CC BY

A Brief Timeline

Paulo Ricardo Lisboa de Almeida

DSBD

UFPR

Timeline

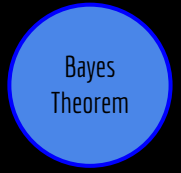
1763

Bayes Theorem

1763 - Thomas Bayes.
Introduced Bayes' Theorem, including the concepts of prior and posterior probabilities.

Timeline

1763



Timeline

1763

Bayes Theorem

1943 - Warren McCulloch & Walter Pitts. Human neuron modeled as an electronic perceptron.

Timeline

1763

Bayes Theorem

1943

Electronic Neuron

Timeline

1763

Bayes Theorem

1943

Electronic Neuron

1957 - Frank Rosenblatt. Implementation of a perceptron on an IBM 704 computer.

Timeline

1763

Bayes Theorem

1943

Electronic Neuron

1957

Rosenblatt's Perceptron

Timeline

1763

Bayes Theorem

1600~1700

Calculus

~1636 - Pierre de Fermat.
Initial ideas of infinitesimal calculus used by Newton and Leibniz.

1943

Electronic Neuron

1957

Rosenblatt's Perceptron

Timeline

1763

Bayes Theorem

1600~1700

Calculus

~1636 - Pierre de Fermat.
Initial ideas of infinitesimal calculus used by Newton and Leibniz.

1943

Electronic Neuron

1957

Rosenblatt's Perceptron

~1687 - Isaac Newton & Gottfried Wilhelm Leibniz.
Creation of modern differential and integral calculus, including its rules and some classical formulas, such as the chain rule.

Timeline

1763

Bayes Theorem

1600~1700

Calculus

1943

Electronic Neuron

1957

Rosenblatt's Perceptron

Timeline

1763

Bayes Theorem

1600~1700

Calculus

1969 - Kunihiko Fukushima proposes the ReLU activation.

1943

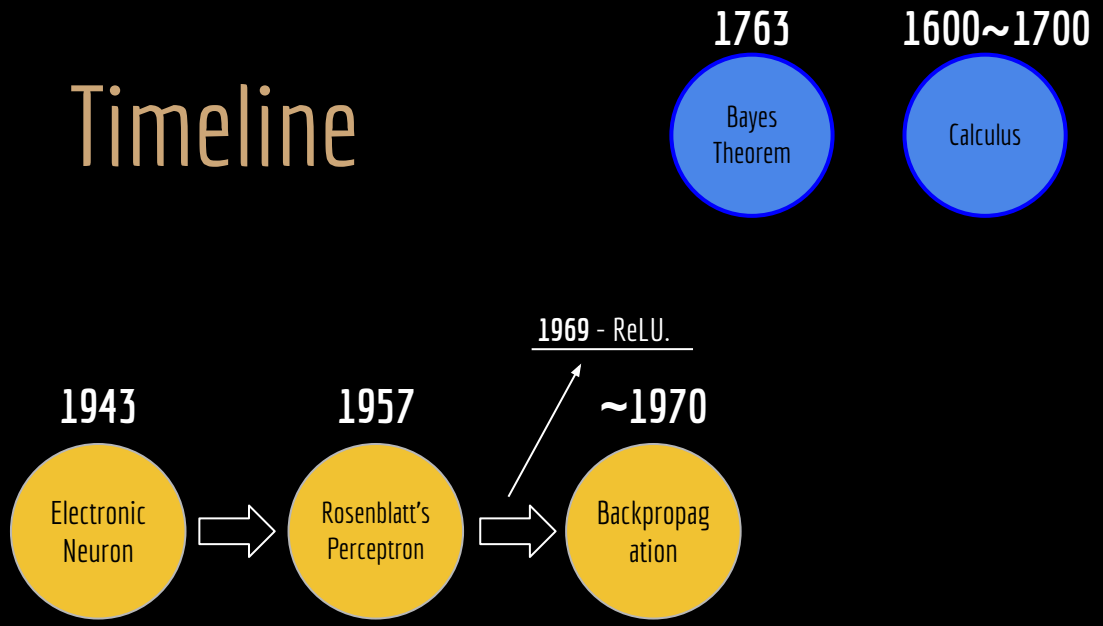
Electronic Neuron

1957

Rosenblatt's Perceptron

1970/1971 - Seppo Linnainmaa & Paul Werbos. Development of backpropagation.

Timeline


1763
Bayes Theorem
1600~1700
Calculus
1969 - ReLU.
1943
Electronic Neuron
1957
Rosenblatt's Perceptron
~1970
Backpropag ation

Timeline

1763
Bayes Theorem

1600~1700
Calculus

1943
Electronic Neuron

1957
Rosenblatt's Perceptron

1969 - ReLU.

~1970
Backpropagation

1989 - Yann LeCun.
CNN to recognize digits without manual tuning.

Timeline

1763
Bayes Theorem

1600~1700
Calculus

1943
Electronic Neuron

1957
Rosenblatt's Perceptron

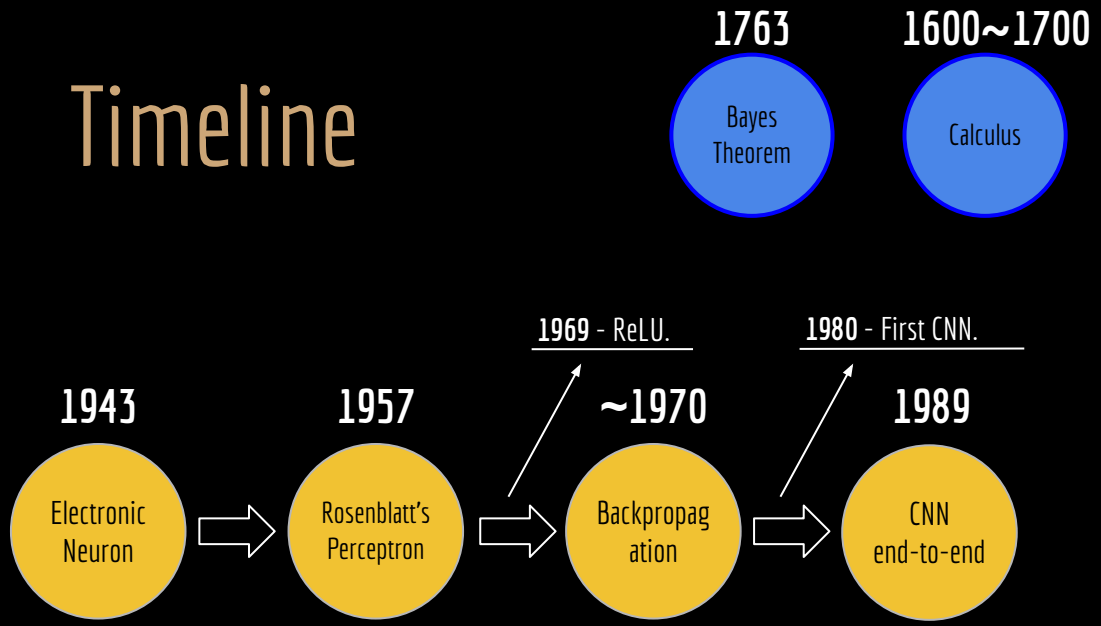
~1970
Backpropagation

1969 - ReLU.

1980 - Kunihiko Fukushima proposes the Neocognitron, the first CNN with convolution and pooling layers.

1989 - Yann LeCun.
CNN to recognize digits without manual tuning.

Timeline


1763
Bayes Theorem
1600~1700
Calculus
1943
Electronic Neuron
1957
Rosenblatt's Perceptron
1969 - ReLU.
~1970
Backpropag ation
1980 - First CNN.
1989
CNN end-to-end

Timeline

1763
Bayes Theorem

1600~1700
Calculus

1943
Electronic Neuron

1957
Rosenblatt's Perceptron

1969 - ReLU.

~1970
Backpropagation

1980 - First CNN.

1989
CNN end-to-end

2012 - Alex Krizhevsky, Ilya Sutskever & Geoffrey Hinton.
AlexNet wins the ImageNet competition.

Timeline

1763 Bayes Theorem

1600~1700 Calculus

1943 Electronic Neuron

1957 Rosenblatt's Perceptron

~1970 Backpropagation

1969 - ReLU.

1989 CNN end-to-end

1980 - First CNN.

2009 - Rajat Raina, Anand Madhavan, and Andrew Y. Ng use 30 GTX 280 GPUs to train a 100M-parameter model.

2012 - Alex Krizhevsky, Ilya Sutskever & Geoffrey Hinton.
AlexNet wins the ImageNet competition.

Timeline

1763
Bayes Theorem

1600~1700
Calculus

2011 - Xavier Glorot, Antoine Bordes, and Yoshua Bengio discover that ReLU works well in deep networks.

2009 - Rajat Raina, Anand Madhavan, and Andrew Y. Ng use 30 GTX 280 GPUs to train a 100M-parameter model.

1969 - ReLU.

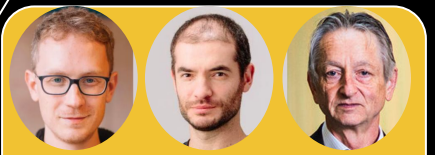
1980 - First CNN.

1943
Electronic Neuron

1957
Rosenblatt's Perceptron

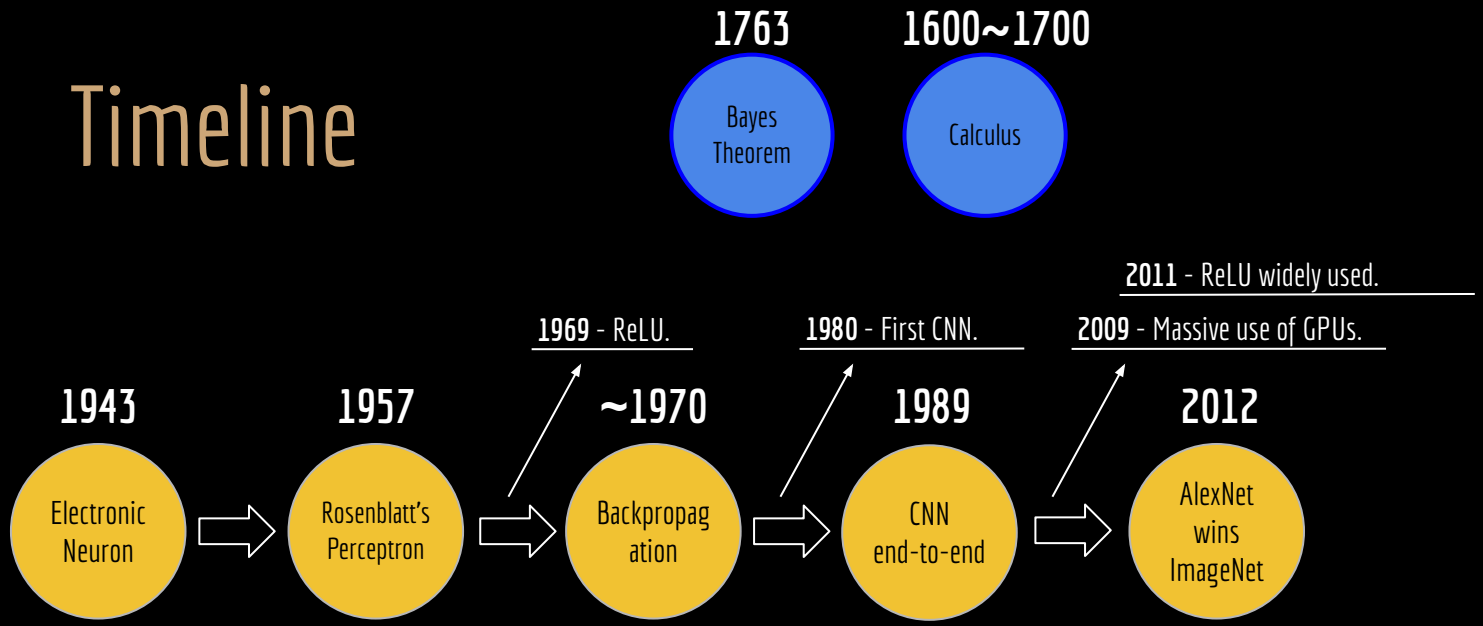
~1970
Backpropagation

1989
CNN end-to-end



2012 - Alex Krizhevsky, Ilya Sutskever & Geoffrey Hinton.
AlexNet wins the ImageNet competition.

Timeline


1763
Bayes Theorem
1600~1700
Calculus
1943
Electronic Neuron
1957
Rosenblatt's Perceptron
~1970
Backpropag ation
1969 - ReLU.
1989
CNN end-to-end
1980 - First CNN.
2012
AlexNet wins ImageNet
2009 - Massive use of GPUs.
2011 - ReLU widely used.

Timeline

1763 Bayes Theorem

1600~1700 Calculus

~1948 Communication Theory

~1948 - Claude Shannon.
Paper *A Mathematical Theory of Communication.*

1943 Electronic Neuron

1957 Rosenblatt's Perceptron

1969 - ReLU.

~1970 Backpropagation

1980 - First CNN.

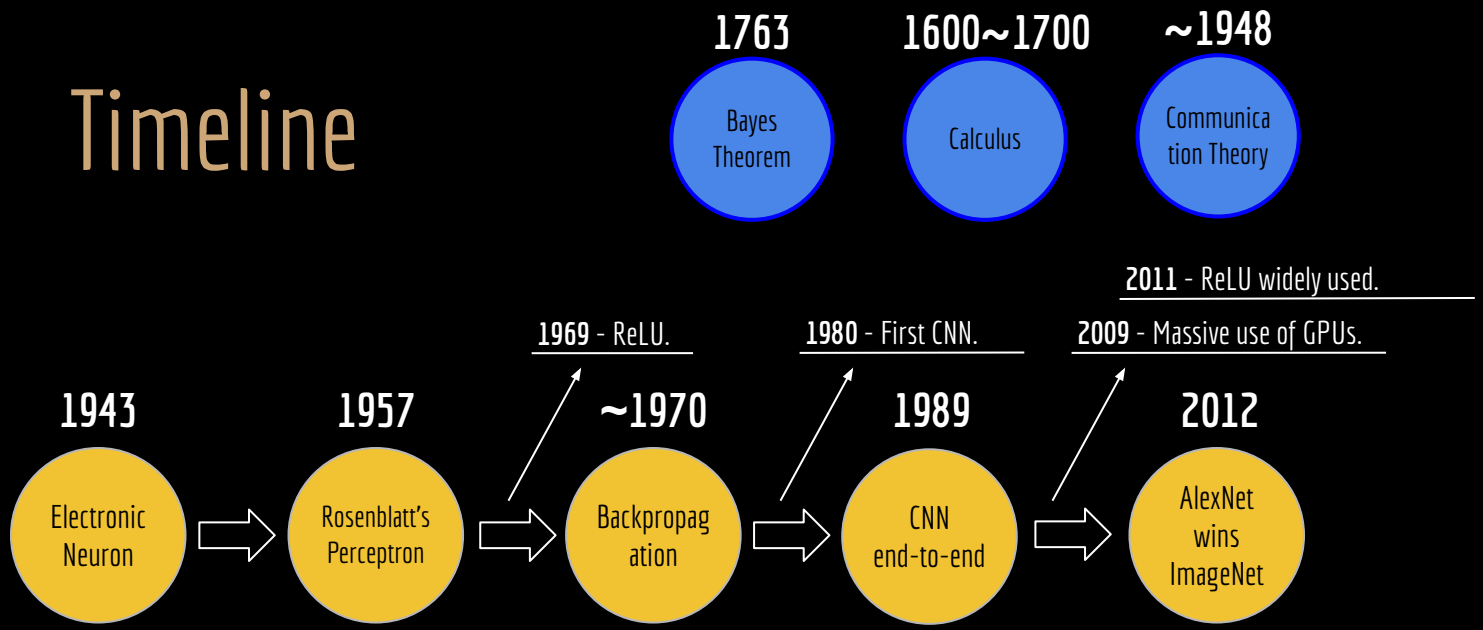
1989 CNN end-to-end

2009 - Massive use of GPUs.

2012 AlexNet wins ImageNet

2011 - ReLU widely used.

Timeline


1763
Bayes Theorem
1600~1700
Calculus
~1948
Communication Theory
2011 - ReLU widely used.
1969 - ReLU.
1980 - First CNN.
2009 - Massive use of GPUs.
1943
Electronic Neuron
1957
Rosenblatt's Perceptron
~1970
Backpropagation
1989
CNN end-to-end
2012
AlexNet wins ImageNet

Timeline

1763 Bayes Theorem

1600~1700 Calculus

~1948 Communication Theory

1950 The Imitation Game

1950 - Alan Turing.
Proposed the Imitation Game (the Turing Test) to assess whether machines exhibit human-like behavior.

1943 Electronic Neuron

1957 Rosenblatt's Perceptron

1969 - ReLU.

~1970 Backpropagation

1980 - First CNN.

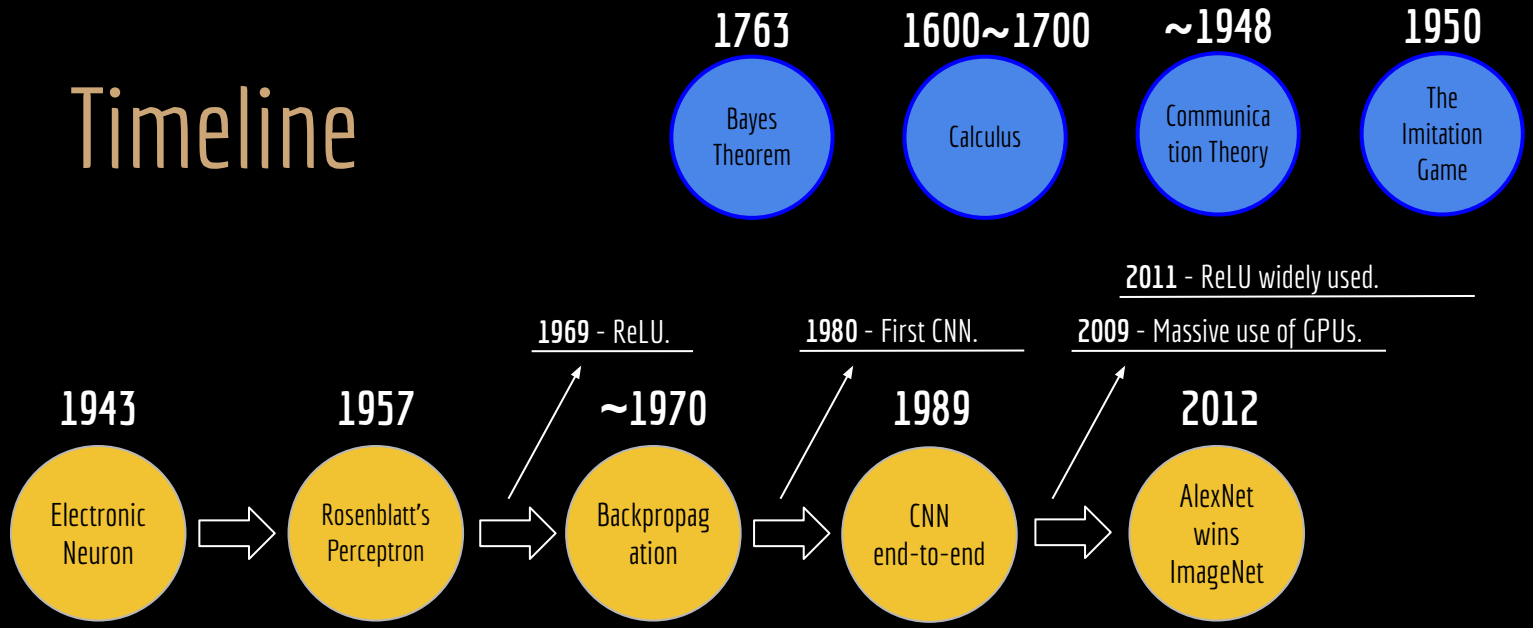
1989 CNN end-to-end

2009 - Massive use of GPUs.

2012 AlexNet wins ImageNet

2011 - ReLU widely used.

Timeline


1763
Bayes Theorem
1600~1700
Calculus
~1948
Communication Theory
1950
The Imitation Game
2011 - ReLU widely used.
1969 - ReLU.
1980 - First CNN.
2009 - Massive use of GPUs.
1943
Electronic Neuron
1957
Rosenblatt's Perceptron
~1970
Backpropagation
1989
CNN end-to-end
2012
AlexNet wins ImageNet

Timeline

1763 Bayes Theorem

1600~1700 Calculus

~1948 Communication Theory

1950 The Imitation Game

1943 Electronic Neuron

1957 Rosenblatt's Perceptron

~1970 Backpropagation

1969 - ReLU.

1989 CNN end-to-end

1980 - First CNN.

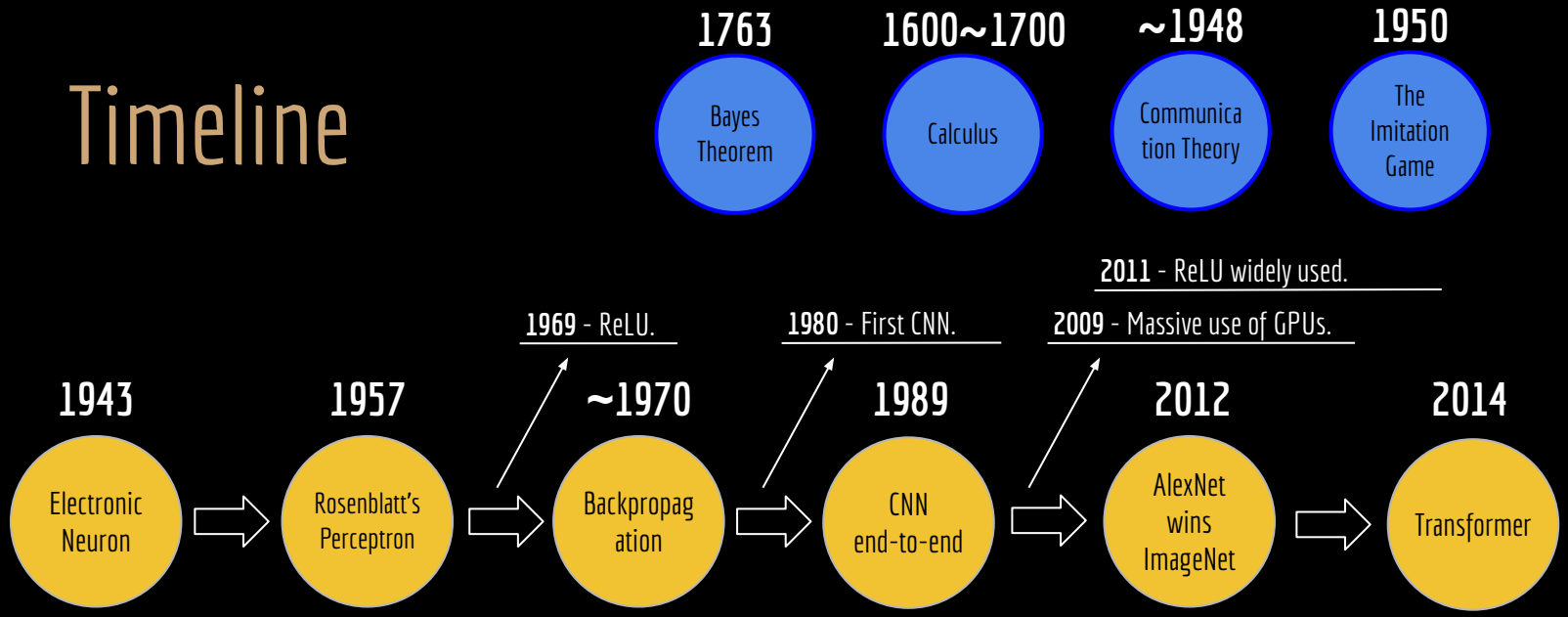
2012 AlexNet wins ImageNet

2009 - Massive use of GPUs.

2011 - ReLU widely used.

2014 - Dzmitry Bahdanau, KyungHyun Cho & Yoshua Bengio
Transformers for text translation.

Timeline


1763
Bayes Theorem
1600~1700
Calculus
~1948
Communication Theory
1950
The Imitation Game
2011 - ReLU widely used.
1969 - ReLU.
1980 - First CNN.
2009 - Massive use of GPUs.
1943
Electronic Neuron
1957
Rosenblatt's Perceptron
~1970
Backpropagation
1989
CNN end-to-end
2012
AlexNet wins ImageNet
2014
Transformer

Timeline

1763 Bayes Theorem

1600~1700 Calculus

~1948 Communication Theory

1950 The Imitation Game

1943 Electronic Neuron

1957 Rosenblatt's Perceptron

1969 - ReLU.

~1970 Backpropagation

1980 - First CNN.

1989 CNN end-to-end

2009 - Massive use of GPUs.

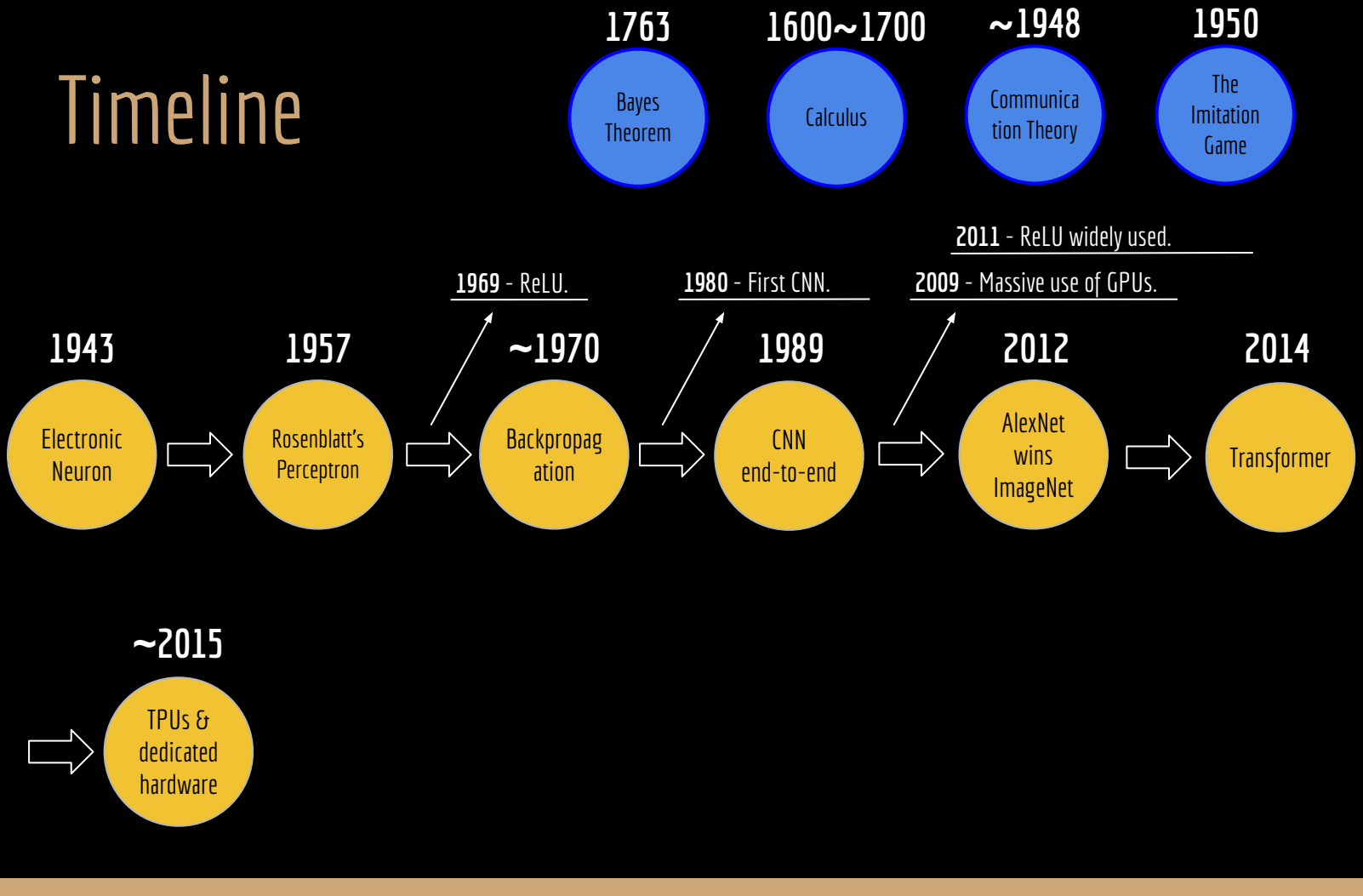
2011 - ReLU widely used.

2012 AlexNet wins ImageNet

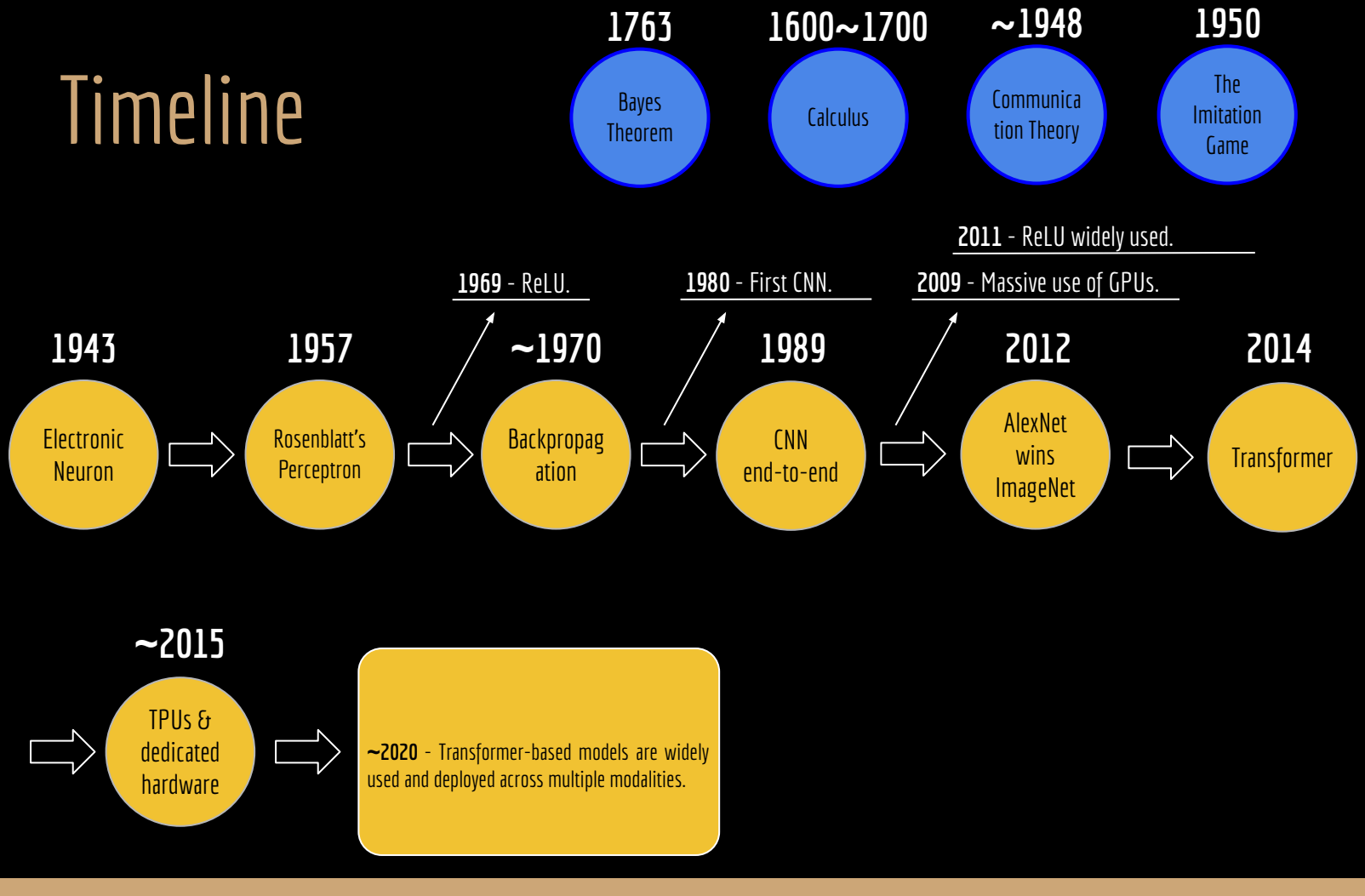
2014 Transformer

2015 - Patterson et. al.
Specialized hardwares, such as TPUs.

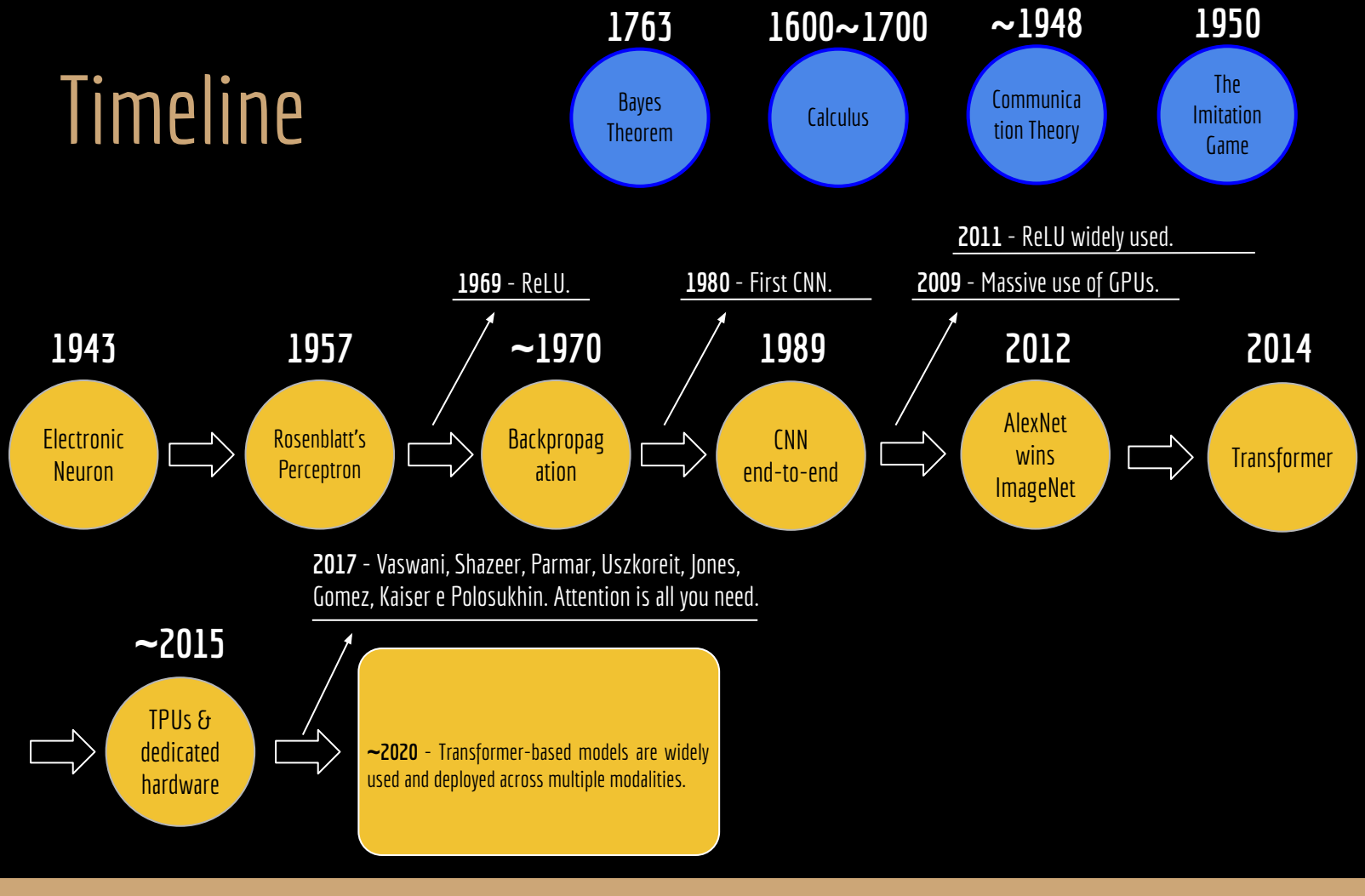
Timeline


1763
Bayes Theorem
1600~1700
Calculus
~1948
Communication Theory
1950
The Imitation Game
2011 - ReLU widely used.
1969 - ReLU.
1980 - First CNN.
2009 - Massive use of GPUs.
1943
Electronic Neuron
1957
Rosenblatt's Perceptron
~1970
Backpropagation
1989
CNN end-to-end
2012
AlexNet wins ImageNet
2014
Transformer
~2015
TPUs & dedicated hardware

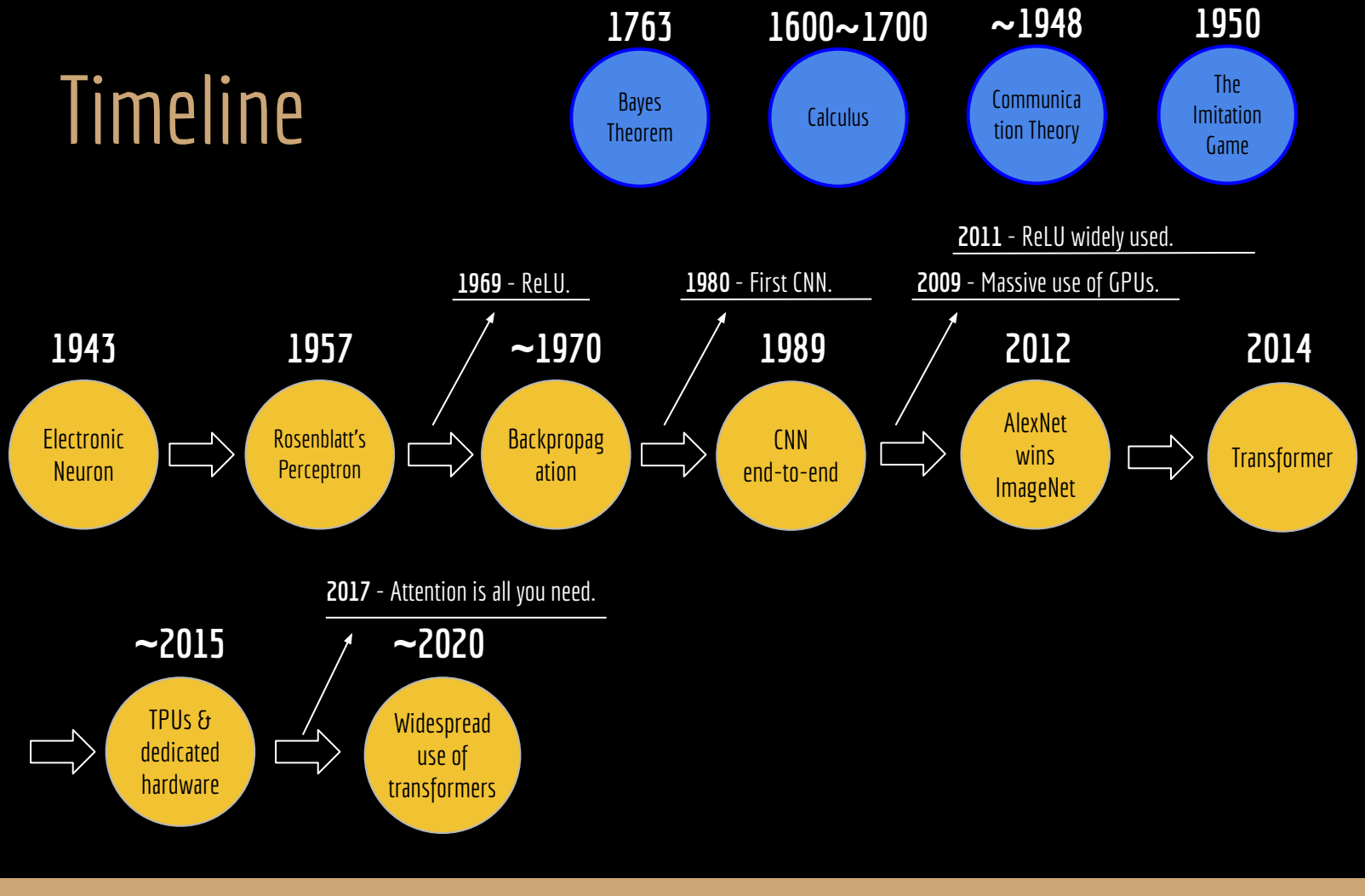
Timeline


1763
Bayes Theorem
1600~1700
Calculus
~1948
Communication Theory
1950
The Imitation Game
2011 - ReLU widely used.
1969 - ReLU.
1980 - First CNN.
2009 - Massive use of GPUs.
1943
Electronic Neuron
1957
Rosenblatt's Perceptron
~1970
Backpropagation
1989
CNN end-to-end
2012
AlexNet wins ImageNet
2014
Transformer
~2015
TPUs & dedicated hardware
~2020 - Transformer-based models are widely used and deployed across multiple modalities.

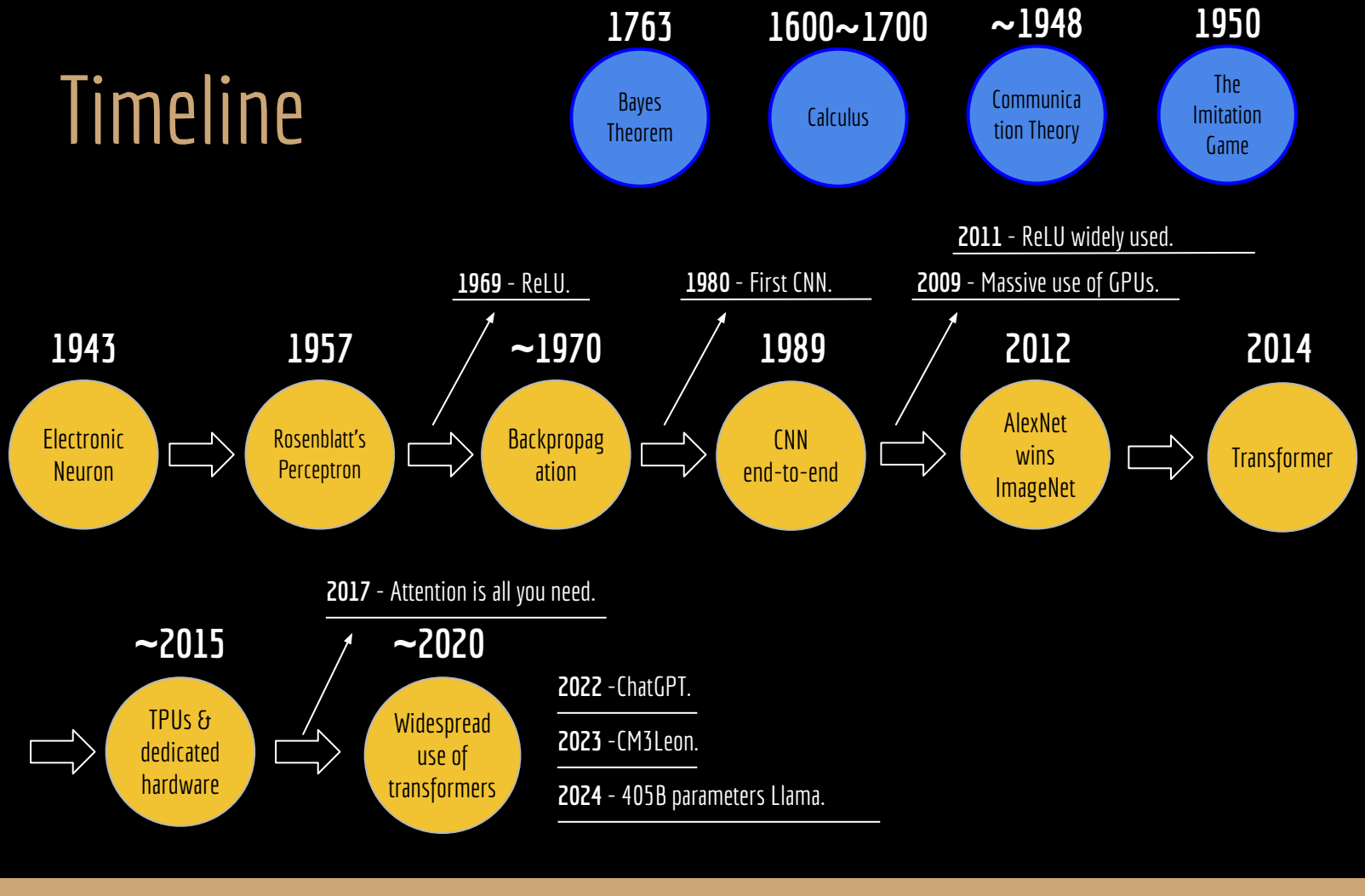
Timeline


1763
Bayes Theorem
1600~1700
Calculus
~1948
Communication Theory
1950
The Imitation Game
2011 - ReLU widely used.
1969 - ReLU.
1980 - First CNN.
2009 - Massive use of GPUs.
1943
Electronic Neuron
1957
Rosenblatt's Perceptron
~1970
Backpropagation
1989
CNN end-to-end
2012
AlexNet wins ImageNet
2014
Transformer
2017 - Vaswani, Shazeer, Parmar, Uszkoreit, Jones, Gomez, Kaiser e Polosukhin. Attention is all you need.
~2015
TPUs & dedicated hardware
~2020 - Transformer-based models are widely used and deployed across multiple modalities.

Timeline


1763
Bayes Theorem
1600~1700
Calculus
~1948
Communication Theory
1950
The Imitation Game
1943
Electronic Neuron
1957
Rosenblatt's Perceptron
1969 - ReLU.
~1970
Backpropagation
1980 - First CNN.
1989
CNN end-to-end
2009 - Massive use of GPUs.
2011 - ReLU widely used.
2012
AlexNet wins ImageNet
2014
Transformer
~2015
TPUs & dedicated hardware
2017 - Attention is all you need.
~2020
Widespread use of transformers

Timeline


1763
Bayes Theorem
1600~1700
Calculus
~1948
Communication Theory
1950
The Imitation Game
2011 - ReLU widely used.
1969 - ReLU.
1980 - First CNN.
2009 - Massive use of GPUs.
1943
Electronic Neuron
1957
Rosenblatt's Perceptron
~1970
Backpropagation
1989
CNN end-to-end
2012
AlexNet wins ImageNet
2014
Transformer
2017 - Attention is all you need.
~2015
TPUs & dedicated hardware
~2020
Widespread use of transformers
2022 -ChatGPT.
2023 -CM3Leon.
2024 - 405B parameters Llama.

1763 Bayes Theorem

1600~1700 Calculus

~1948 Communication Theory

1950 The Imitation Game

1943 Electronic Neuron

1957 Rosenblatt's Perceptron

1969 - ReLU.

~1970 Backpropagation

1980 - First CNN.

1989 CNN end-to-end

2009 - Massive use of GPUs.

2011 - ReLU widely used.

2012 AlexNet wins ImageNet

2014 Transformer

2017 - Attention is all you need.

~2015 TPUs & dedicated hardware

~2020 Widespread use of transformers

2022 -ChatGPT.

2023 -CM3Leon.

2024 - 405B parameters Llama.

Some questions for › 2025:

How to reduce the computational cost?

How to learn continuously?

How to embed reasoning into models?

How to model causality?

...

Licence

This work is licensed under CC BY 4.0.