



# Deep Learning

## Chapter 1: History of Neural Networks

**Mina Rezaei**

Department of Statistics – LMU Munich

Winter Semester 2020

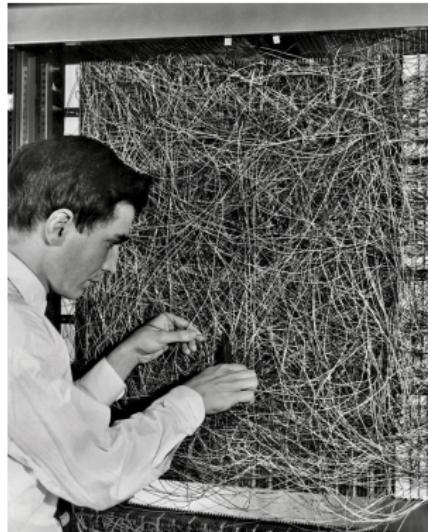


# A BRIEF HISTORY OF NEURAL NETWORKS

- **1943:** The first artificial neuron, the "Threshold Logic Unit (TLU)", was proposed by Warren McCulloch & Walter Pitts.
  - In this model the neuron fires a  $+1$  if the input exceeds a certain threshold  $\phi$ .
  - However, this model did not have adjustable weights, so learning could only be achieved by changing the threshold  $\phi$ .

# A BRIEF HISTORY OF NEURAL NETWORKS

- 1957: The perceptron was invented by Frank Rosenblatt.



**Figure:** Frank Rosenblatt with a Mark I Perceptron machine, the first implementation of the perceptron algorithm in 1960.  
(<https://en.wikipedia.org/wiki/Perceptron>)

# A BRIEF HISTORY OF NEURAL NETWORKS

- **1960:** ADALINE by Bernard Widrow & Tef Hodd.
  - Weights are now adjusted according to the weighted sum of the inputs.
- **1965:** Group method of data handling (also known as polynomial neural networks) by Alexey Ivakhnenko. The first learning algorithm for supervised deep feedforward multilayer perceptrons.
- **1969:** The first “AI Winter” kicked in.
  - Marvin Minsky & Seymour Papert proved that a perceptron cannot solve the XOR-Problem (linear separability).
  - Less funding ⇒ Standstill in AI/DL research

# A BRIEF HISTORY OF NEURAL NETWORKS

- **1985:** Multi-layered perceptron with backpropagation by David Rumelhart, Geoffrey Hinton and Ronald Williams.
  - Method to efficiently compute derivatives of differentiable composite functions.
  - Backpropagation was developed already in 1970 by Seppo Linnainmaa.
- **1985:** The second “AI Winter” kicked in.
  - Overly optimistic/exaggerated expectations concerning potential of AI/DL.
  - Angering investors, the phrase “AI” even reached a pseudoscience status.
  - Kernel machines and graphical models both achieved good results on many important tasks.
  - Some of the fundamental mathematical difficulties in modeling long sequences were identified.

# A BRIEF HISTORY OF NEURAL NETWORKS

- **2006:** Age of deep neural networks began.
  - Geoffrey Hinton showed that a deep belief network could be efficiently trained using a strategy called greedy layer-wise pretraining.
  - This wave of neural networks research popularized the use of the term deep learning to emphasize that researchers were now able to train deeper neural networks than had been possible before.
  - At this time, deep neural networks outperformed competing AI systems based on other machine learning technologies as well as hand-designed functionality.

# A BRIEF HISTORY OF NEURAL NETWORKS

- Why now and not earlier?
  - Significantly bigger datasets.
  - Better algorithms (optimization chapter).
    - Vanishing gradient problem (ReLU).
  - Better regularization (regularization chapter).
  - Unsupervised representation learning (autoencoder chapter).
- More layers inevitably lead to a significant increase of parameters.

# A BRIEF HISTORY OF NEURAL NETWORKS

- Back then, processing power was simply not capable to handle such huge amounts of parameters.  
⇒ Nowadays, deep neural networks are trained on GPUs (graphic processing units), not on CPUs (central processing units).
- Investment by industries and universities.
- Deep learning tools make learning and applying deep learning easier.

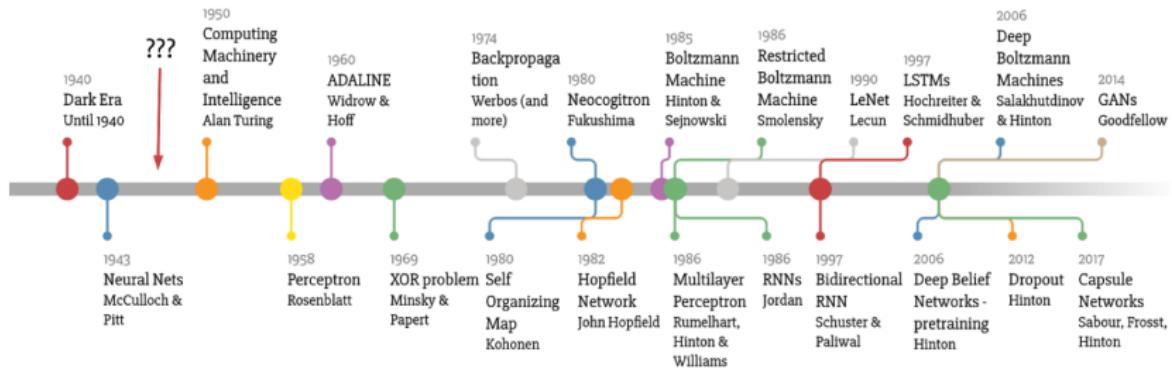
# A BRIEF HISTORY OF NEURAL NETWORKS

## History of DL Tools

- 1960: Mark 1 Perceptron
- 2002: Torch
- 2007: CUDA
- 2008: Theano
- 2014: Caffe
- 2015: TensorFlow 0.1
- 2017: PyTorch 0.1
- 2017: MXNet
- 2017: Caffe 2.0
- 2019: TensorFlow 2.0

# A BRIEF HISTORY OF NEURAL NETWORKS

## Deep Learning Timeline



**Figure:** Credit: Favio Vazquez

# A BRIEF HISTORY OF NEURAL NETWORKS



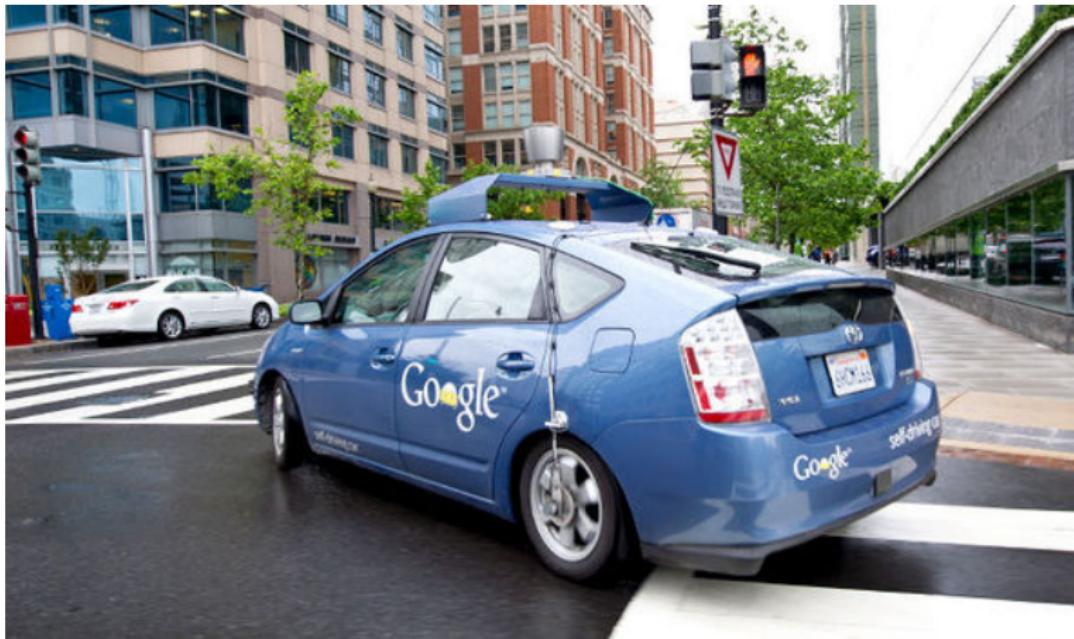
**Figure:** Boston Dynamic

# A BRIEF HISTORY OF NEURAL NETWORKS



**Figure:** IBM Supercomputer

# A BRIEF HISTORY OF NEURAL NETWORKS



**Figure:** Google Self driving car (Waymo)