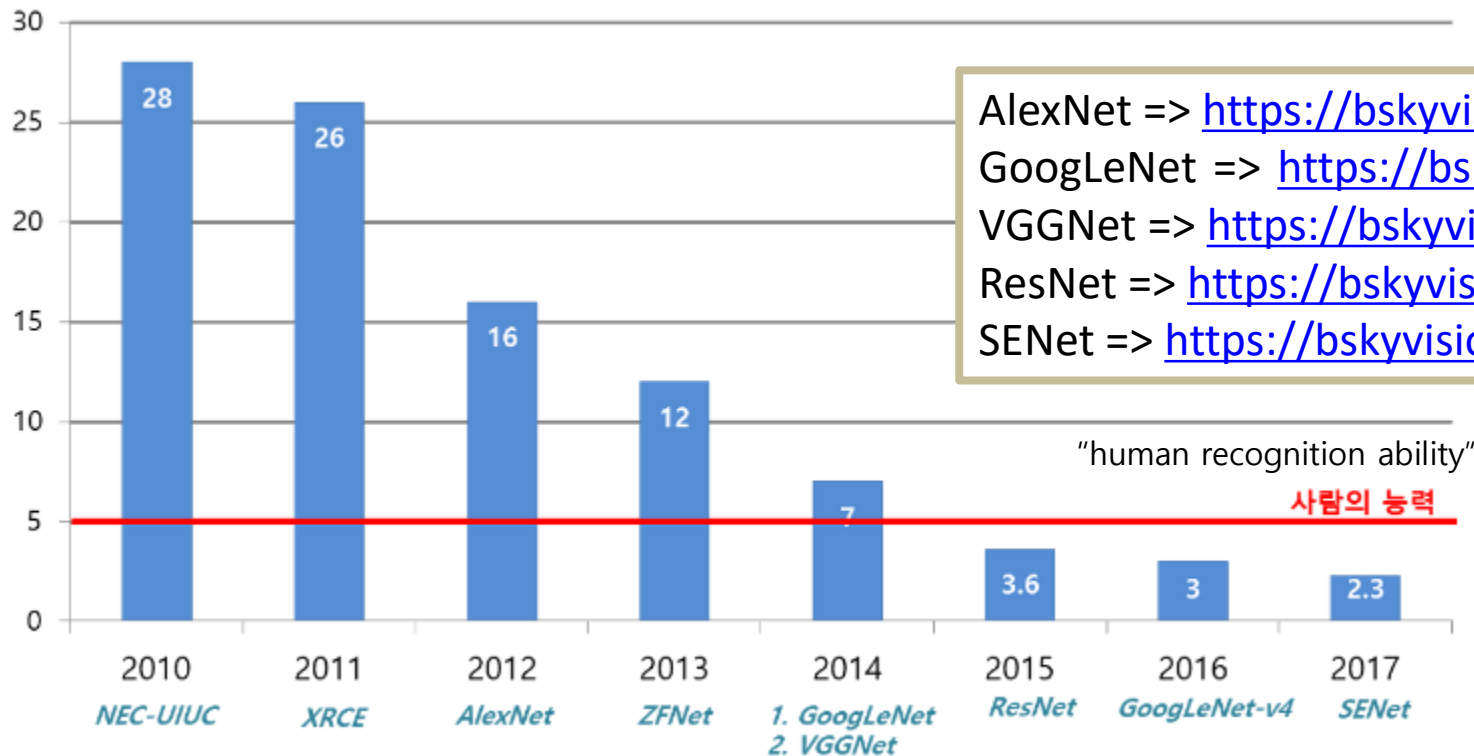
- Difficulties in training the MLP
(sigmoid function)
- Not enough data
- Long learning time
- Overfitting
- etc.

Deep Learning in 2000s

3rd Rosy Period



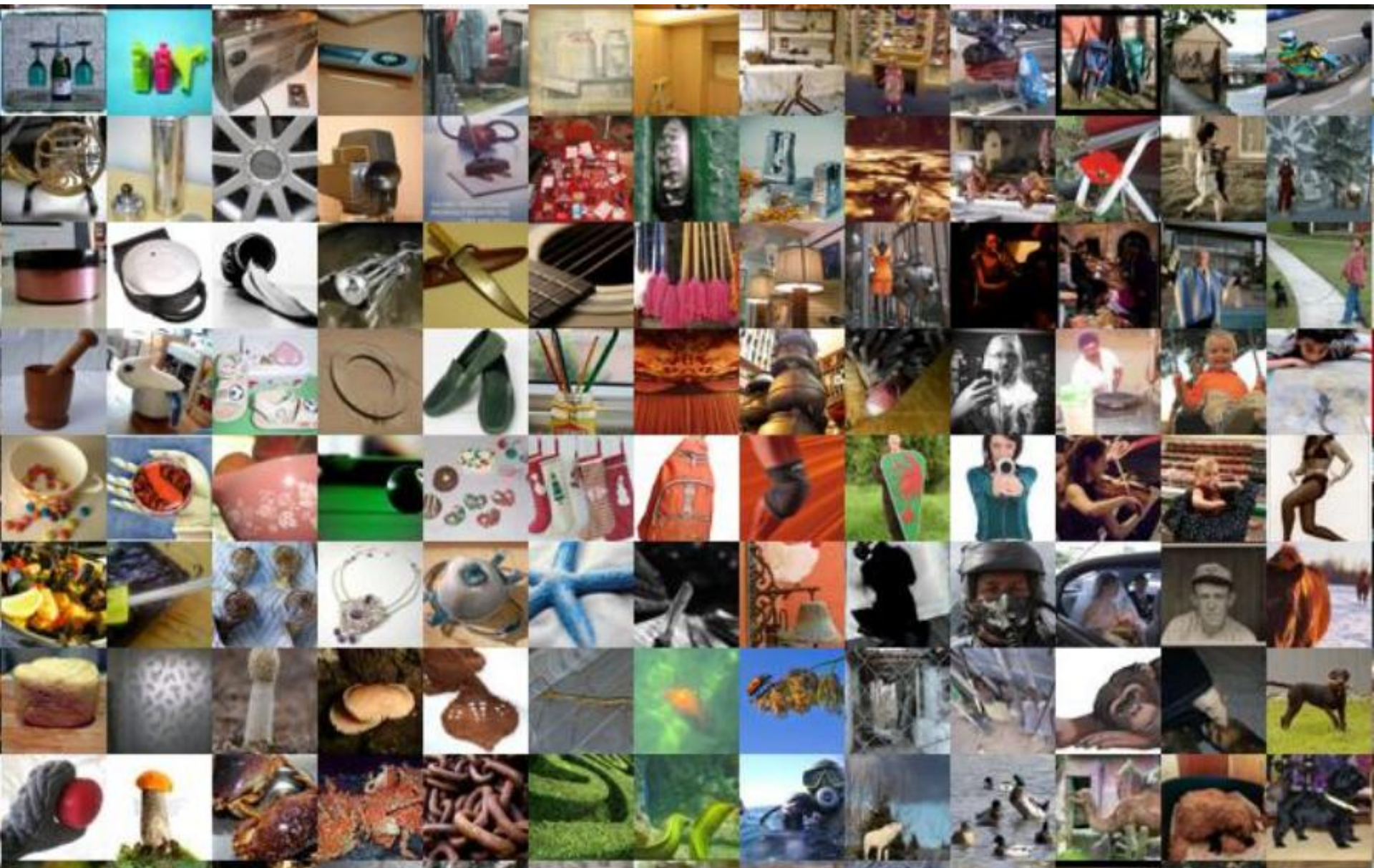
Error (%)



AlexNet => <https://bskyvision.com/421>
GoogLeNet => <https://bskyvision.com/539>
VGGNet => <https://bskyvision.com/504>
ResNet => <https://bskyvision.com/644>
SENet => <https://bskyvision.com/640>

"human recognition ability"

사람의 능력





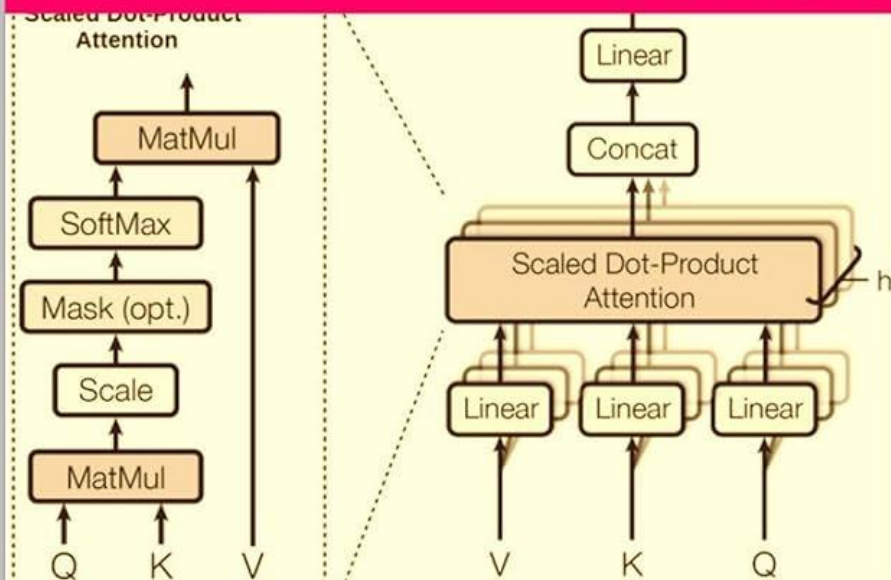
AlphaGo

2016

Attention is All You Need

The Game-Changing Paper That Transformed NLP

HENRI VAN MAARSEVEEN



Attention Is All You Need

Ashish Vaswani*
Google Brain
avaswani@google.com

Noam Shazeer*
Google Brain
noam@google.com

Niki Parmar*
Google Research
nikip@google.com

Jakob Uszkoreit*
Google Research
usz@google.com

Llion Jones*
Google Research
llion@google.com

Aidan N. Gomez* †
University of Toronto
aidan@cs.toronto.edu

Lukasz Kaiser*
Google Brain
lukaszkaizer@google.com

Illia Polosukhin* ‡
illia.polosukhin@gmail.com

2017

Transformer

implementing attention mechanism

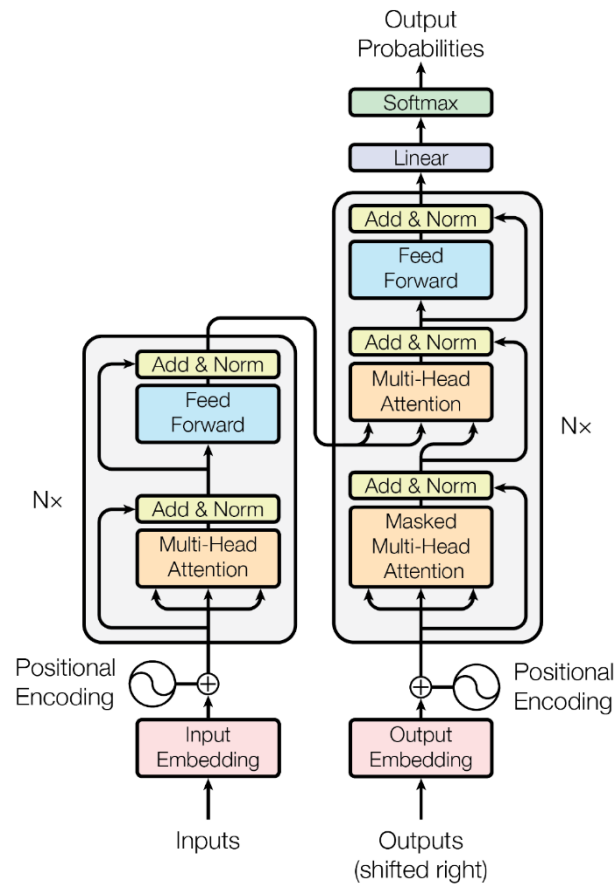


Figure 1: The Transformer - model architecture.

GPT

GPT(Generative Pre-trained Transformer)

- A representative model developed based on the Transformer architecture.
- It primarily utilizes the decoder part of the Transformer to pre-train on vast amounts of text data.
- Then, it is fine-tuned for specific tasks as a generative AI model.

LLMs



deepseek

Physical AI



Let's study the **principles** of
AI, understand **where we**
are, and determine **where**
we need to go.

Schedule

Week	Subject
1	(Course change period)
2	Introduction of the class
3	Self-introduction (students)
4	Brain and Artificial Neural Networks
5	Linear Regression
6	Minimizing Cost
7	Logistic Classification and Softmax #1
8	(Exam period)

Schedule

Week	Subject
9	Multilayer NN
10	Deep Learning
11	Convolutional Neural Network
12	RNN, LSTM, Seq2seq, Transformer
13	Presentation#2
14	Presentation#3
15	Presentation#4

Self-Introduction

(More than 10 slides in English)

- Personal information (Lab., Marital Status, Family Info., etc)
- Nationality/Birthplace/Tourism Place (Photo)
- Hobbies / Specialty
- Education and Experience
- Current Major and Research Interests
- Your dream / future plan
- and others