100 Years The History of Neural Network

Special edition with

<DEV>
MOUNTAIN
TECH FESTIVAL

All About Me Boyd

Sorratat Sirirattanajakarin, Boyd

Senior DS/AI Researcher

@American Express Global Business Travel (30secondstofly)

"Data is a new UX"

Education

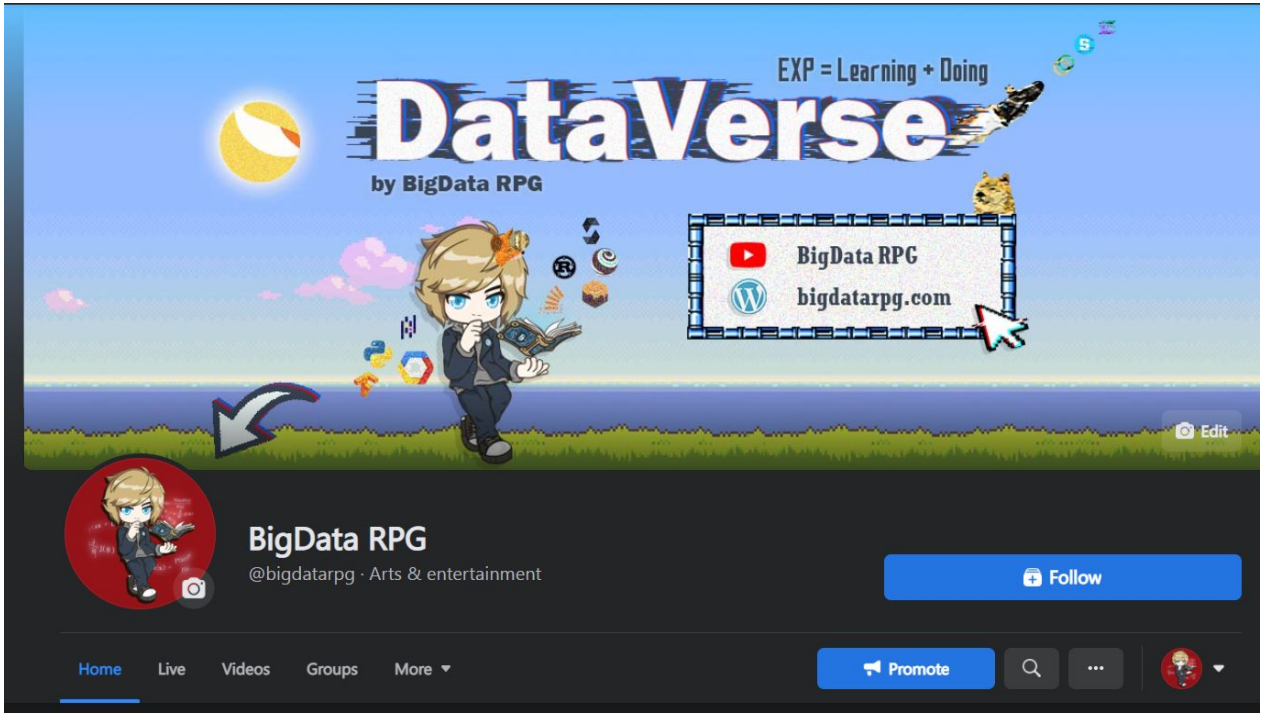
1. Bachelor's degree in **Biotechnology**, A.I., KU. (2010-2013)
2. Master's degree in **Big Data Engineering**, CITE, DPU. (2017-2019)
3. Knowledge Transfer Ambassador **Natural Language Understanding** VISTEC – SCB. (2020-2021)

Work Experiences

1. Food and Beverage Industry: since 2013
2. Digital Marketing: since 2016
3. Software House & Agency: since 2017
4. Bank Industry: since 2018
5. Logistic Supply Chain: since 2021
6. Travel Industry since 2021

Boyd BigData RPG

BigData RPG

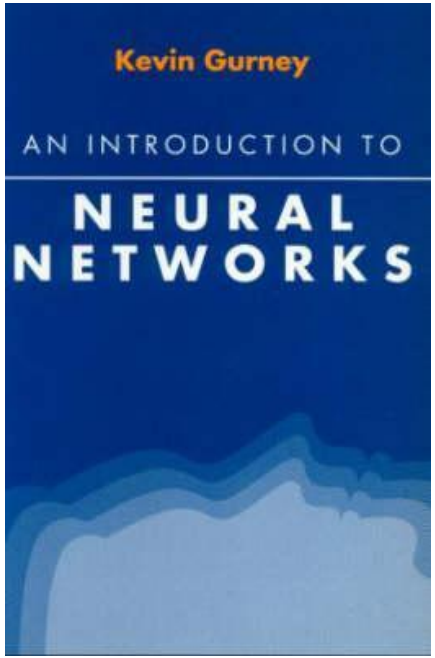


Boyd BigData RPG

PRELIM

Neural networks, something like brains and is potentially laden with the science fiction connotations of the Frankenstein mythos.

By: Kevin Gurney, University of Sheffield



Their name and structure are **inspired by the human brain**, mimicking the way that biological neurons signal to one another.

By: IBM Cloud Education

1950

I, ROBOT

1920-1992 (72)

ISAAC ASIMOV,
AN AMERICAN WRITER AND PROFESSOR OF BIOCHEMISTRY AT BOSTON UNIVERSITY

1. หุ่นยนต์มิอาจกระทำการอันตรายต่อผู้ที่เป็นมนุษย์ หรือนิ่งเฉยปล่อยให้ผู้ที่เป็นมนุษย์ตกอยู่ในอันตรายได้
   (A robot may not harm a human being, or, through inaction, allow a human being to come to harm.)
2. หุ่นยนต์ต้องเชื่อฟังคำสั่งที่ได้รับจากผู้ที่เป็นมนุษย์ เว้นแต่คำสั่งนั้นๆ ขัดแย้งกับกฎข้อแรก
   (A robot must obey the orders given to it by human beings, except where such orders would conflict with the First Law.)
3. หุ่นยนต์ต้องปกป้องสถานะความมีตัวตนของตนไว้ ตราบเท่าที่การกระทำนั้นมิได้ขัดแย้งต่อกฎข้อแรกหรือกฎข้อที่สอง
   (A robot must protect its own existence, as long as such protection does not conflict with the First or Second Law.)

ROBOT !

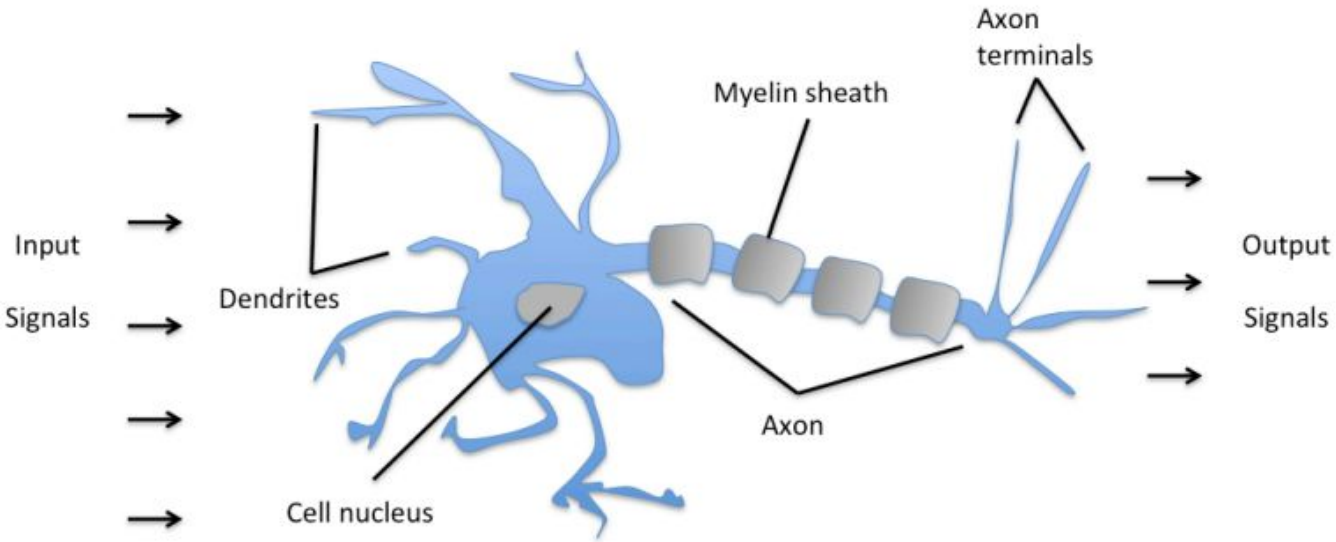

ISAAC ASIMOV'S ROBOT CITY
BESTSELLING AUTHOR OF ROBOTS AND EMPIRE
ISAAC ASIMOV
ISAAC
ROBOTS AND ALIENS
Changeling by Stephen Leigh
Renegade by Cordell Scotte
Intruder by Robert Thurston
Alliance by Jerry Oltion
Book 6 Perihelion by William F. Wu
Maverick by Bruce Bethke
Humanity by Jerry Oltion

How to create an artificial intelligence ?

Trust me I can!!!

Dr. Albert W. Wily,
Scientist, (1900-???)

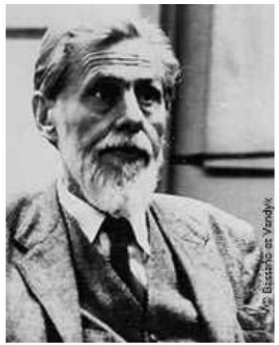
Neural Network !



Schematic of a biological neuron.

1943

Neural Network



1898-1969 (70)

Warren McCulloch,
an American neurophysiologist and cybernetician

1923-1969 (46)

Walter Pitts,
a logician who worked in the field of computational neuroscience

(1943). A logical calculus of the ideas immanent in nervous activity.

Mathematical formulation of a biological neuron,could solve AND, OR, NOT problems

1957

Neural Network

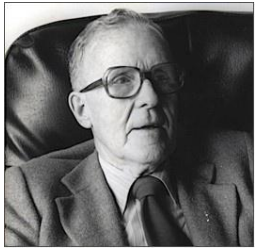
1928-1971

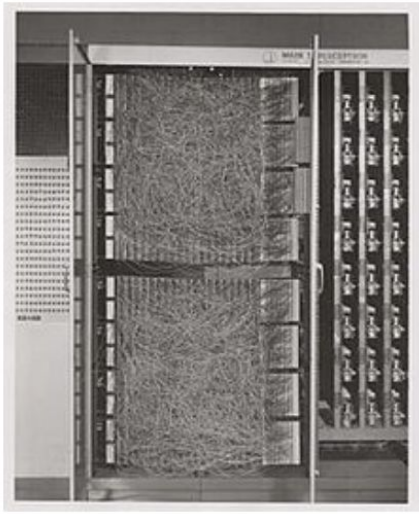
Frank Rosenblatt

An American psychologist notable in the field of artificial intelligence.

Inspired by

McCulloch, W. S., & Pitts, W. (1943).
Hebb, D. O. (1949).





Mark I Perceptron

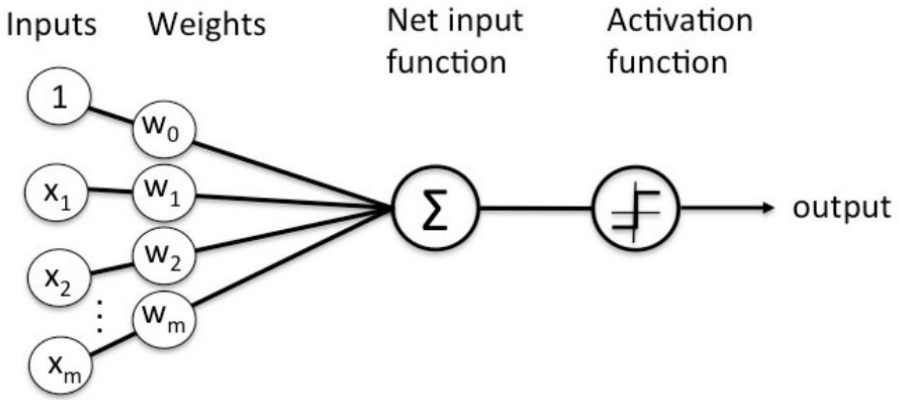
1957

Neural Network

1928-1971

Frank Rosenblatt

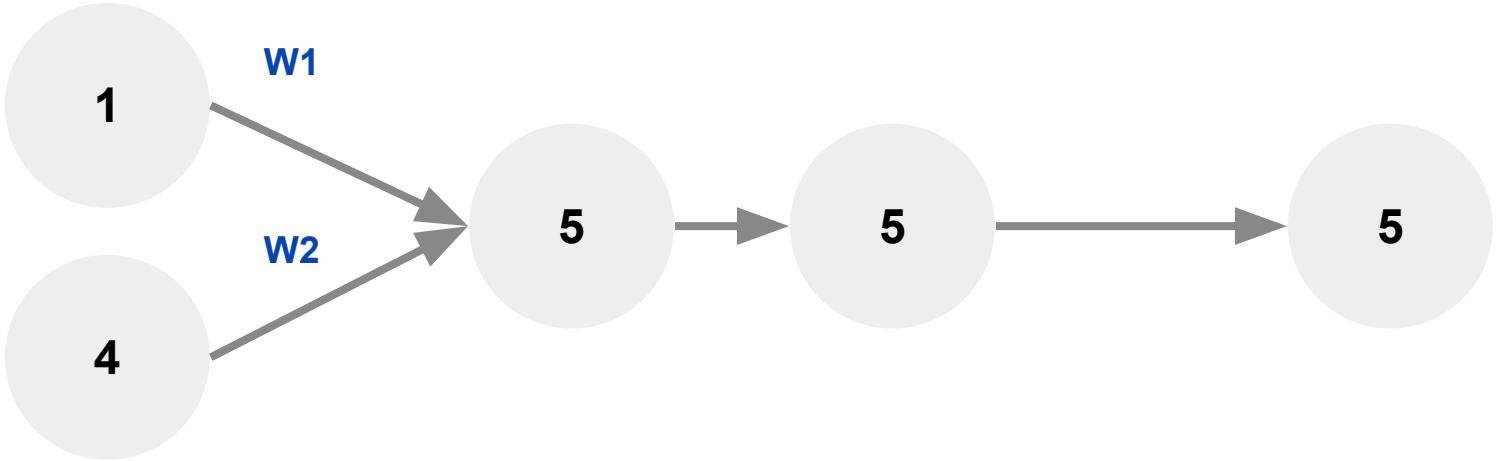
An American psychologist notable in the field of artificial intelligence.



Perceptron

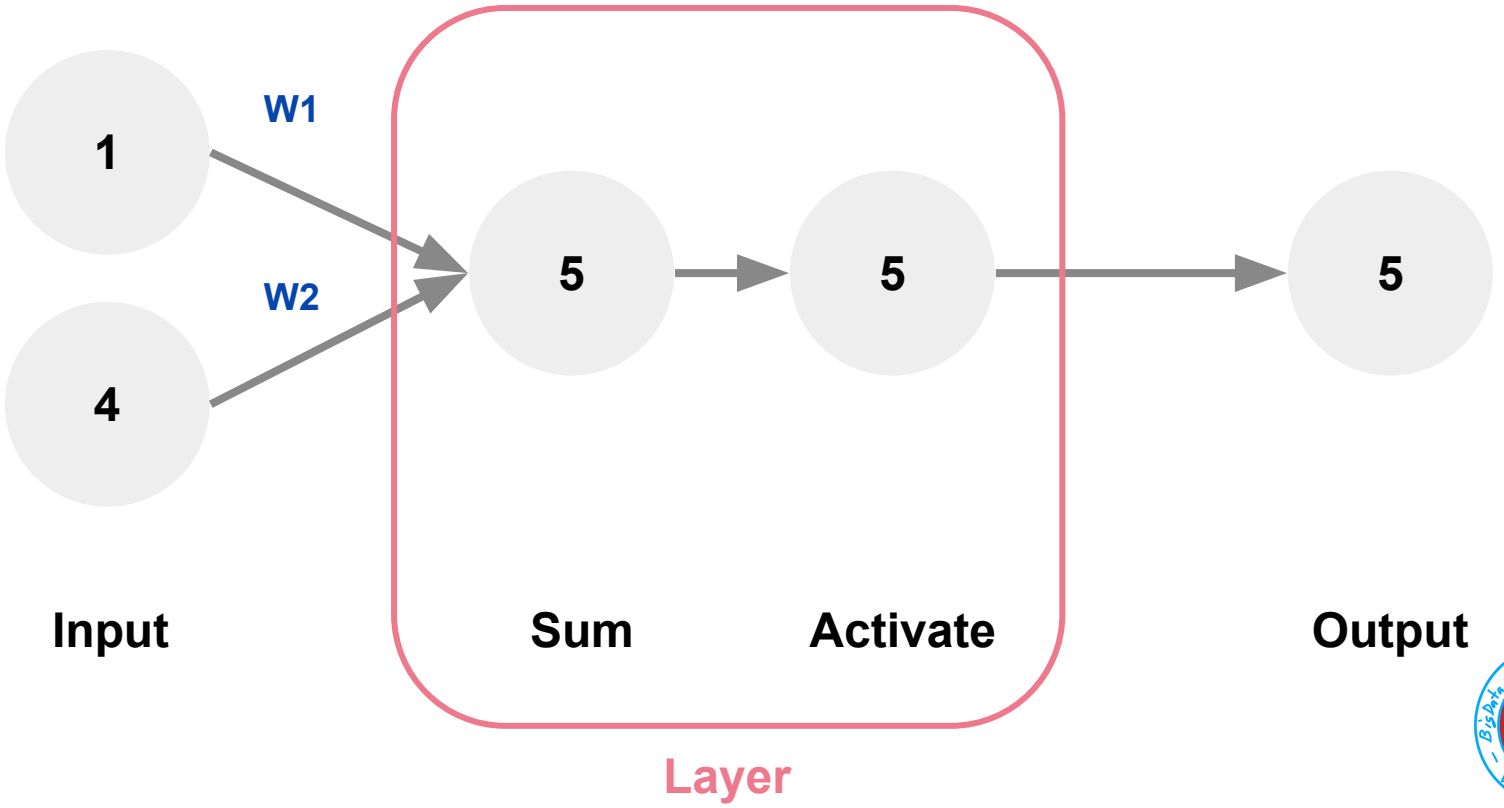
1957

Perceptron concept



1957

Perceptron Concept



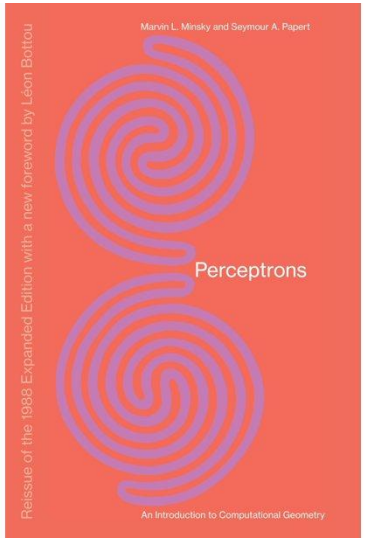
1969

Neural Network

1927-2016 (88)

Marvin Minsky

An American cognitive and computer scientist



Perceptron book

Could not solve XOR problems!

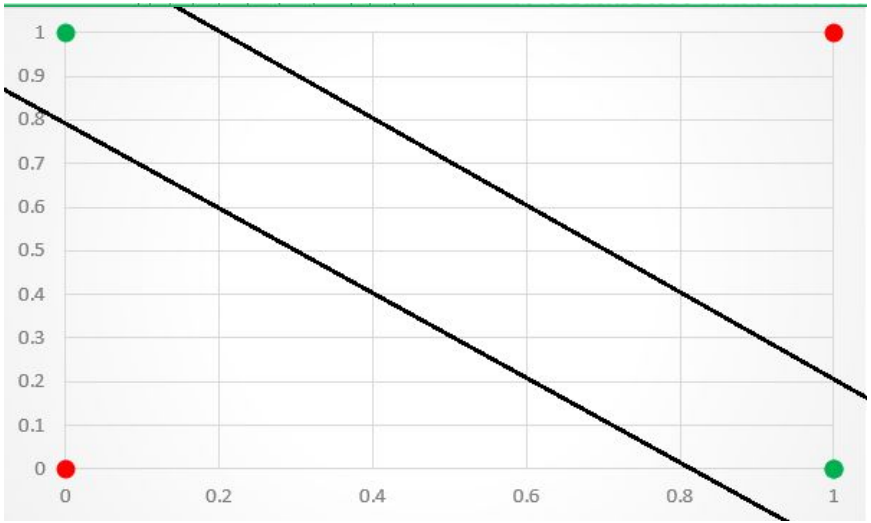
"AI WINTER"

Neural Network

Could not solve XOR problems!

Solved by

Hidden layers and non-linear activation



Neural Network

HARD TO TRAIN

Neural Network

1970

1945-Present

Seppo Linnainmaa
Finnish mathematician and computer scientist

Automatic Differentiation (Backpropagation)

Explicit, efficient error backpropagation in arbitrary, discrete, possibly sparsely connected, neural networks-like networks was first described in a 1970 master's thesis (Linnainmaa, 1970, 1976), albeit without reference to NNs,[3] when Linnainmaa introduced the reverse mode of automatic differentiation (AD), in order to efficiently compute the derivative of a differentiable composite function that can be represented as a graph, by recursively applying the chain rule to the building blocks of the function.

1986

Neural Network

You just get the output layer to reproduce the input layer, and then you don't need a separate teaching signal. Then the hidden units are representing some code for the input.

In late 1985, I actually had a deal with Dave Rumelhart that I would write a short paper about backpropagation, which was his idea, and he would write a short paper about autoencoders, which was my idea. It was always better to have someone who didn't come up with the idea write the paper because he could say more clearly what was important.

So I wrote the short paper about backpropagation, which was the *Nature* paper that came out in 1986, but Dave still hasn't written the short paper about autoencoders. I'm still waiting.

What he did do was tell Dave Zipser about the idea of autoencoders and

– Geoffrey Hinton in Talking Nets - An Oral History of Neural Networks, pg. 380

1942-2011 (68)

David Rumelhart

An American Psychologist

1947-Present (73)

Geoffrey Hinton

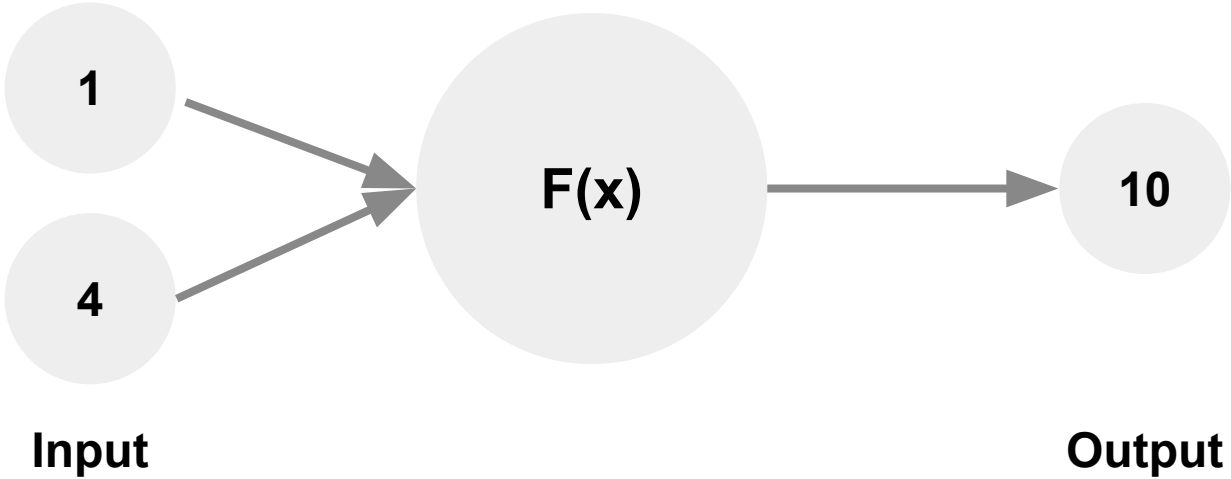
A British-Canadian Cognitive Psychologist

1957

BACK PROPAGATION ??

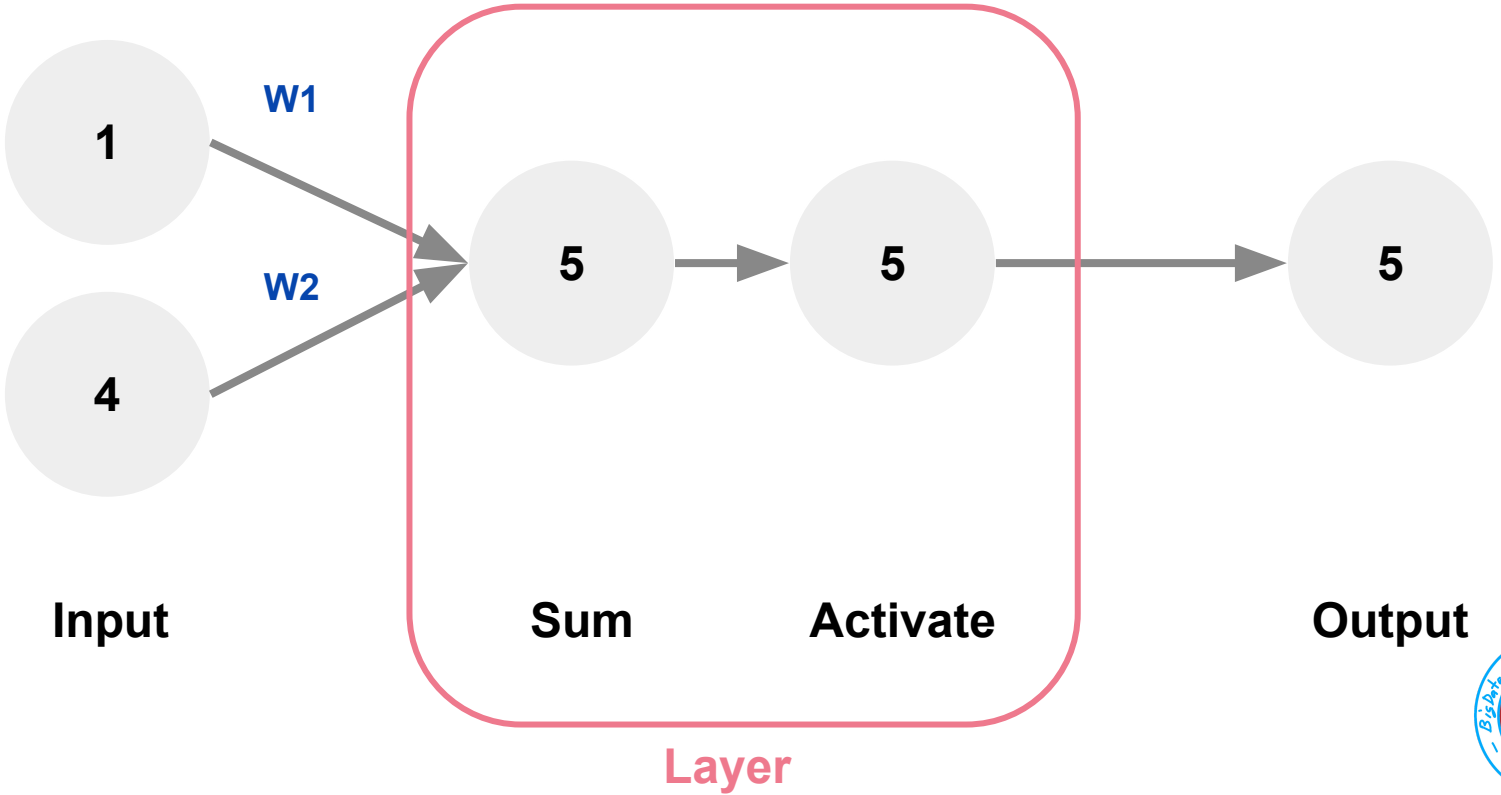
Input: [1, 4]

Output: 10



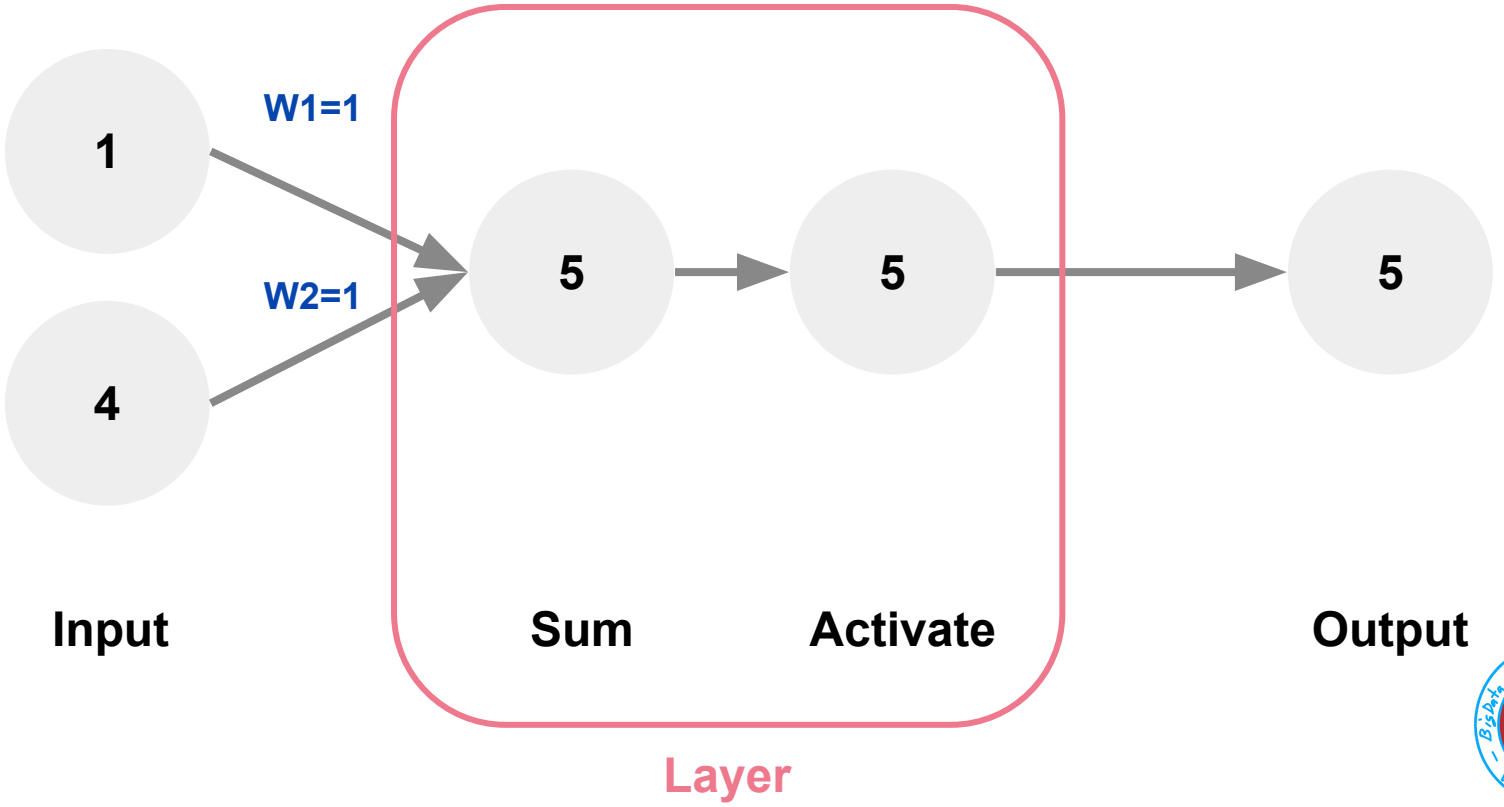
1957

BACK PROPAGATION ??



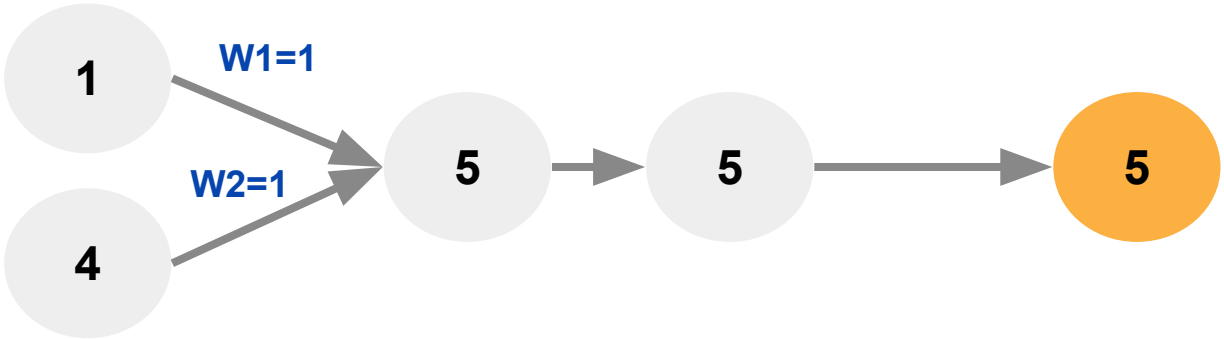
1957

BACK PROPAGATION ??



1957

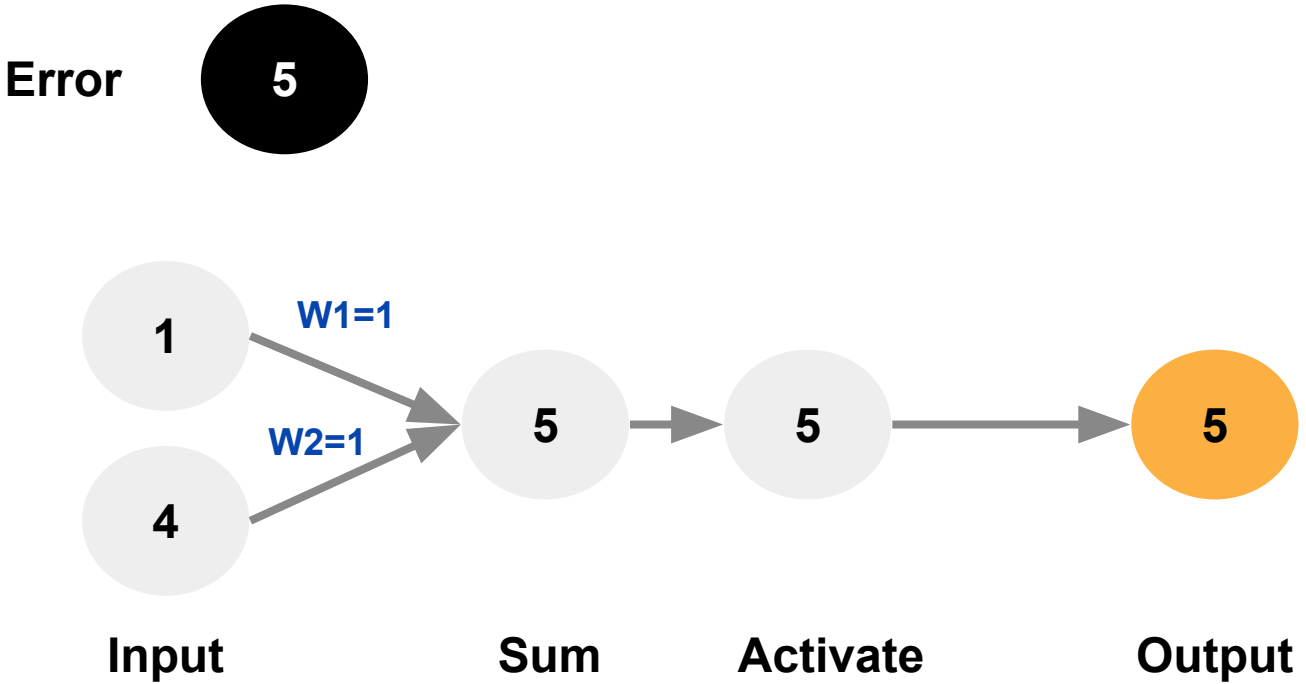
BACK PROPAGATION ??



True 10 Pred 5 Error 5

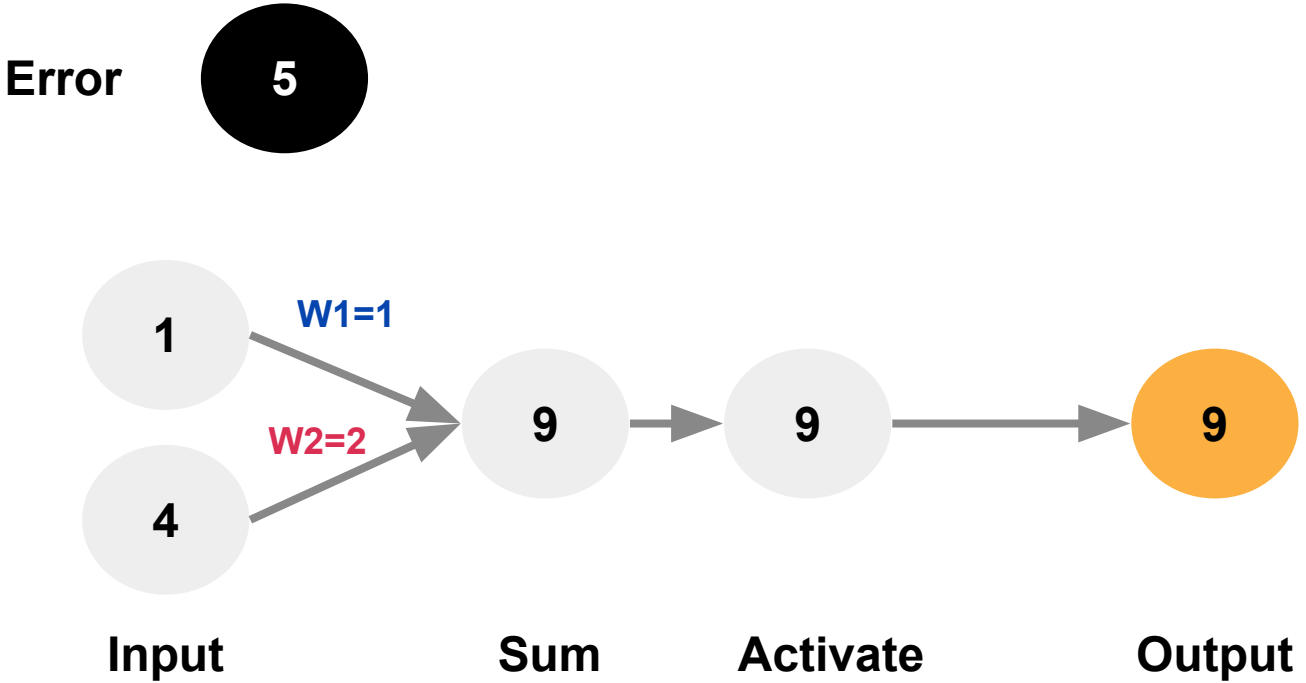
1986

BACK PROPAGATION



1986

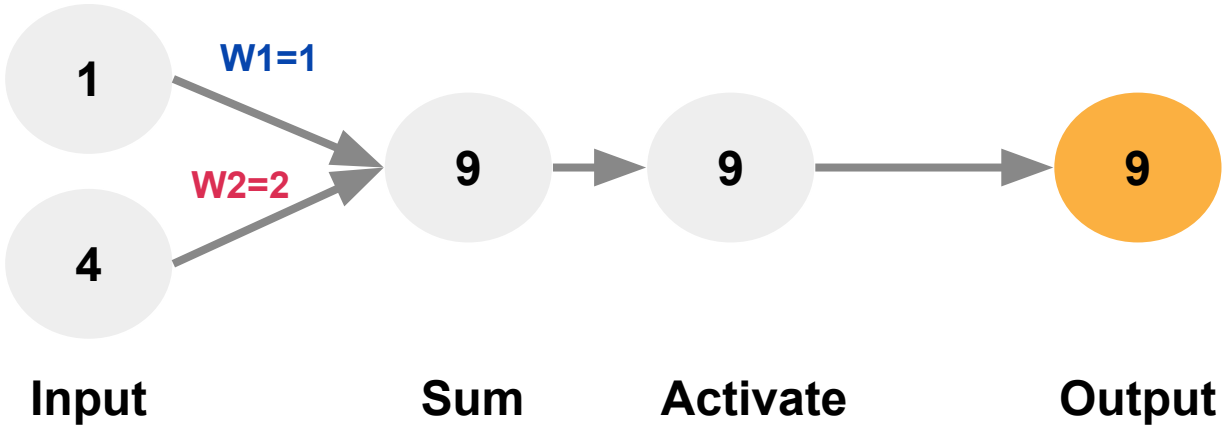
BACK PROPAGATION



1986

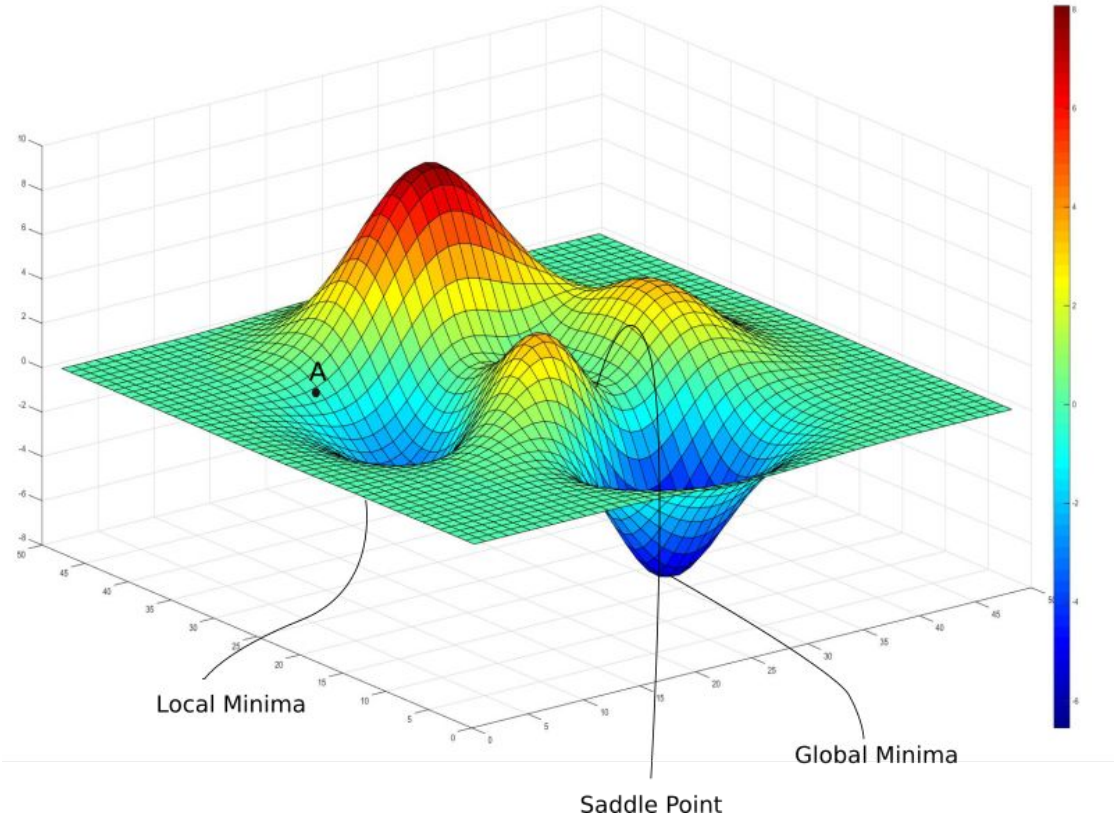
BACK PROPAGATION

True 10 Pred 9 Error 1

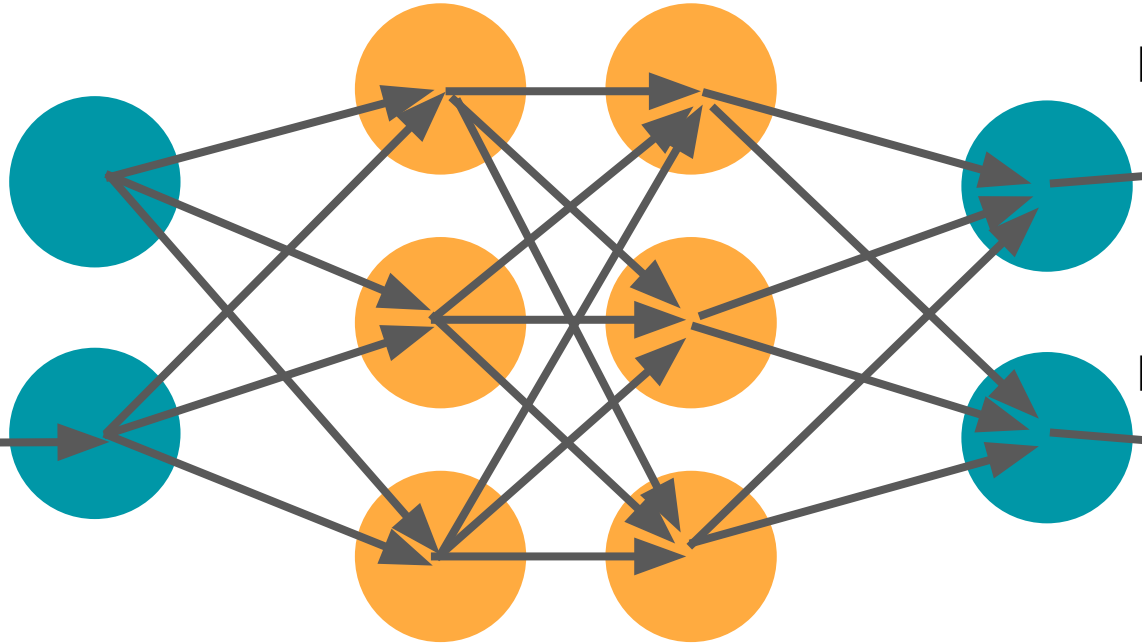


1986

BACK PROPAGATION



Feed Forward Neural Network
Multi-Layer Perceptron



p=0.25

p=0.75

Input

Hidden Layer

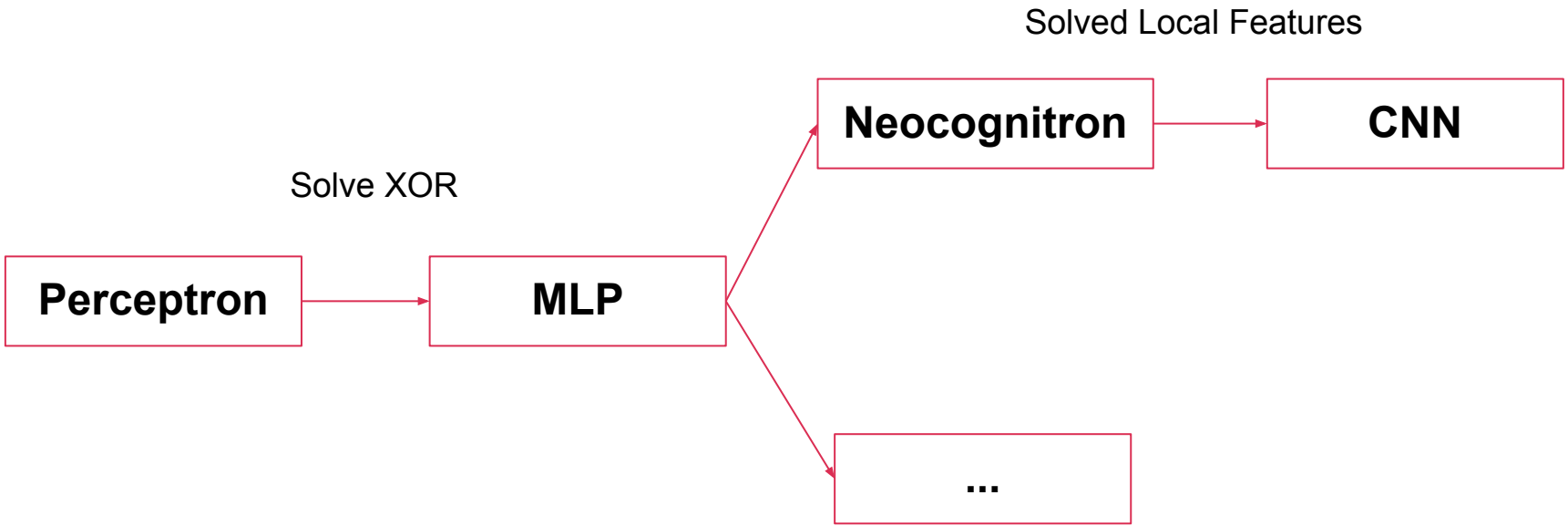
Output

Neural Network

New
Problems

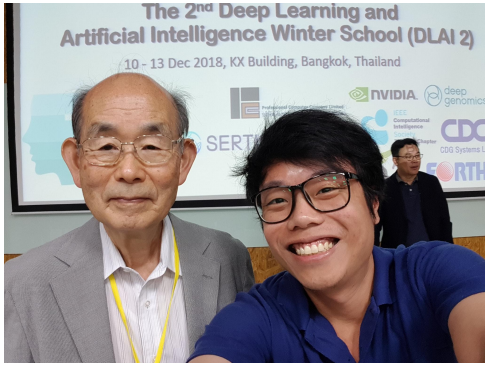
Spatial Features

Neural Network



1980

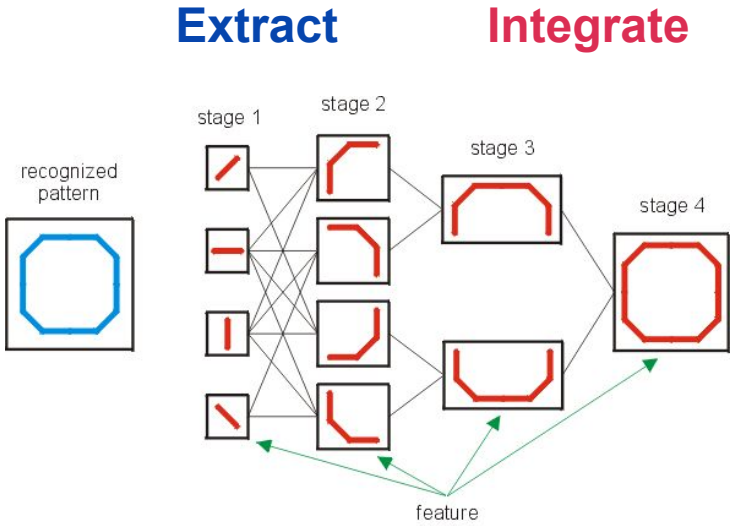
Neural Network



1936-Present

Kunihiko Fukushima

A Japanese computer scientist



In 1980, Fukushima published the neocognitron the original deep convolutional neural network (CNN)

1989

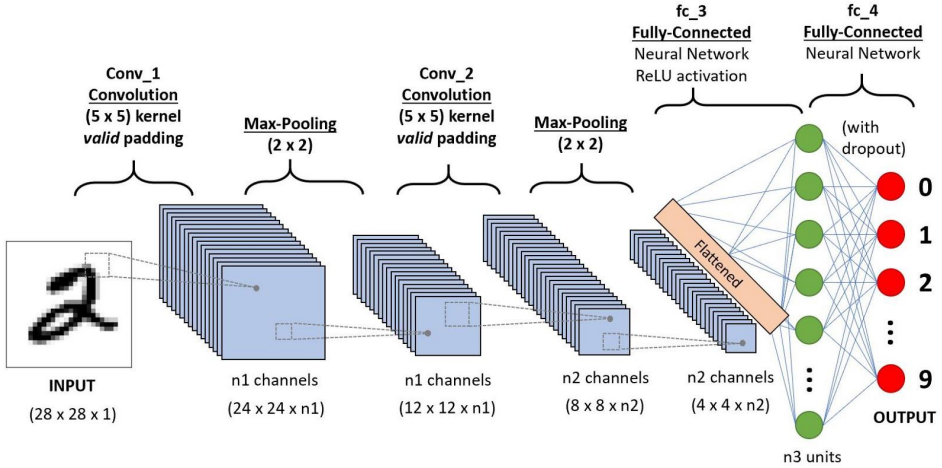
Neural Network

1960-Present (61)

Yann LeCun

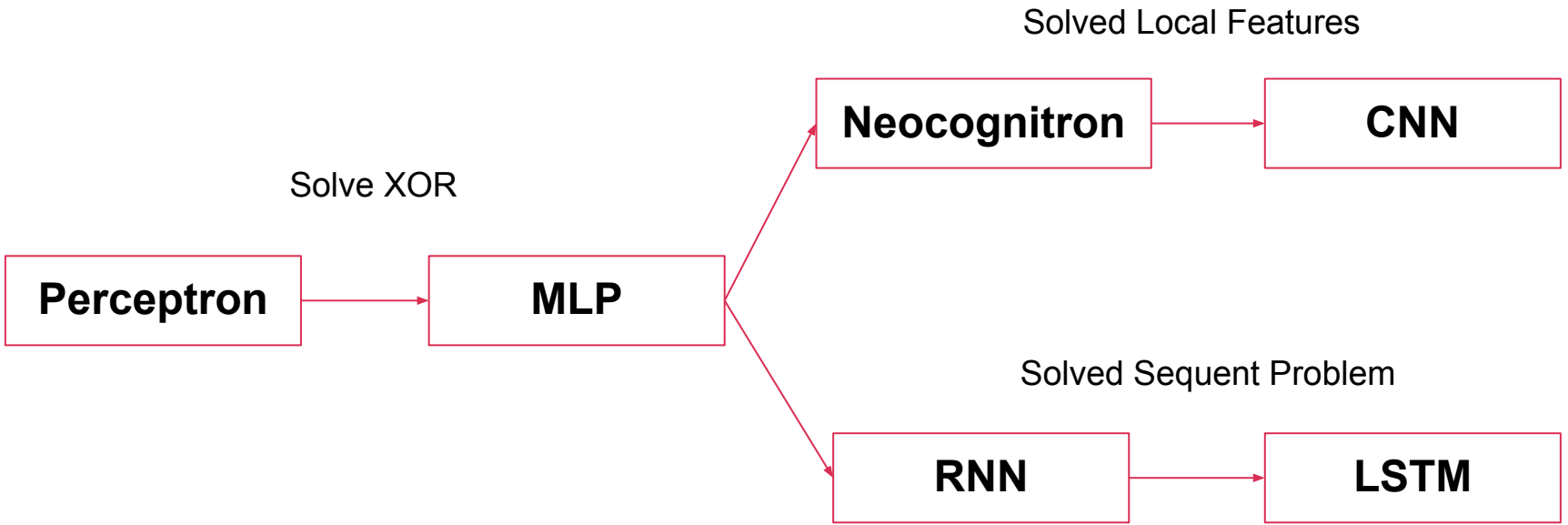
An computer scientist

Weight Sharing + Pooling



LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. Neural computation,
1(4), 541-551.

Neural Network



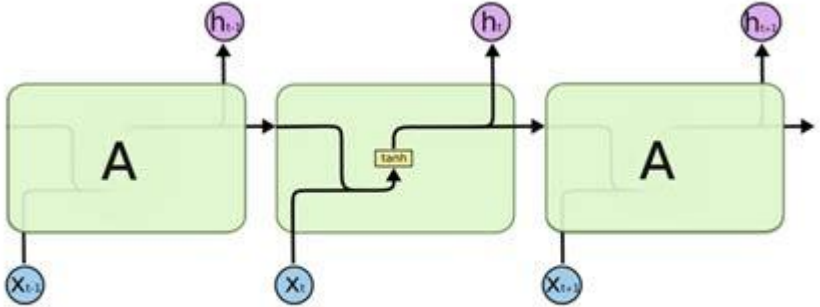
1982

Neural Network

1933-Present

John Hopfield

An American Scientist



A recurrent neural network (RNN) is a class of artificial neural network where connections between nodes form a directed graph along a sequence.

Vanishing Gradients!

1982

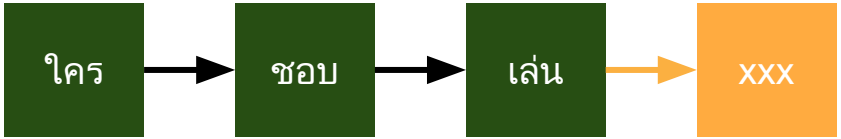
Recurrent Neural Network

Example: ใครชอบเล่นหวย ฉันล่ะสิ ฉันละสิ !

Create Sequence: ใคร-ชอบ-เล่น-หวย- -ฉัน-ล่ะ-สิ- -ฉัน-ละ-สิ !

Learning:

Predict:



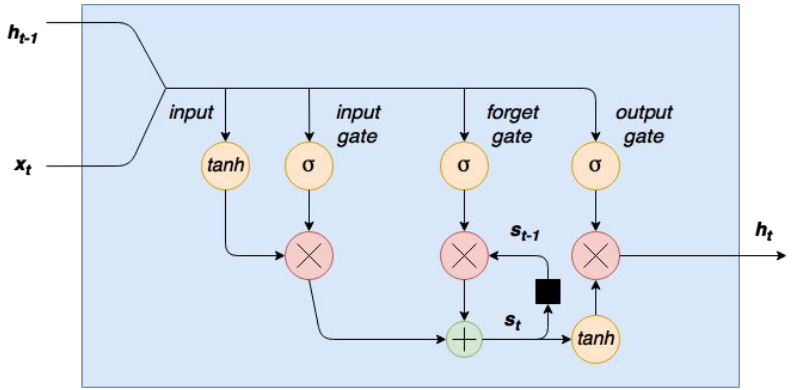
1997

Neural Network

1942-2011 (68)

Sepp Hochreiter

An American Psychologist



Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

1997

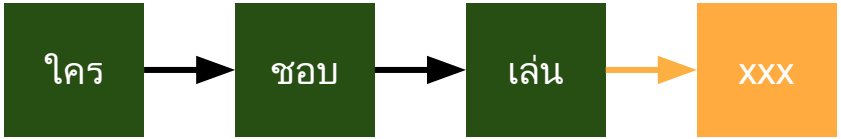
LONG SHORT-TERM MEMORY

Example: ใครชอบเล่นหวย ฉันล่ะสิ ฉันละสิ !

Create Sequence: ใคร-ชอบ-เล่น-หวย- -ฉัน-ล่ะ-สิ- -ฉัน-ละ-สิ !

Learning:

Predict:



1997

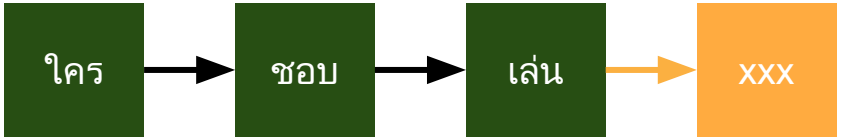
LONG SHORT-TERM MEMORY

Example: ใครชอบเล่นหวย ฉันล่ะสิ ฉันละสิ !

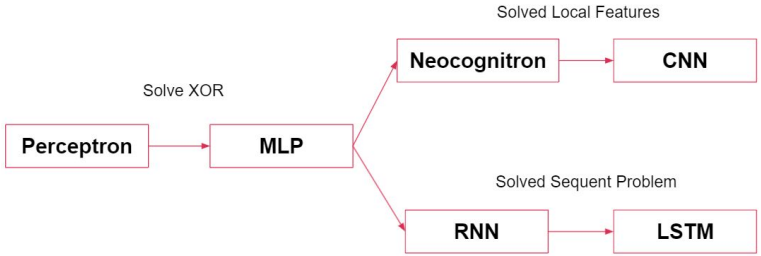
Create Sequence: ใคร-ชอบ-เล่น-หวย- -ฉัน-ล่ะ-สิ- -ฉัน-ละ-สิ !

Learning:

Predict:



Neural Network



"2nd AI Winter"

< 1990

1990-2000

1943-2000 → 57 Years

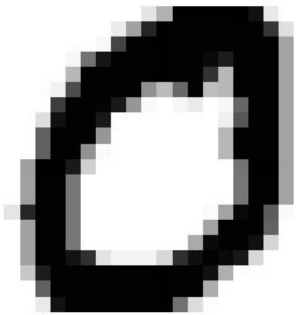
2000

New Millennium

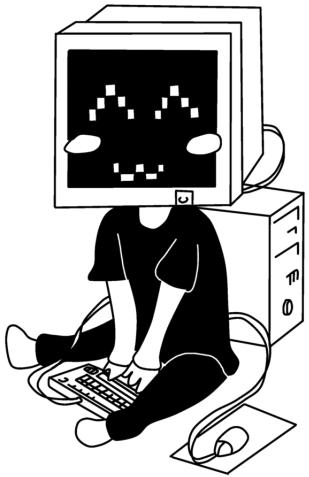
- Pre-trained Model
- Autoencoder
- Generative Adversarial Network
- Attention Model
- Transformer
- BERT

2012

Pre-trained model



Problem

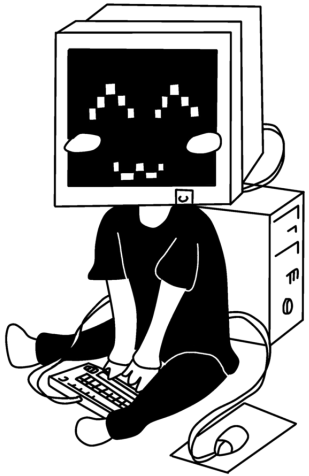


Machine

Result

2012

Pre-trained model



Problem

Machine

Result

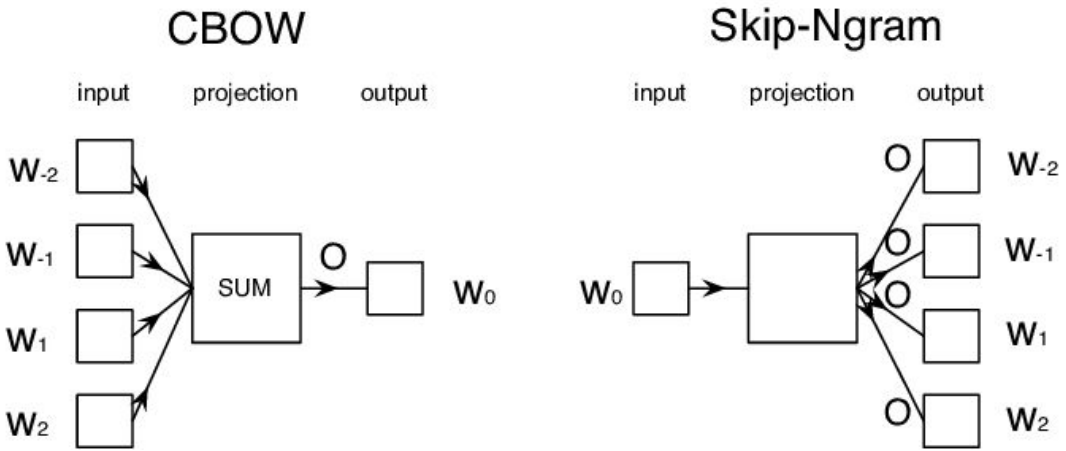
2012

Pre-trained model

1982 - Present

Tomas Mikolov

Computer scientist

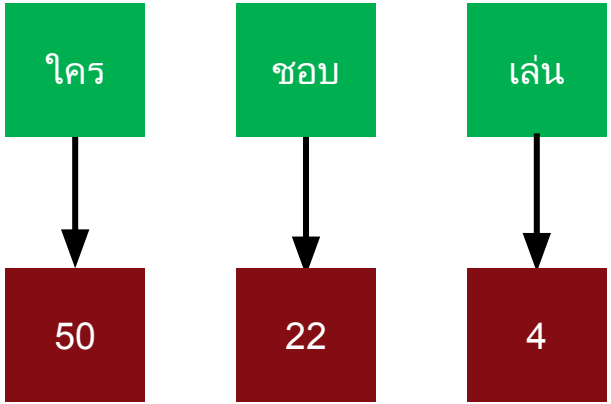


Tomas Mikolov, Ilya Sutskever, Kai Chen, Gregory S. Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality. In NeurIPS, 2013.

2012

Pre-trained model

Text:



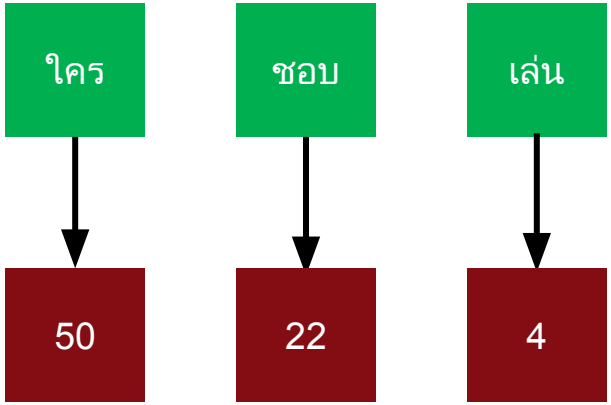
Numerical Data:

ใคร > ชอบ > เล่น

2012

Pre-trained model

Text:

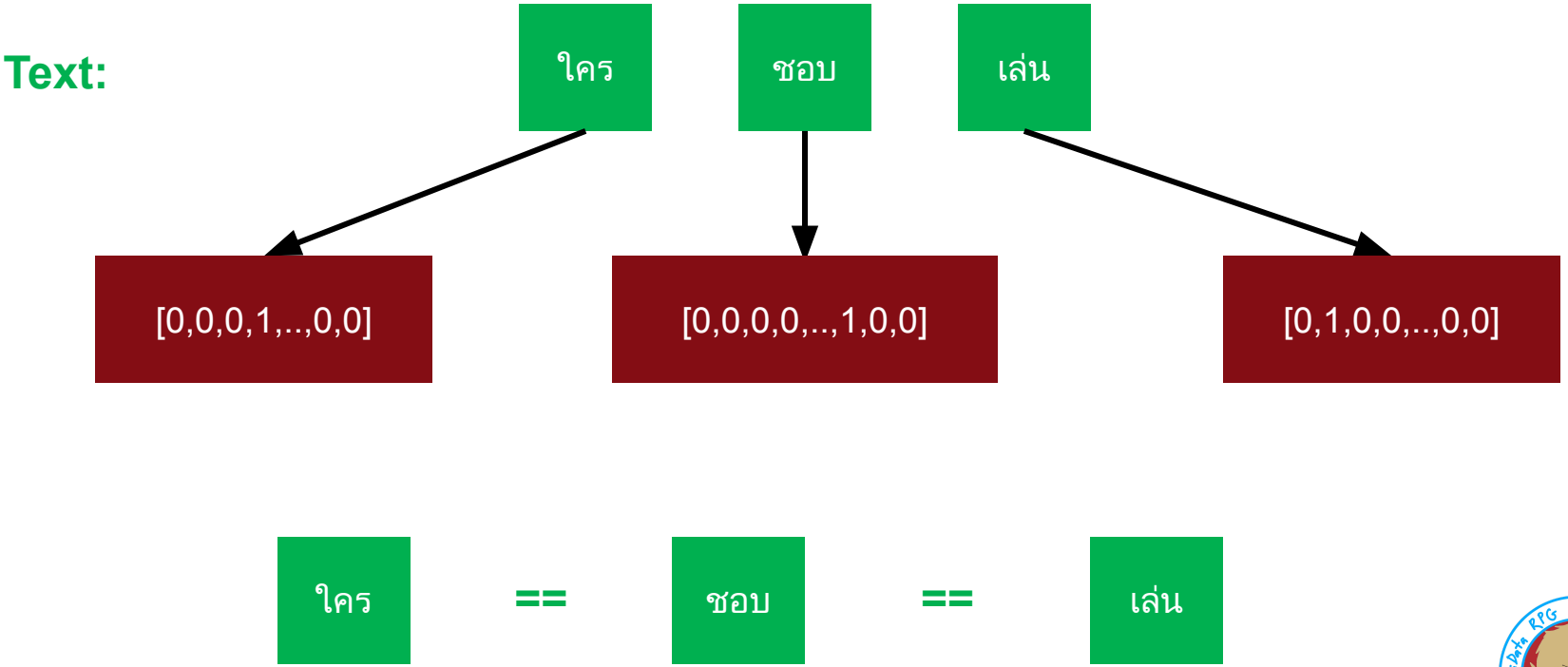


Numerical Data:

ใคร > ชอบ > เล่น

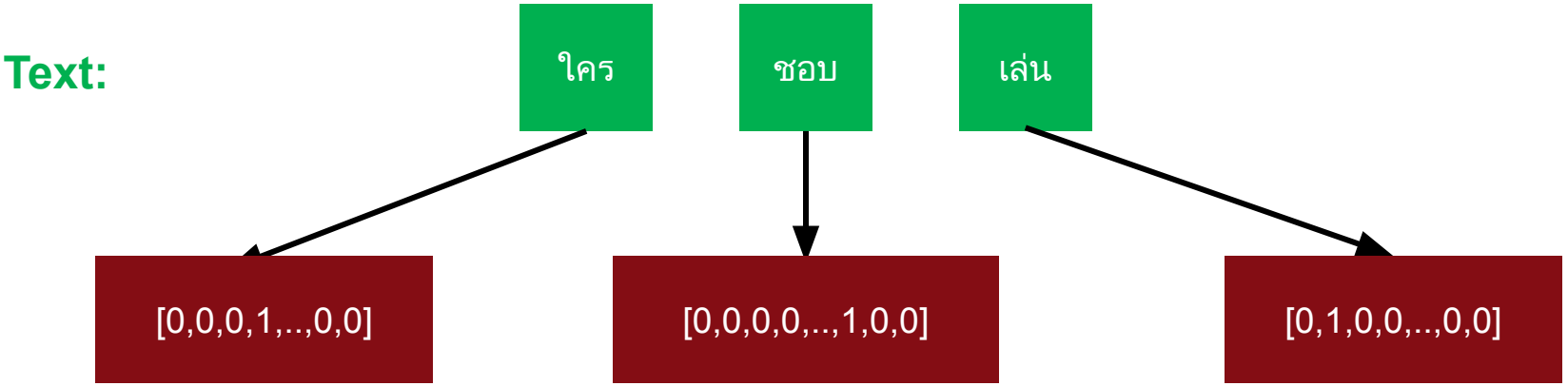
2012

Pre-trained model



2012

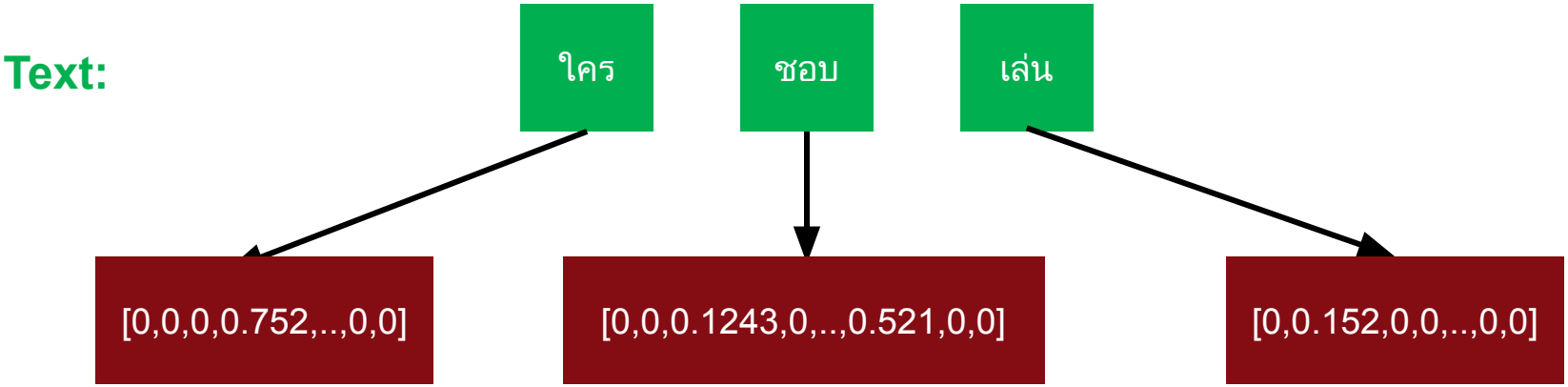
Pre-trained model



High number of unique word
Leads High memory consumption

2012

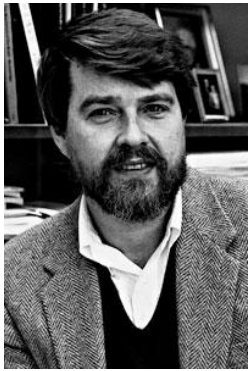
Pre-trained model



Fixed dimensionality with Continuous values

1986

Autoencoder



1942-2011 (68)

David Rumelhart

An American psychologist

1947-present (73)

Geoffrey Hinton

A British-Canadian cognitive psychologist

Autoencoders have been first introduced in as a neural network that is trained to reconstruct its input. Their main purpose is learning in an unsupervised manner an "informative" representation of the data that can be used for various implications such as clustering.

Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1. chap. Learning Internal Representations by Error Propagation, pp. 318–362. MIT Press, Cambridge, MA, USA (1986). URL http://dl.acm.org/citation.cfm?id=104279.104293

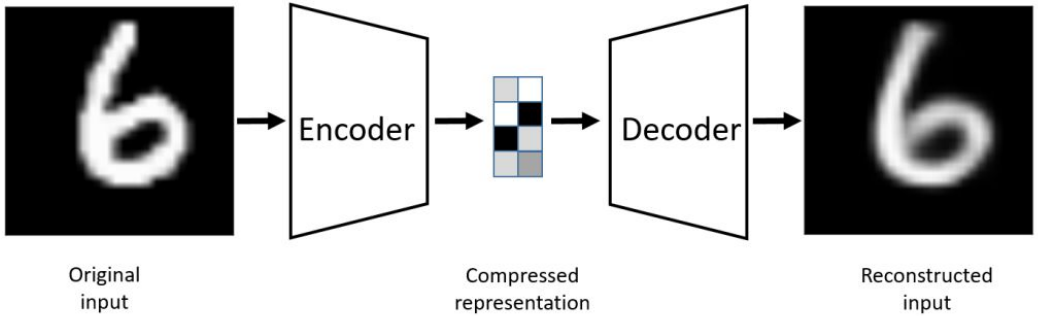
2012

Autoencoder

1957-Present

Pierre Baldi

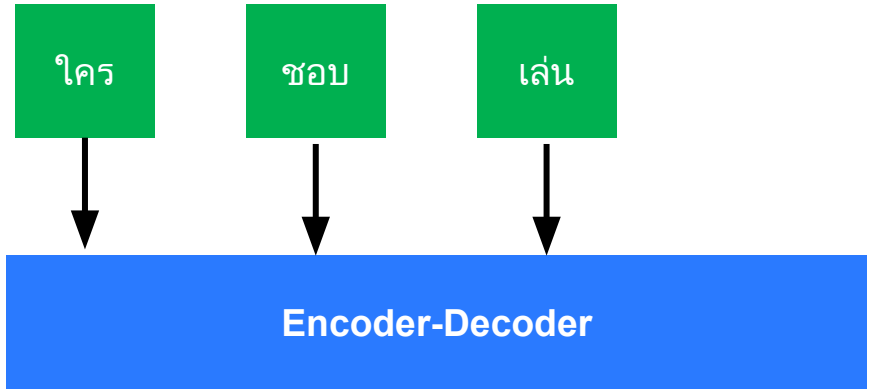
A distinguished professor of computer science at University of California Irvine



Baldi, P.: Autoencoders, unsupervised learning, and deep architectures. In: I. Guyon, G. Dror, V. Lemaire, G. Taylor, D. Silver (eds.) Proceedings of ICML Workshop on Unsupervised and Transfer Learning, Proceedings of Machine Learning Research, vol. 27, pp. 37–49. PMLR, Bellevue, Washington, USA (2012)

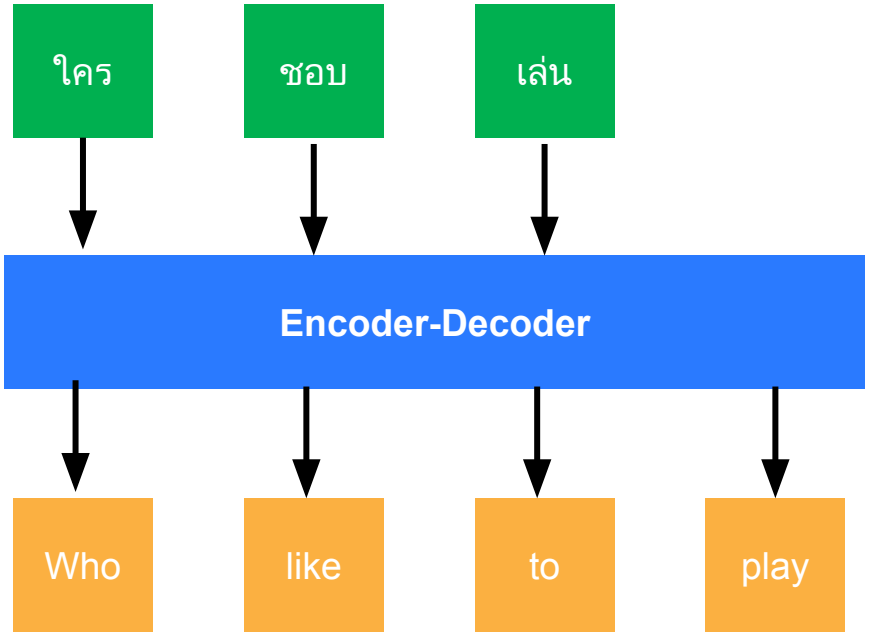
2012

AUTOENCODER



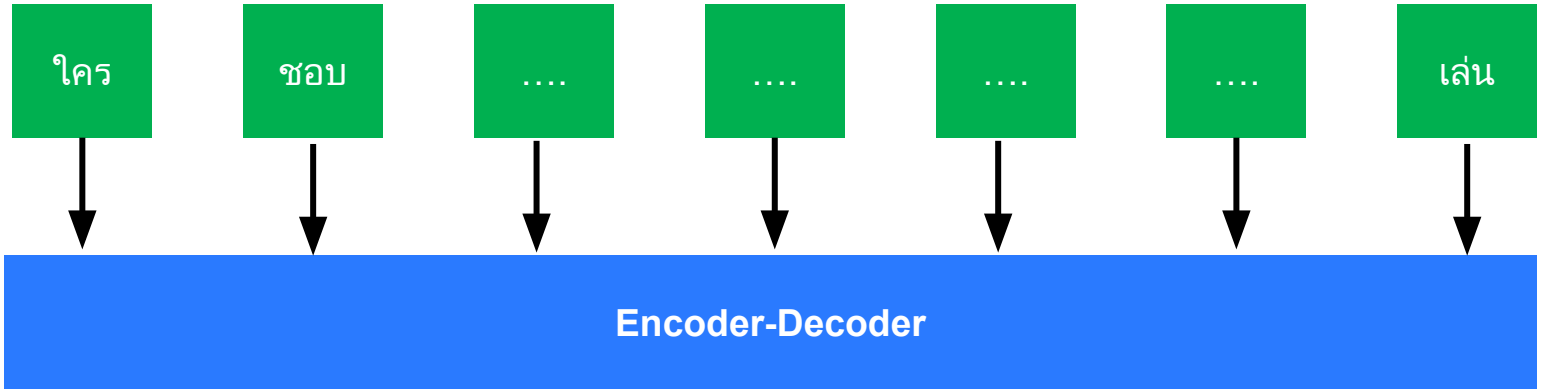
2012

Autoencoder



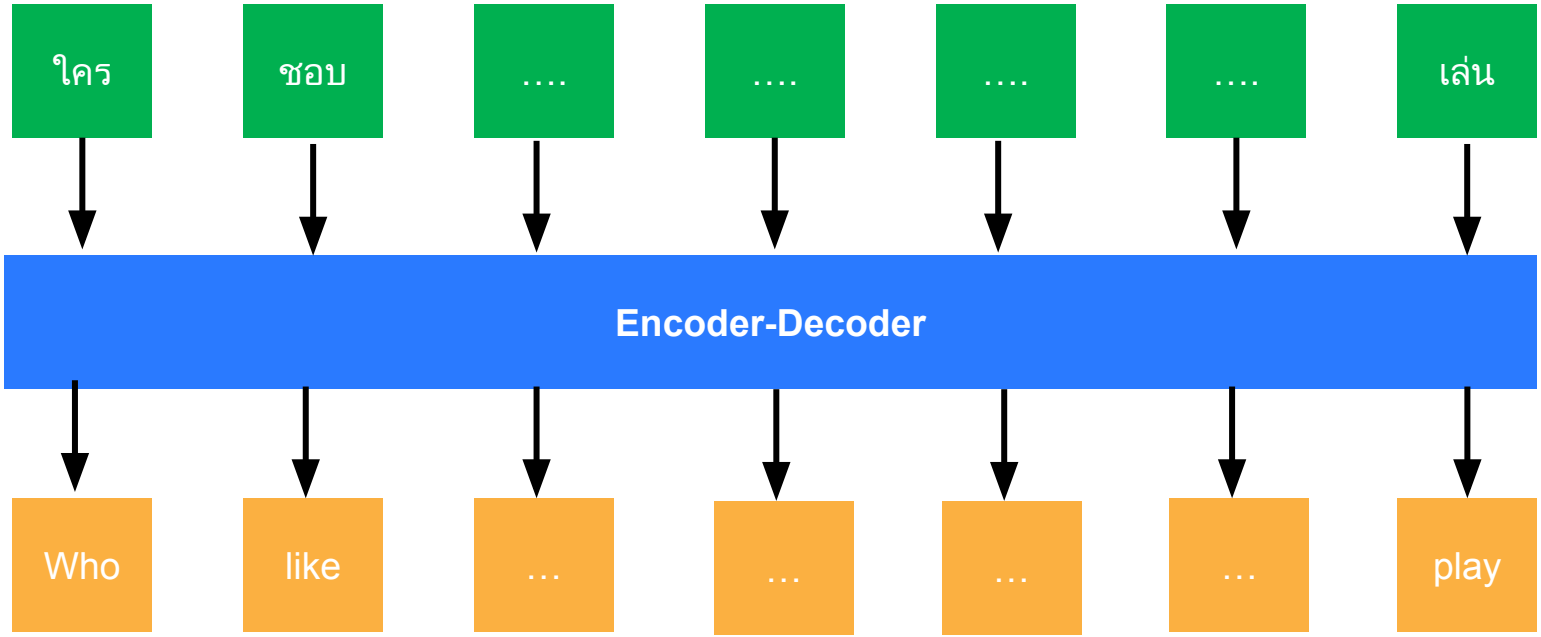
2012

AUTOENCODER



2012

Autoencoder



2014

Generative Adversarial Network

1985-Present

Ian J. Goodfellow

Computer Scientist

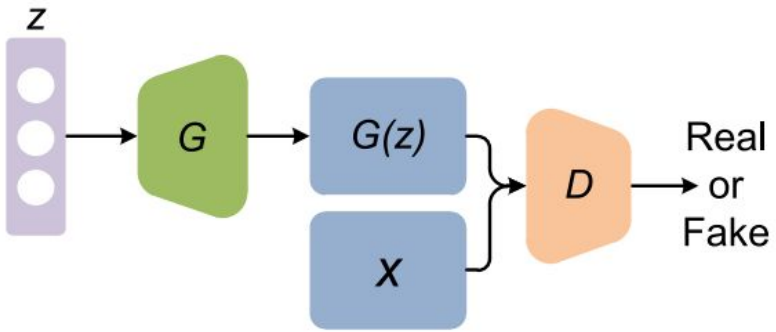


FIGURE 1. The architecture of generative adversarial networks.

I. Goodfellow et al., "Generative adversarial nets," in Proc. Adv. Neural Inf. Process. Syst., 2014, pp. 2672–2680. [Online]. Available: http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

2017

Attention is All You Need

1986-Present

Ashish Vaswani

Staff Research Scientist
at Google Brain

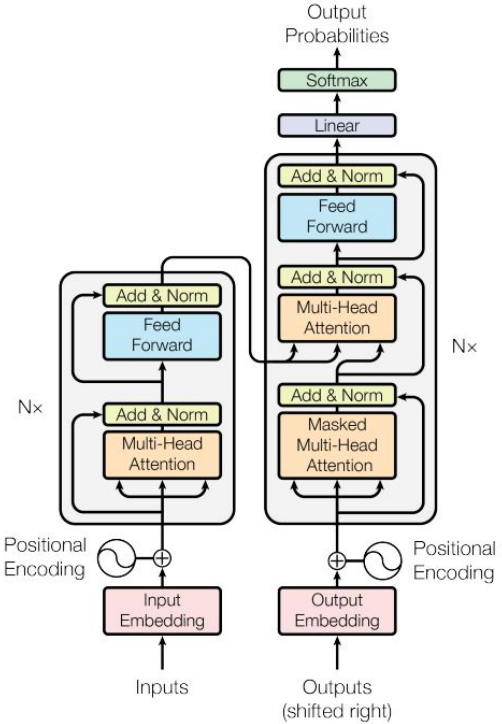
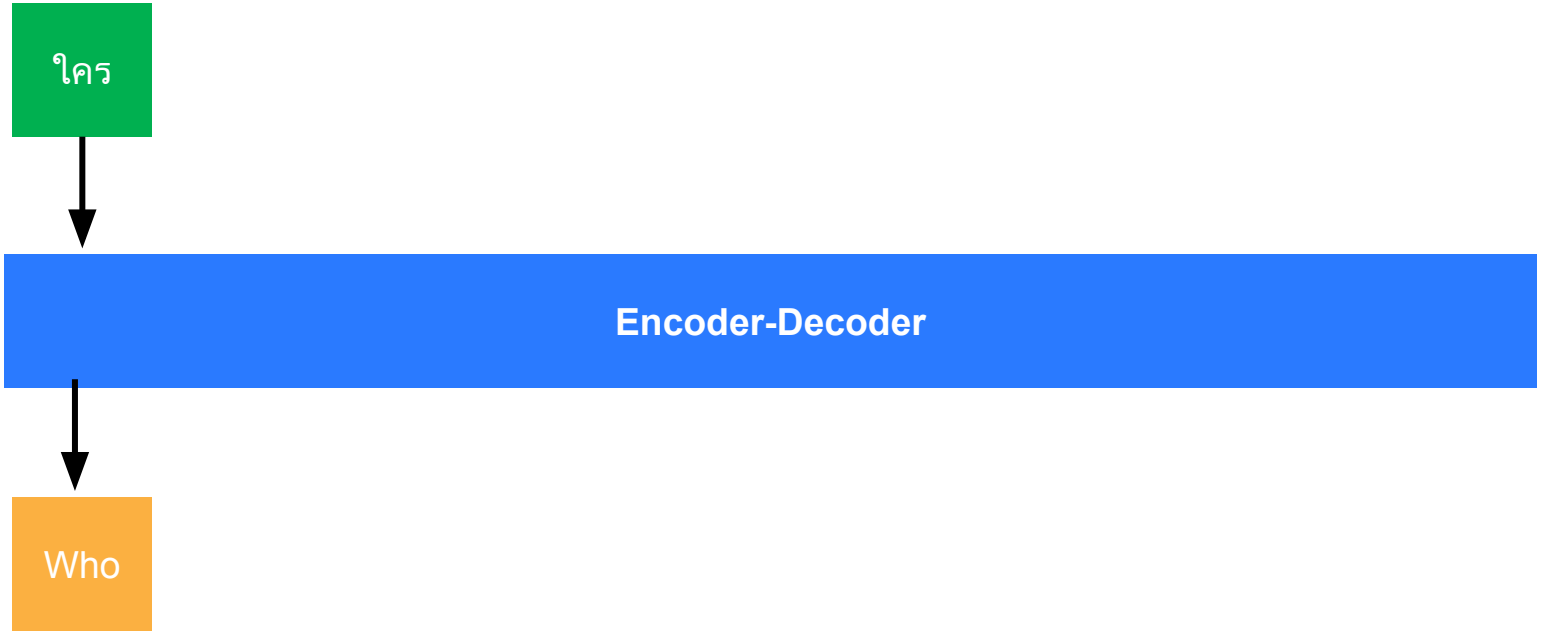


Figure 1: The Transformer - model architecture.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In Advances in Neural Information Processing Systems, pages 6000–6010.

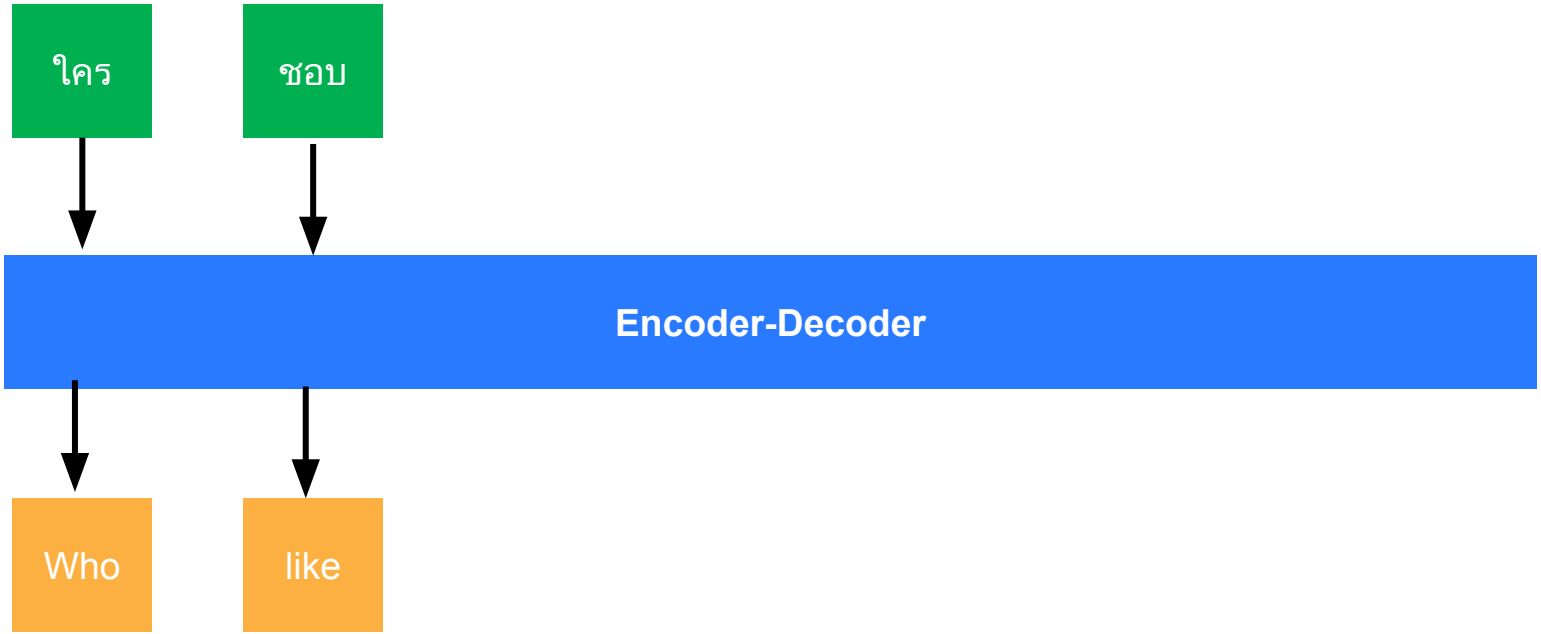
2012

Attention



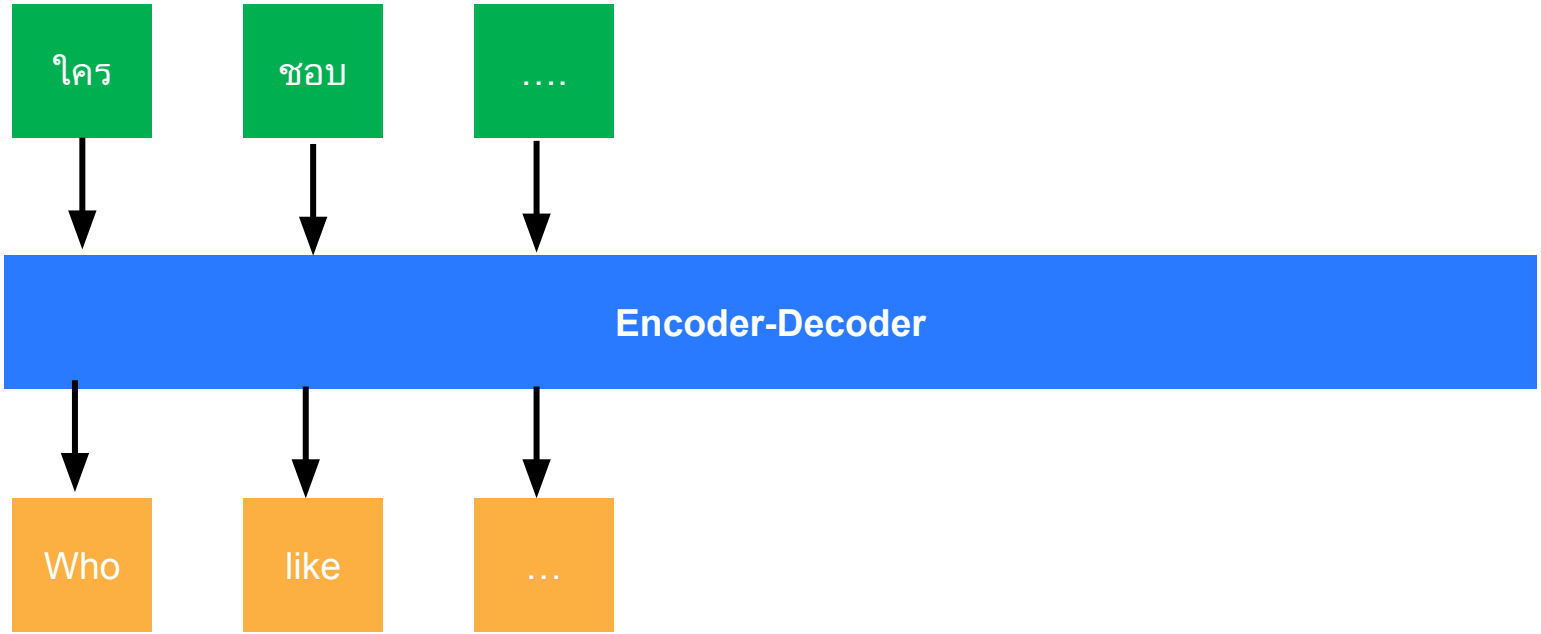
2012

Attention



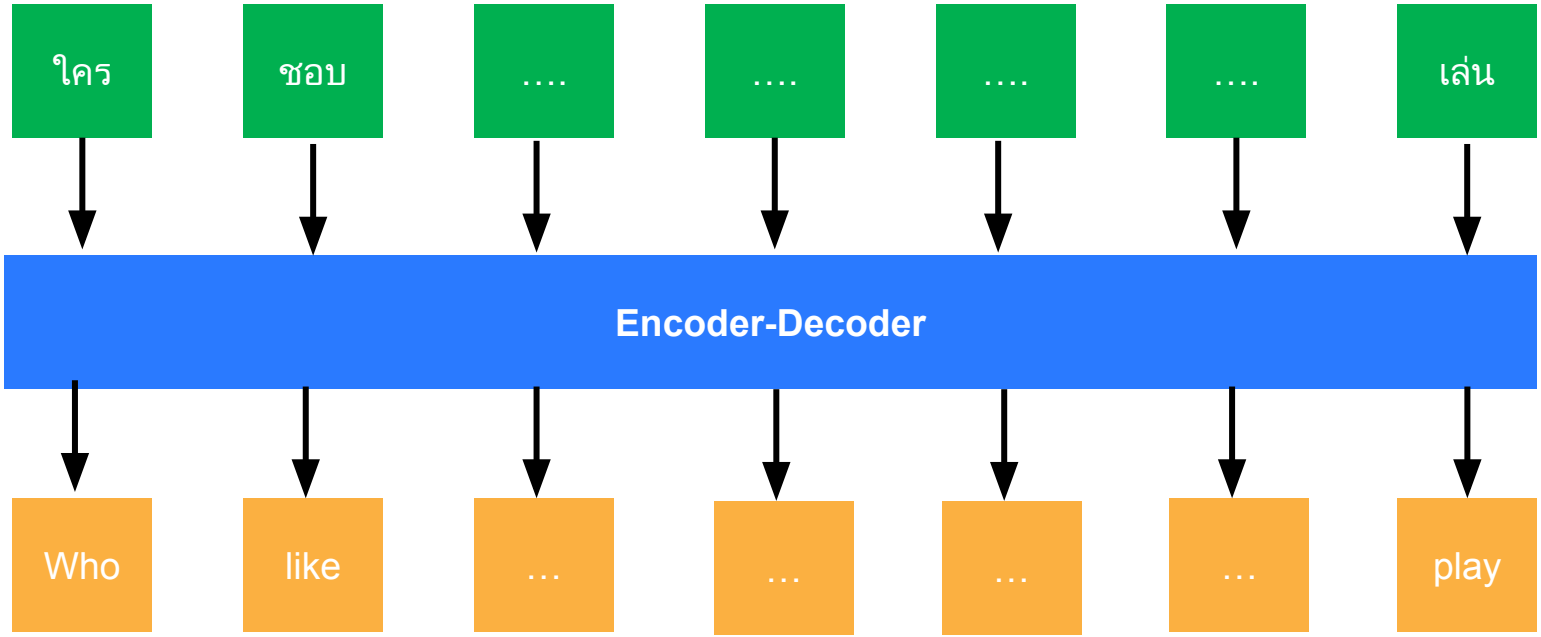
2012

Attention



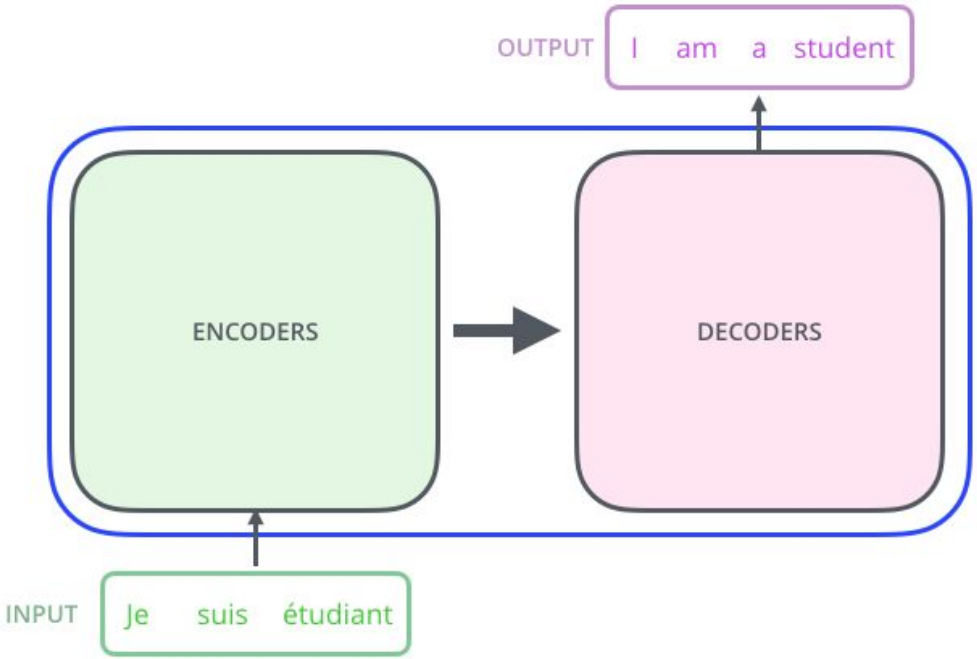
2012

Attention



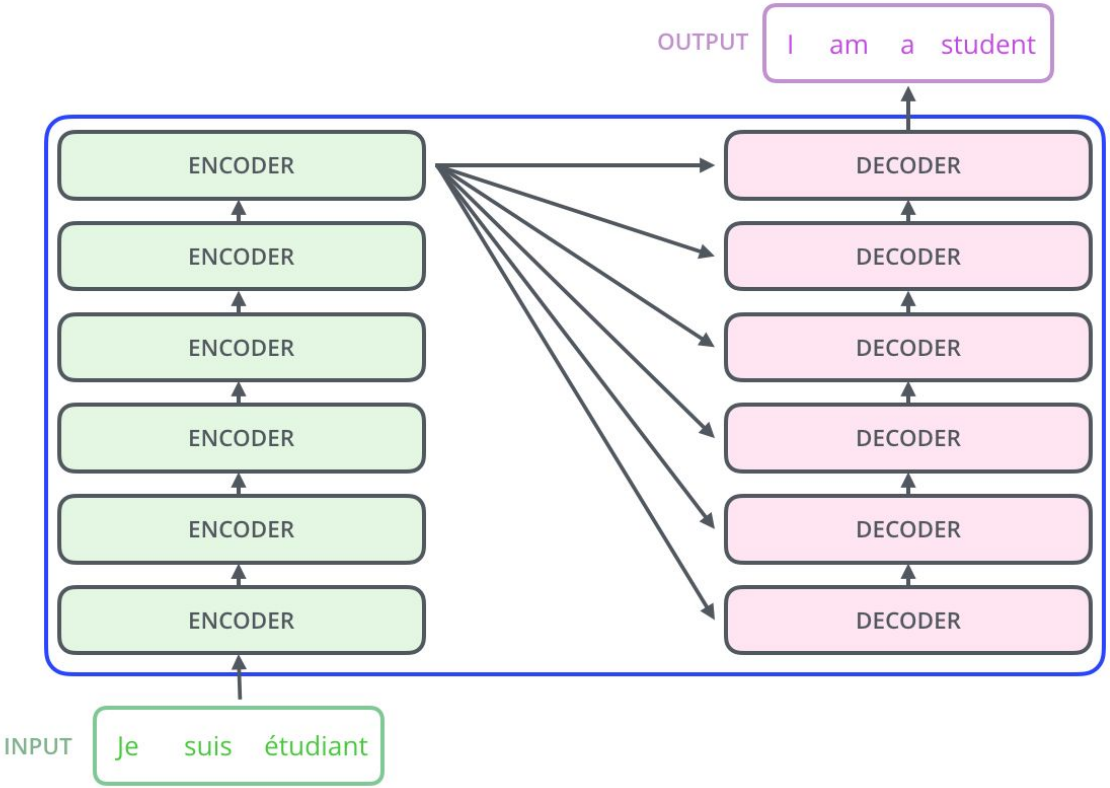
2017

Encoder-Decoder



2017

TRANSFORMER



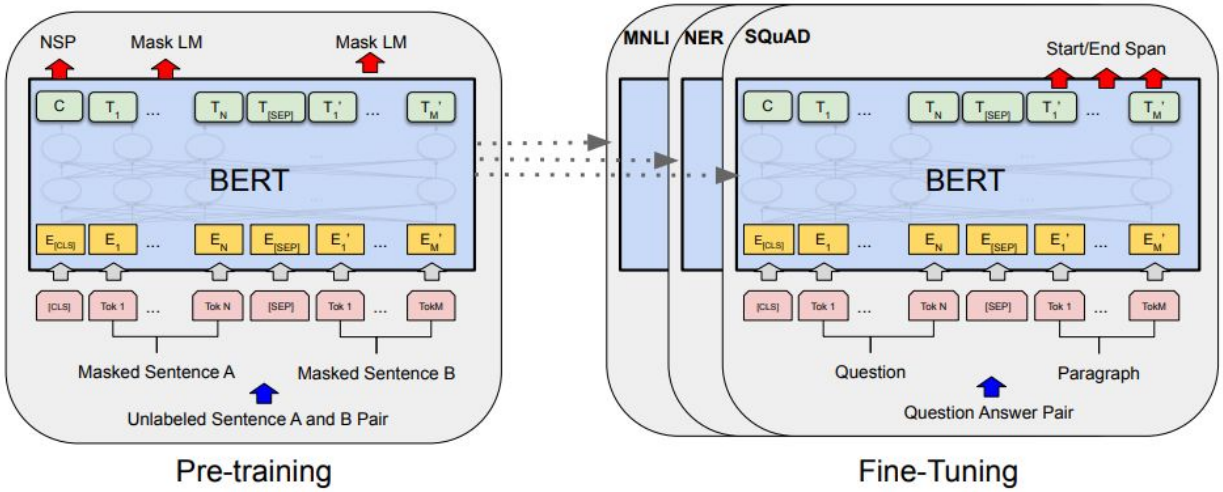
2018

BERT

1986-Present

Jacob Devlin

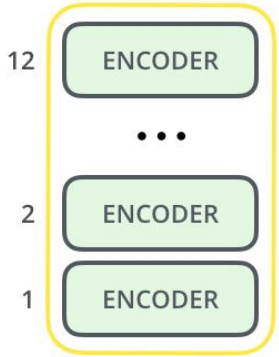
Software Engineer at Google



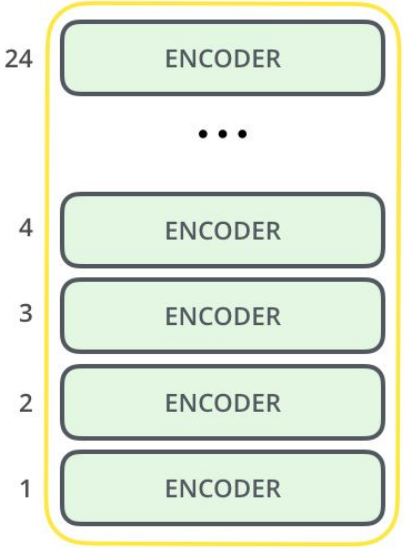
Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805, 2018.

2018

BERT



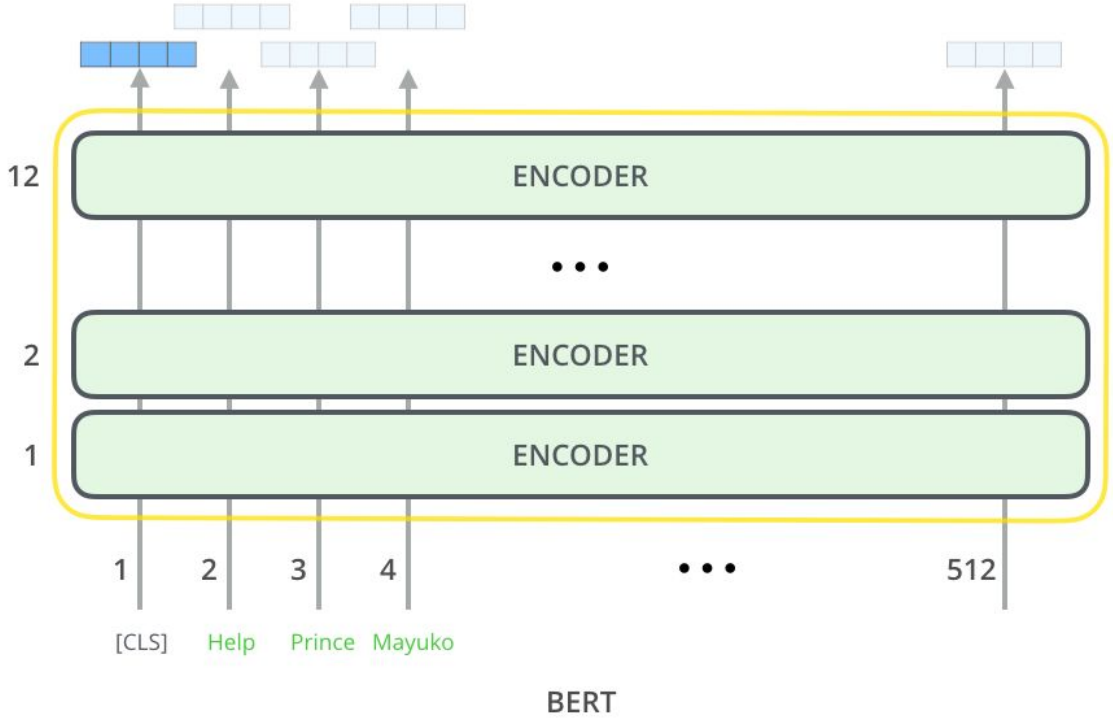
$BERT_{BASE}$



$BERT_{LARGE}$

2018

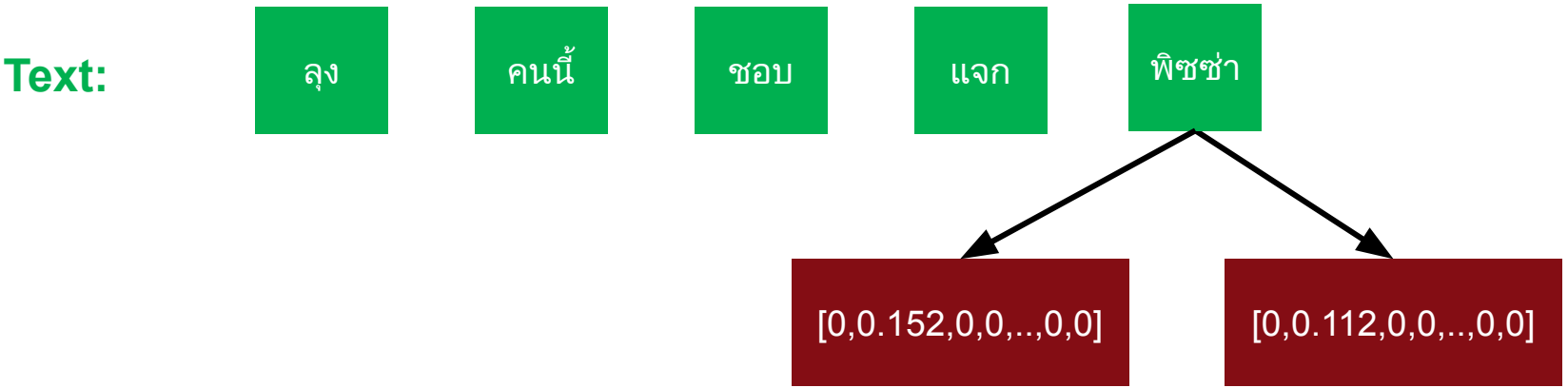
BERT



BERT

2018

BERT



Same word different context represents different vector

Neural Network

1943-2000 → 57 years

2000-2022 → 22 years

~ 79 years

2022 ++

Future and Challenge

- Future Applications
- Big Data and Consense
- Edge Computing
- Low Memory Consumption
- Multi-tasked model prediction
- Explainable Artificial Intelligence (XAI)
- Fairness and Ethic

2022 ??

METAVERSE

2022 ??

METAVERSE

Physical World

Digital Twin

Metaverse

A. Technologies

- Artificial Intelligence
- Blockchain
- Computer Vision
- Network
- Edge Computing
- User Interactivity
- Extended Reality
- IoT & Robotics

B. Ecosystems

- Avatar
- Content Creation
- Virtual Economy
- Social Acceptability
- Security & Privacy
- Trust & Accountability

Neural Network

1943-2000 → 57 years

2000-2022 → 22 years

2022-2043 → 21 years

~ 100 years

THANKS

Special edition with

<DEV>
MOUNTAIN
TECH FESTIVAL

References

- https://www.ibm.com/cloud/learn/neural-networks
- https://www.simplilearn.com/tutorials/deep-learning-tutorial/deep-learning-algorithm
- https://sebastianraschka.com/pdf/lecture-notes/stat479ss19/L02_dl-history_slides.pdf
- https://www.inf.ed.ac.uk/teaching/courses/nlu/assets/reading/Gurney_et_al.pdf
- https://people.idsia.ch/~juergen/who-invented-backpropagation.html
- McCulloch, W. S., & Pitts, W. (1943). Hebb, D. O. (1949).
- In 1980, Fukushima published the neocognitron the original deep convolutional neural network (CNN)
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. Neural computation, 1(4), 541-551.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

References

- Tomas Mikolov, Ilya Sutskever, Kai Chen, Gregory S. Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality. In NeurIPS, 2013.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1. chap. Learning Internal Representations by Error Propagation, pp. 318–362. MIT Press, Cambridge, MA, USA (1986). URL http://dl.acm.org/citation.cfm?id=104279.104293
- Baldi, P.: Autoencoders, unsupervised learning, and deep architectures. In: I. Guyon, G. Dror, V. Lemaire, G. Taylor, D. Silver (eds.) Proceedings of ICML Workshop on Unsupervised and Transfer Learning, Proceedings of Machine Learning Research, vol. 27, pp. 37–49. PMLR, Bellevue, Washington, USA (2012)
- I. Goodfellow et al., ‘‘Generative adversarial nets,’’ in Proc. Adv. Neural Inf. Process. Syst., 2014, pp. 2672–2680. [Online]. Available: http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In Advances in Neural Information Processing Systems, pages 6000–6010.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805, 2018.