BIA-667-A: Introduction to Deep Learning and Business Applications

References:

deeplearningbook.org
Dive into Deep Learning, https://d2l.ai/index.html
introtodeeplearning.com
deeplearningindaba.com

Agenda

- Introduction
  - Tell us about yourself, e.g. major, interest, expectation from this class, …
- Syllabus
- Hands-on Labs
  - Try it yourself!
  - Recitation schedule
  - Homework assignments will be similar to labs
- Poll Everywhere
- Introduction to Deep Learning
- Review of Probability Theory

What is Deep Learning?

A. If sender contains "xxx.com" -> move to Junk
B. Detect spam emails based on email content
C. Use Recurrent Neural Network to detect spam emails

ARTIFICIAL INTELLIGENCE

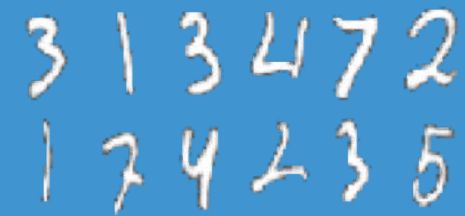Any technique that enables computers to mimic human behavior
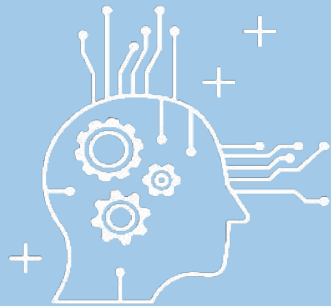
MACHINE LEARNING

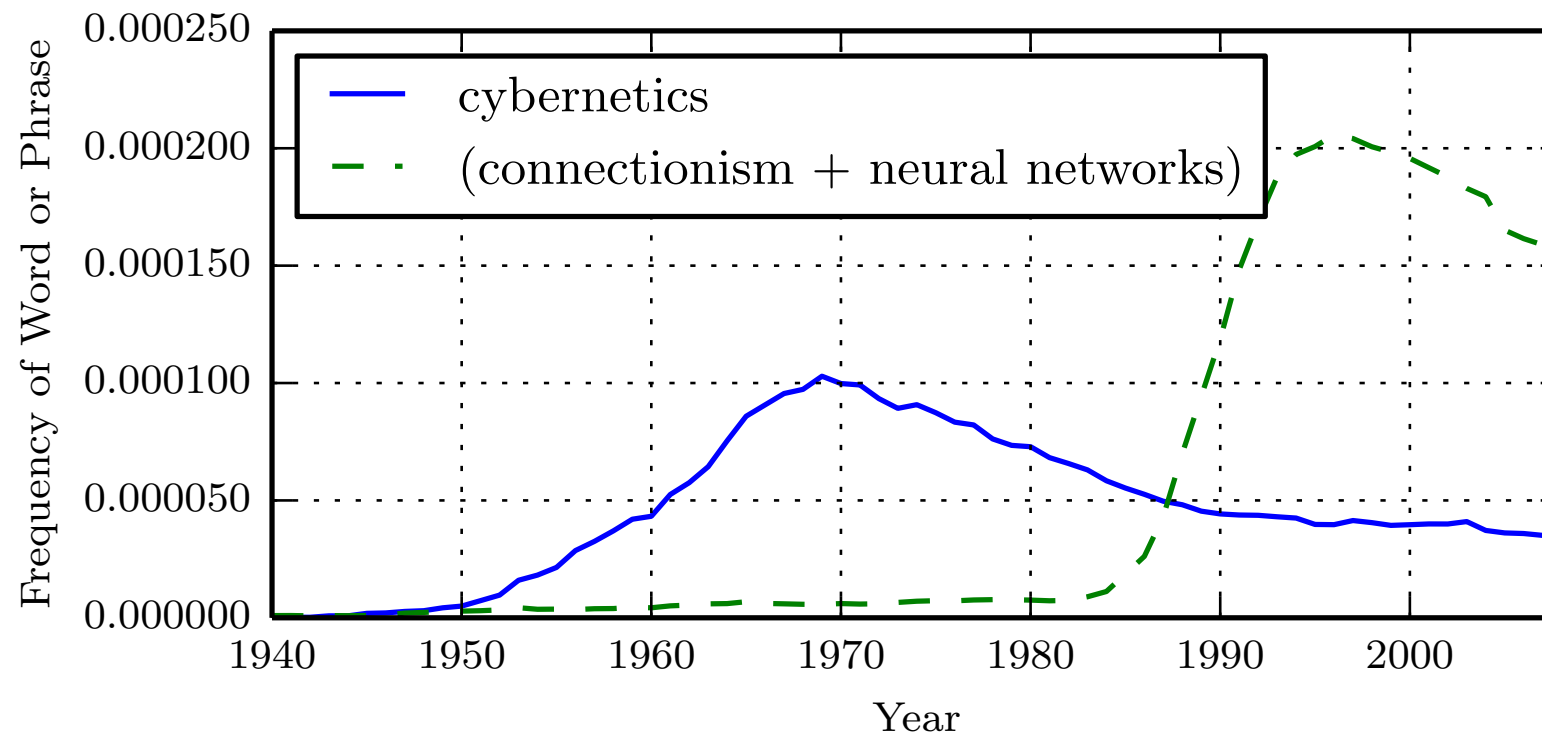Ability to learn without explicitly being programmed

DEEP LEARNING

Extract patterns from data using neural networks

Massachusetts Institute of Technology

Historical Waves

Three historical waves:
- Cybernetics (1940s–1960s): development of theories of biological learning and implementation of the first models such as the perceptron
                            allowing the training of a single neuron
- Connectionism or Neural Networks (1980–1995): back-propagation to train a neural network with one or two hidden layers
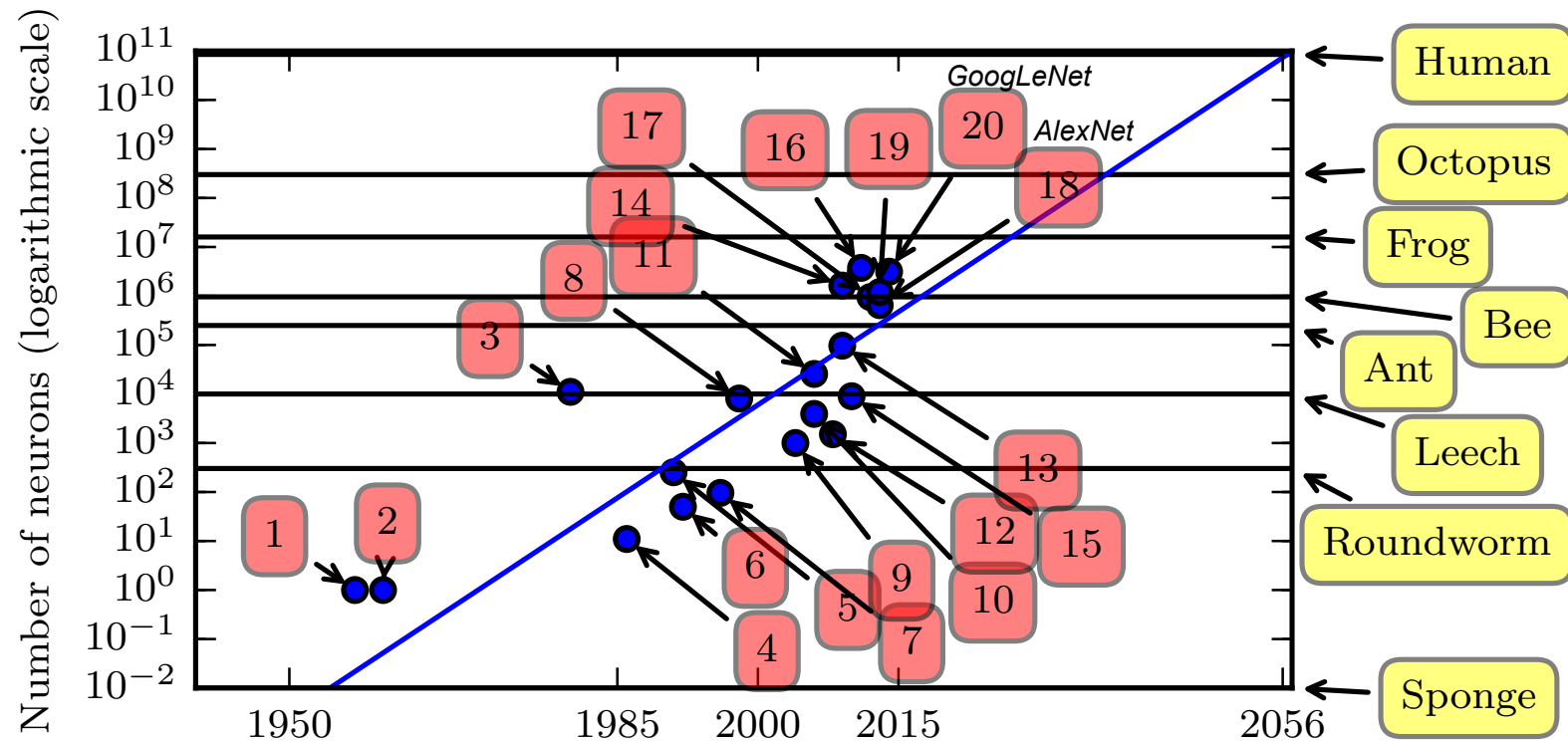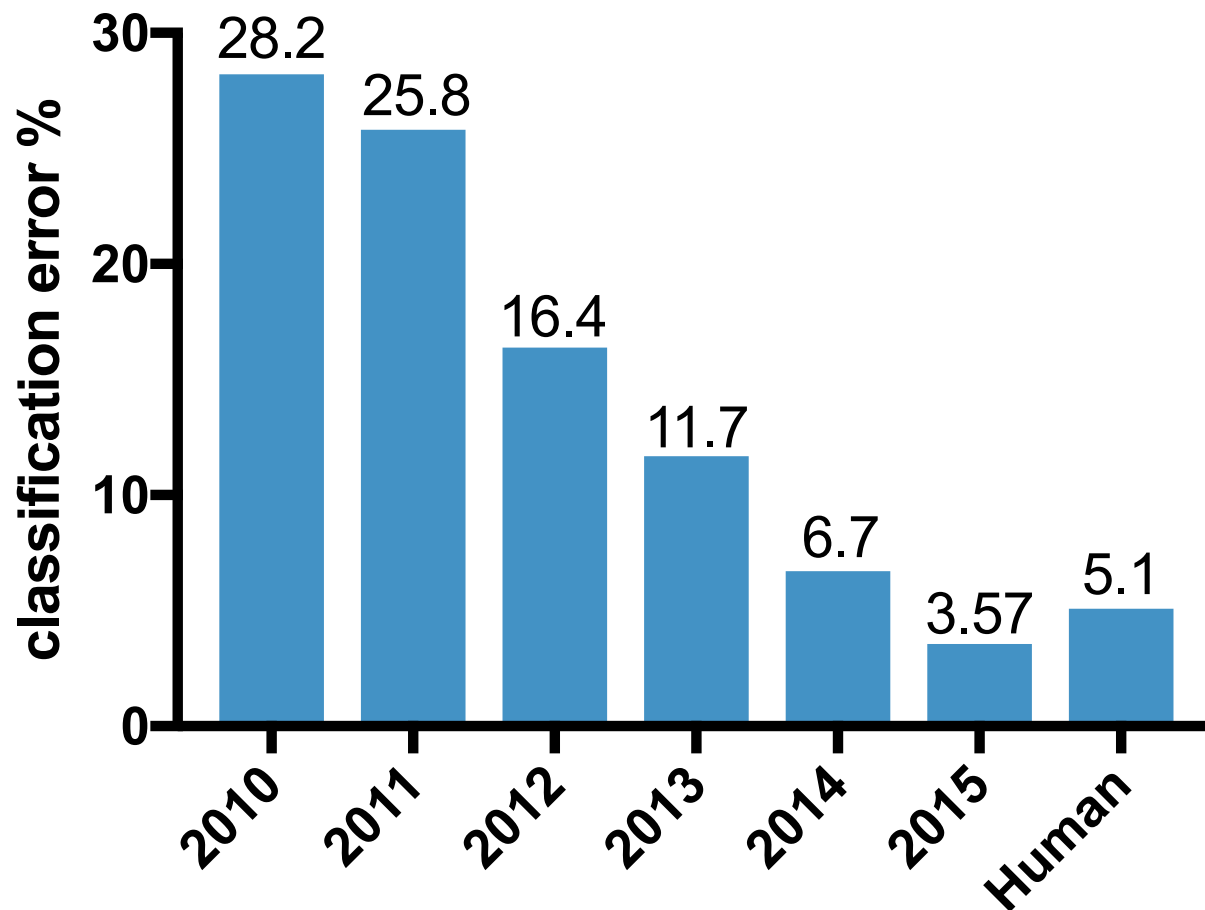- Deep learning (2006 -)

Number of Neurons

Figure 1.11

ImageNet Challenge: Classification Task

2012: AlexNet. First CNN to win.
- 8 layers, 61 million parameters

2013: ZFNet
- 8 layers, more filters

2014: VGG
- 19 layers

2014: GoogLeNet
- "Inception" modules
- 22 layers, 5million parameters
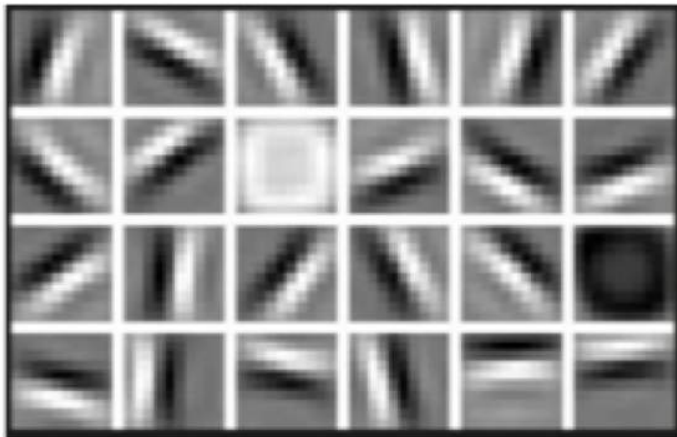
2015: ResNet
- 152 layers

Why Deep Learning and Why Now?

Why Deep Learning?

Hand engineered features are time consuming, brittle and not scalable in practice

Can we learn the **underlying features** directly from data?
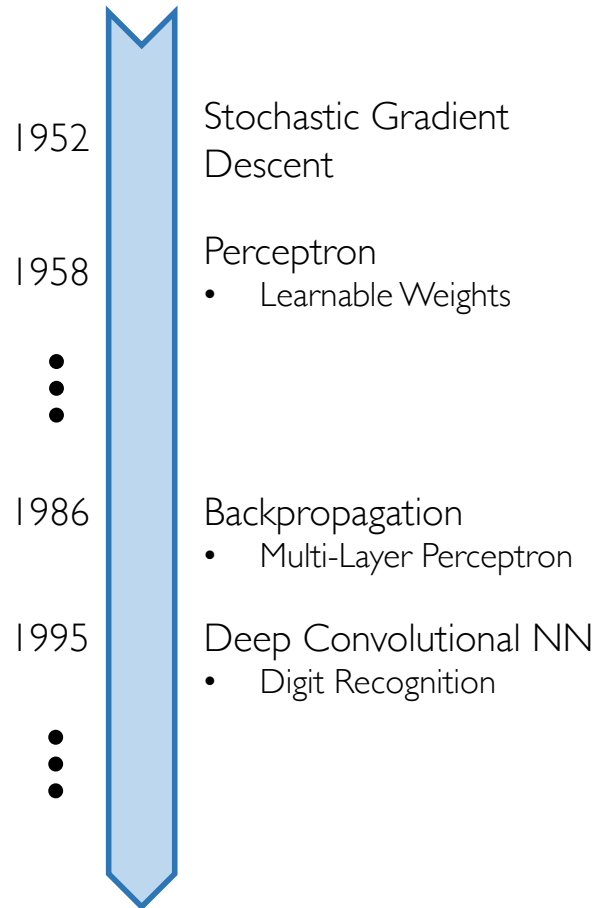
Low Level Features	**Mid Level Features**	**High Level Features**
Lines & Edges	Eyes & Nose & Ears	Facial Structure

Why Now?

Neural Networks date back decades, so why the resurgence?

1952 Stochastic Gradient
      Descent

1958 Perceptron
      • Learnable Weights

⋮

1986 Backpropagation
      • Multi-Layer Perceptron

1995 Deep Convolutional NN
      • Digit Recognition

⋮

1. Big Data

- Larger Datasets
- Easier Collection & Storage

IM**A**GENET

WIKIPEDIA
The Free Encyclopedia

2. Hardware

- Graphics Processing Units (GPUs)
- Massively Parallelizable

3. Software

- Improved Techniques
- New Models
- Toolboxes

TensorFlow

Massachusetts Institute of Technology

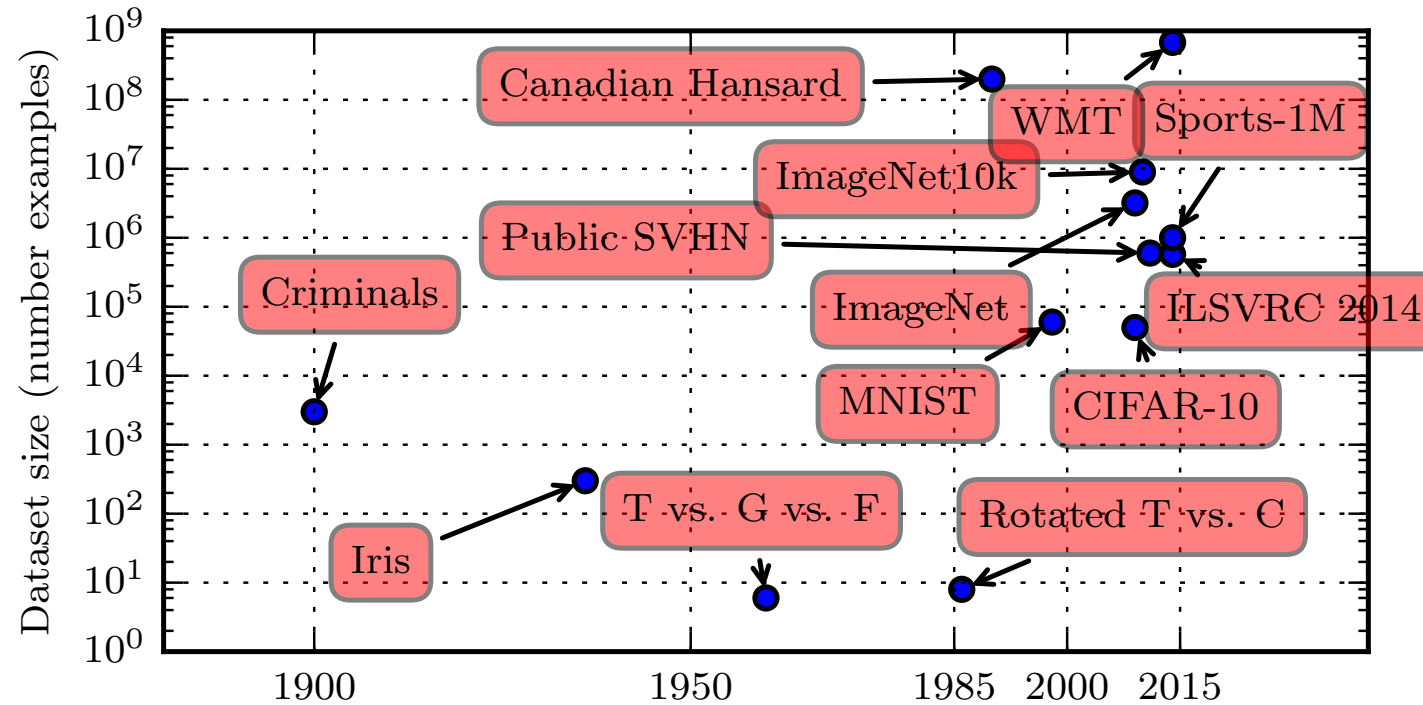# Historical Trends: Growing Datasets

Figure 1.8

Kinds of Deep Learning Problems
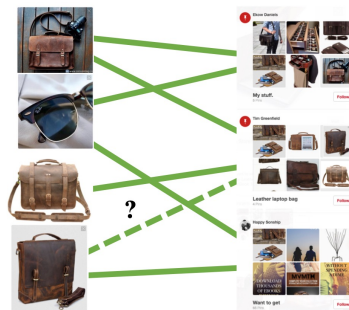

Regression: How much a house costs?

Fig. 1.3.2 Death cap—do not eat!
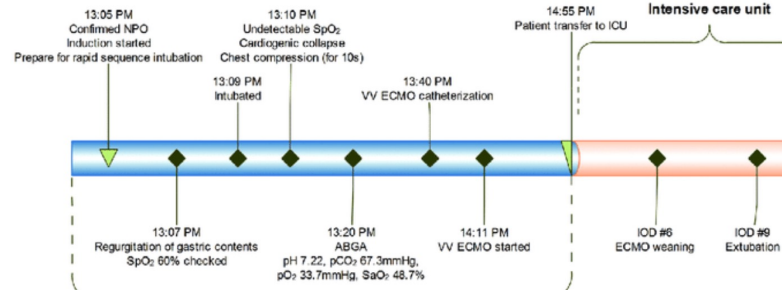
Classification: Is mushroom poisonous?

Sequence learning: Patient clinical event prediction

Sample Generation: Generative Adversarial Networks

. 1.3.3 A donkey, a dog, a cat, and a rooster.

Multi-label classification: Tagging

Recommendation

Automatically captioned

A dog is sitting on the beach next to a dog.
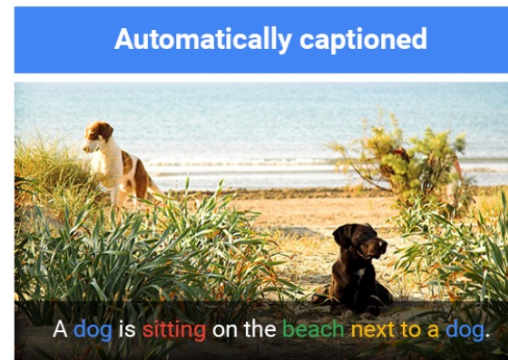
Sequence learning: Image Captioning

Sequence to sequence learning: Translation
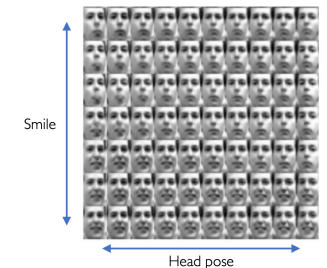
Unsupervised learning: self supervised learning