Historical Trends in Deep Learning

Sargur N. Srihari

srihari@cedar.buffalo.edu

This is part of lecture slides on Deep Learning:
http://www.cedar.buffalo.edu/~srihari/CSE676
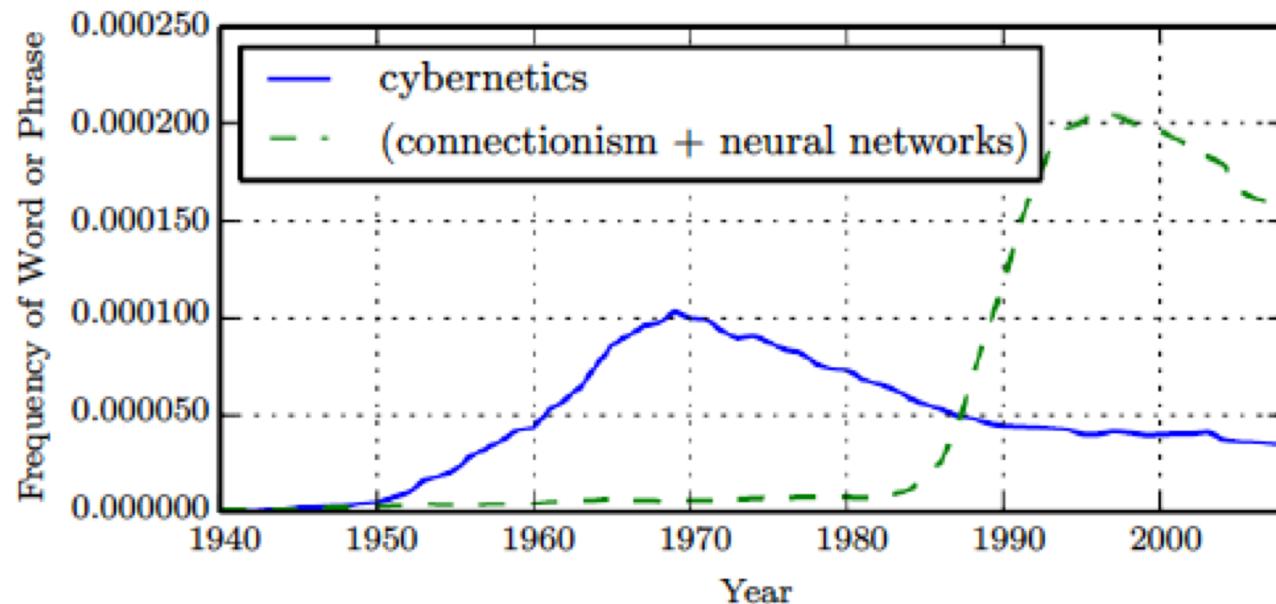
Logic systems vs Deep learning

- The difference between failed logic systems and deep learning is *scale*

- The connectionists' neural networks, unlike the rules-based, logic-based approach, were able to scale up to larger and larger problems as computers got more and more powerful and data more plentiful

Time-frame of growth

- People who went for logic had fifty years to show that it didn't scale

- And now, we had thirty years, from the eighties to today, to show that connectionism does scale

- Here, at least with some patterns, with pattern recognition, with reinforcement learning and so forth, we have something that scales

History of Deep Learning

- ## Dates back to 1940s

- ## Three historical waves:
  - Cybernetics, peaked in 1970
  - connectionism/neural networks, peaked in 1995
  - Deep learning, 2006+ (layerwise training, big data)

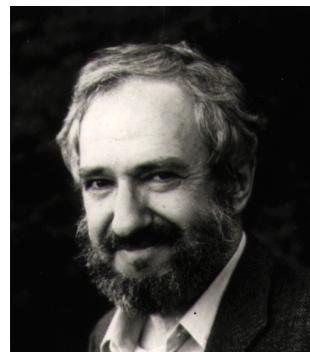# Beginnings of Deep Learning

Perceptron
Invented at Calspan Buffalo, NY
Rosenblatt, Frank,
The Perceptron--a perceiving and
recognizing automaton.
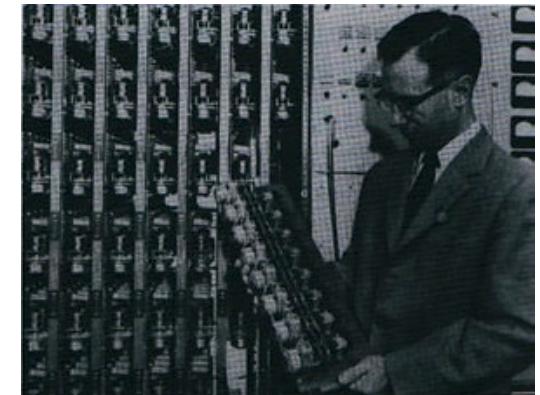Report 85-460-1, 1957
Cornell Aeronautical Laboratory

Minsky and Papert
dedicated book to him
Minsky M. L. and Papert S. A. 1969
Perceptrons, MIT Press

Perceptron Hardware (Analog)

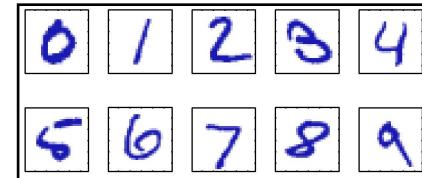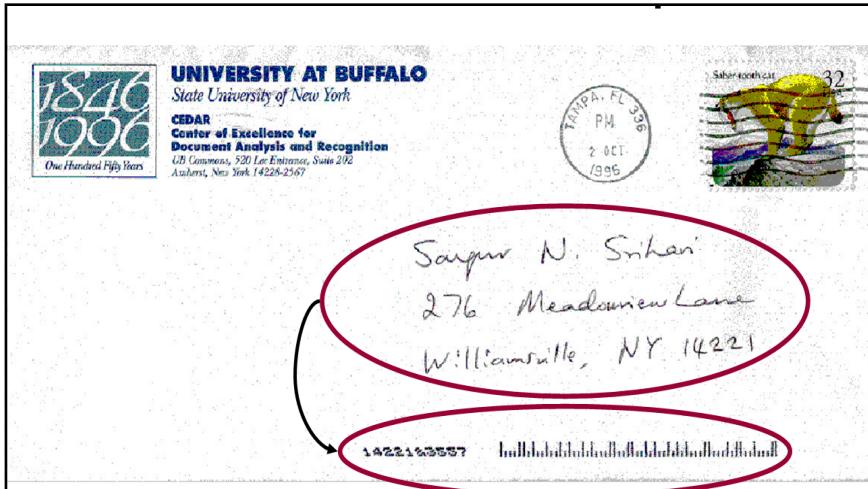**Learning to discriminate shapes of characters**
20x20 cell
Image of character

Patch-board
to allow
different
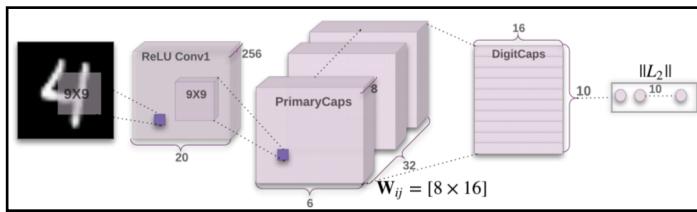configurations of
input features φ

Racks of
Adaptive Weights
Implemented
as potentiometers

Known as Mark 1 Perceptron. It is now in the Smithsonian

6

Handwriting recognition as "fruit-fly of AI"

- Many handcrafted rules and exceptions
- Better learn from training set
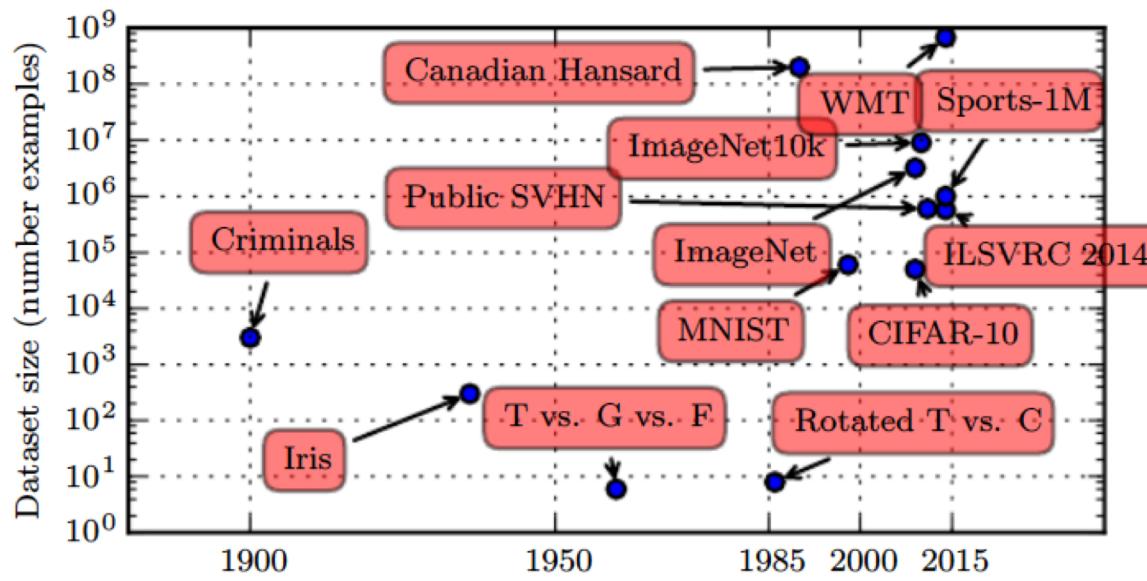- Handwriting rec cannot be done without ML!

NYU, Toronto, Montreal, Baidu

Why this resurgence?

- Data set sizes needed to succeed
- Increasing computational resources (model sizes)
  - Dense connectivity
  - Increased no. of neurons

Large Data Sets

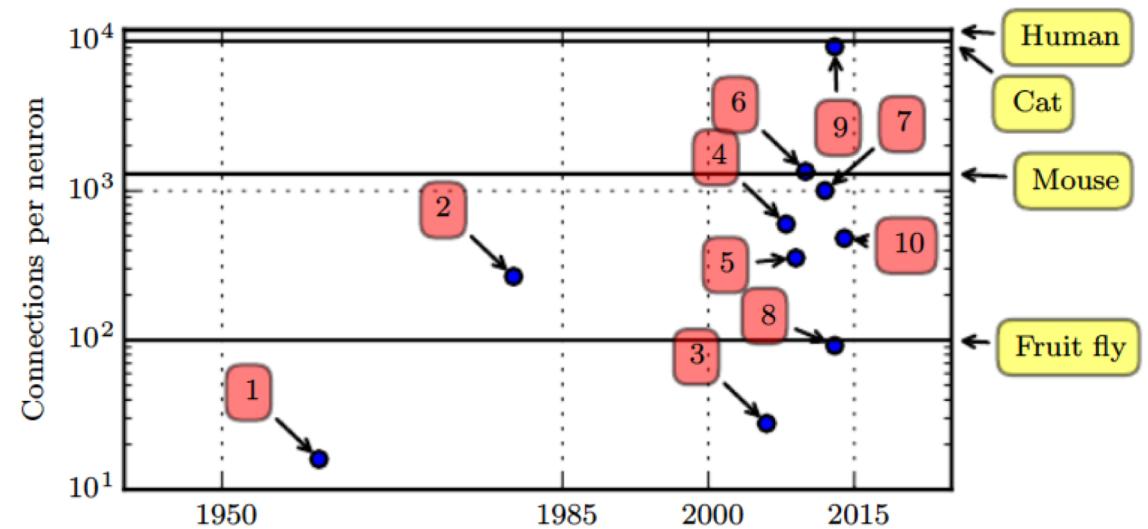- Algorithms can be provided with data sets needed to succeed

Increased Connectivity

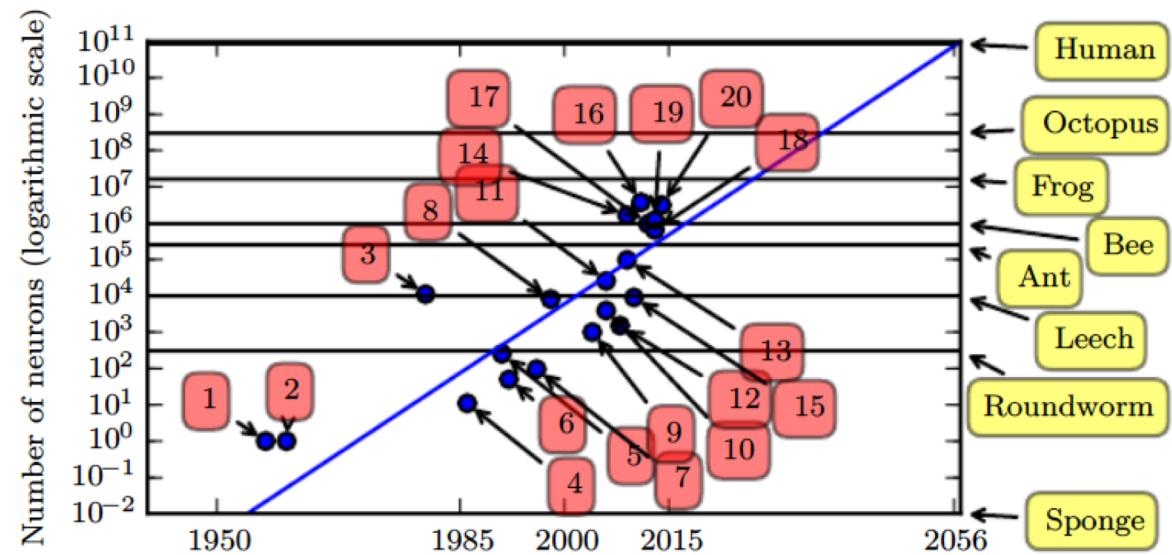– Intelligence achieved when many neurons work together

1. Adaptive Linear element (Adaline), 1960
2. Neocognitron
3. GPU CNN
4. Deep Boltzmann
5. Unsup. CNN
6. GPU MLP
7. Distributed Autc
8. Multi-GPU CNN
9. COTS Unsupervised CNN
10. GoogLeNet
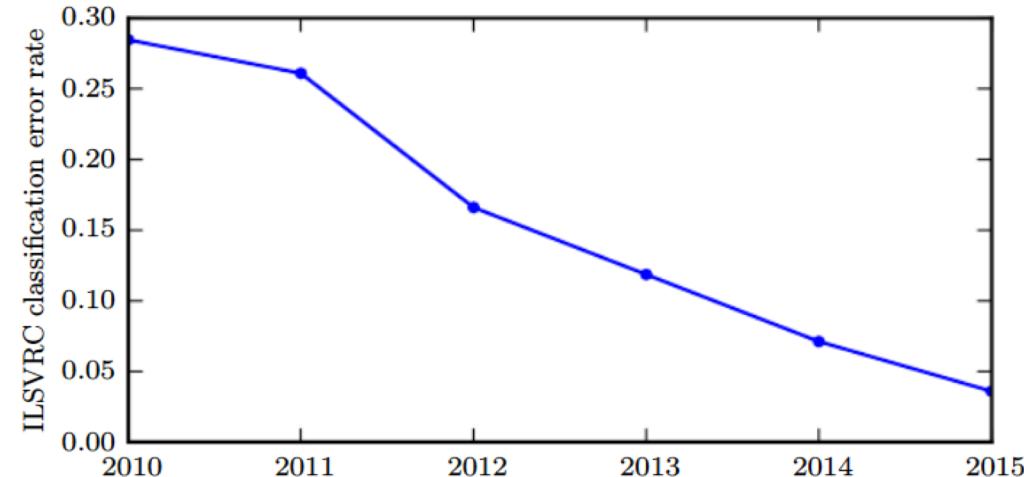

10

Increased no. of neurons

- Since introduction of hidden units, doubling of neurons every 2.4 years

1. Perceptron
2. Adaline
3. Neocognitron
4. Early backpro
5. RNN speech
6. MLP speech
7. Mean-field Belief Net
8. LeNet5
9. Echo-state network
10. Deep Belief Net
11. GPU CNN

11

Increasing Accuracy

- Image Net Large Scale Visual Recognition Challenge

ILSVRC 2012 tasks

1. Classification
2. Classification with localization
3. Fine-grained classification

IMAGENET CHALLENGE 2017
Identify and label objects in images and videos.