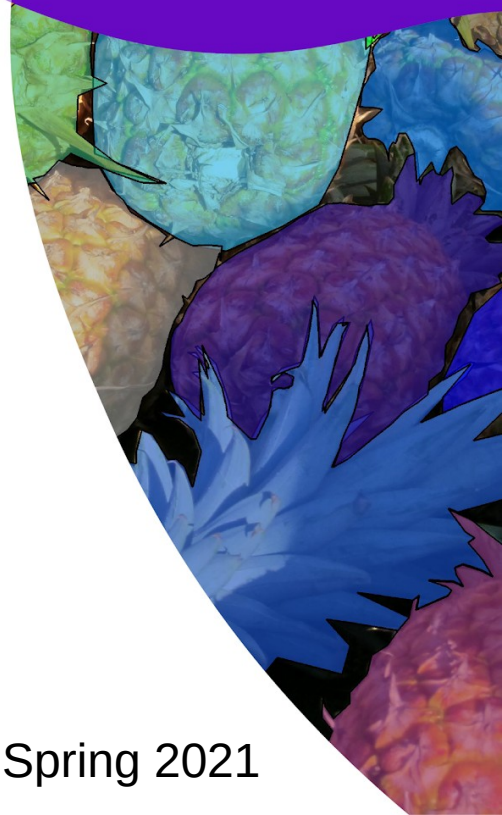
NEW YORK UNIVERSITY

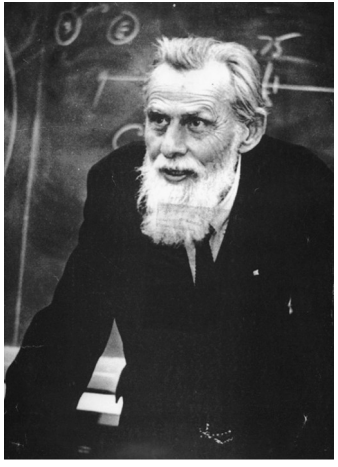
# Deep Learning

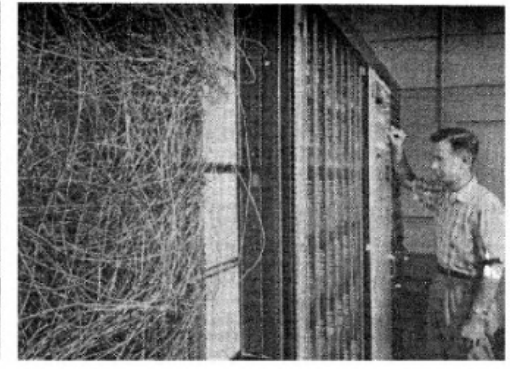


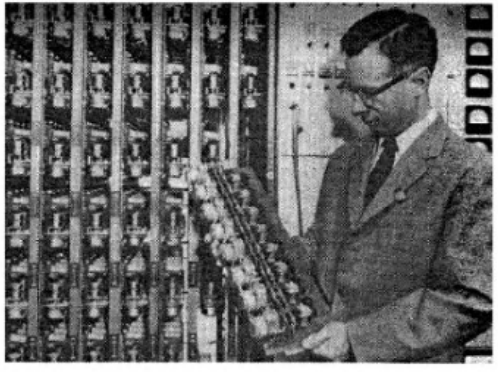
Alfredo Canziani, Yann LeCun
NYU - Courant Institute & Center for Data Science

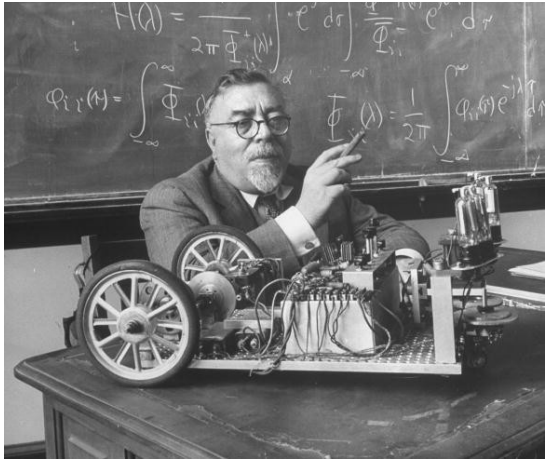
# Inspiration for Deep Learning: The Brain!

- 1943: McCulloch & Pitts, networks of binary neurons can do logic
- 1947: Donald Hebb, Hebbian synaptic plasticity
- 1948: Norbert Wiener, cybernetics, optimal filter, feedback, autopoïesis, auto-organization.
- 1957: Frank Rosenblatt, Perceptron
- 1961: Bernie Widrow, Adaline
- 1962: Hubel & Wiesel, visual cortex architecture
- 1969: Minsky & Papert, limits of the Perceptron





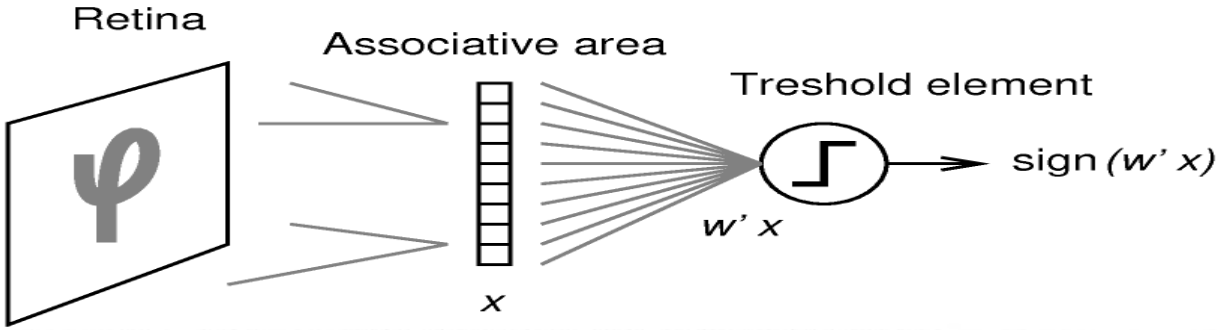


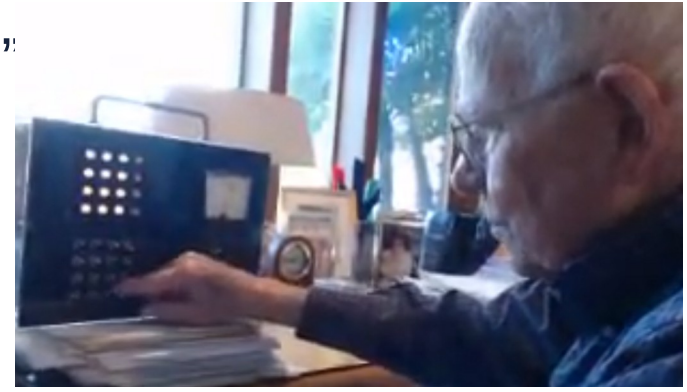


# Supervised Learning goes back to the Perceptron & Adaline

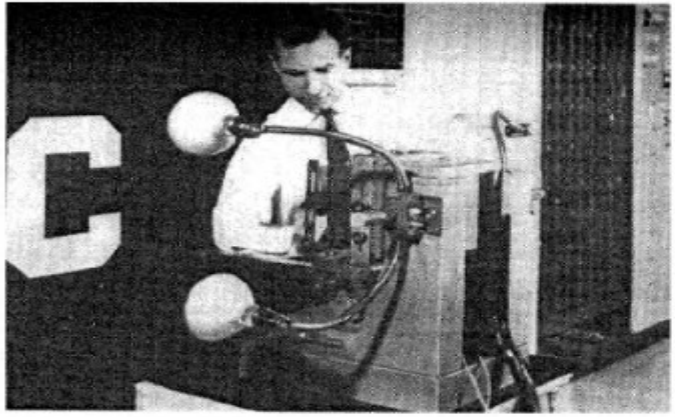
- **The McCulloch-Pitts Binary Neuron**
  - Perceptron: weights are motorized potentiometers
  - Adaline: Weights are electrochemical "memistors"

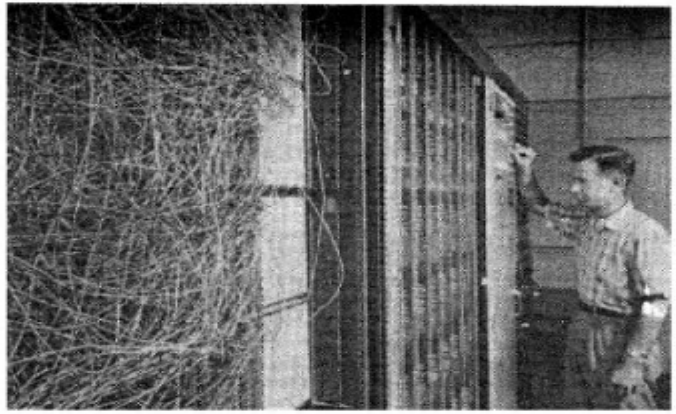
$$y = sign\left(\sum_{i=1}^{N} W_i X_i + b\right)$$

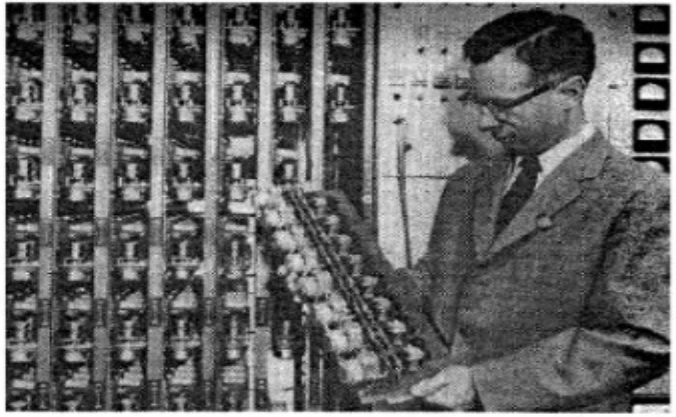


https://youtu.be/X1G2g3SiCwU







# More History

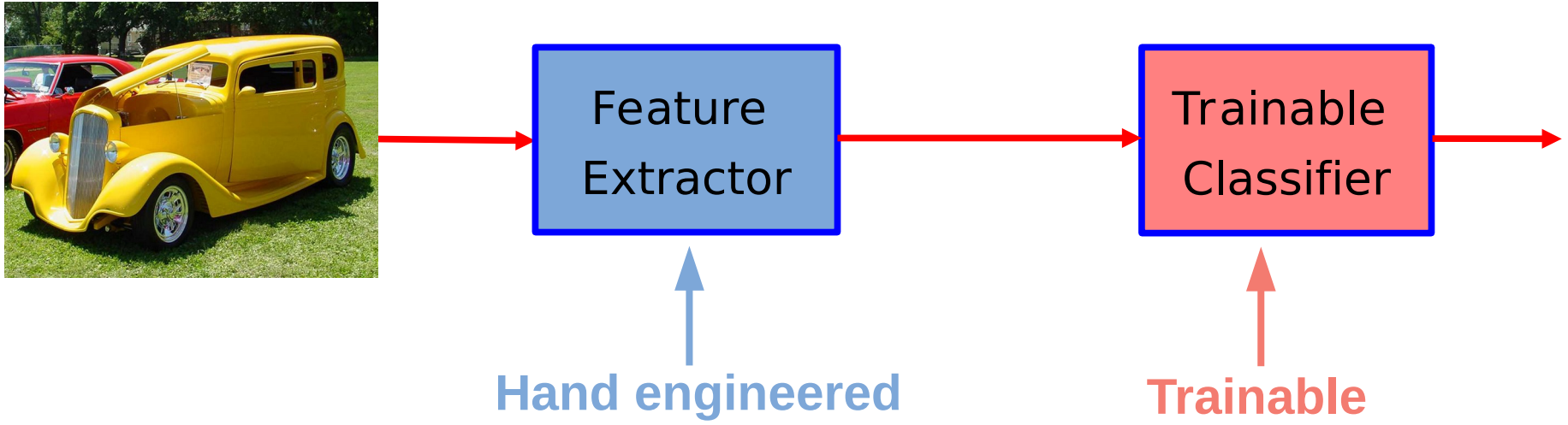
- 1970s: statistical patter recognition (Duda & Hart 1973)
- 1979: Kunihiko Fukushima, Neocognitron
- 1982: Hopfield Networks
- 1983: Hinton & Sejnowski, Boltzmann Machines
- 1985/1986: Practical Backpropagation for neural net training
- 1989: Convolutional Networks
- 1991: Bottou & Gallinari, module-based automatic differentiation
- 1995: Hochreiter & Schmidhuber, LSTM recurrent net.
- 1996: structured prediction with neural nets, graph transformer nets
- …..
- 2003: Yoshua Bengio, neural language model
- 2006: Layer-wise unsupervised pre-training of deep networks
- 2010: Collobert & Weston, self-supervised neural nets in NLP

# More History

- **2012: AlexNet / convnet on GPU / object classification**
- **2015: I. Sutskever, neural machine translation with multilayer LSTM**
- **2015: Weston, Chopra, Bordes: Memory Networks**
- **2016: Bahdanau, Cho, Bengio: GRU, attention mechanism**
- **2016: Kaiming He, ResNet**

# The Standard Paradigm of Pattern Recognition

- **...since the 1960s**
- **...and "traditional" Machine Learning**
  - until the "Deep Learning Revolution" (circa 2012)



# Multilayer Neural Nets and Deep Learning

## Traditional Machine Learning

Feature Extractor

Trainable Classifier

Hand engineered

Trainable

## Deep Learning

Trainable

Low-Level Features

Mid-Level Features

High-Level Features

Trainable Classifier