ECBM E4040 Neural Networks and Deep Learning

Lecture #1: Introduction and Overview

Course Instructor: Zoran Kostic
slides by Yiyin Zhou

Columbia University
Department of Electrical Engineering

Sep. 4th, 2016

Outline

Course Location and Time

Classroom: 207 Mathematics

Fridays, 10:10 am - 12:40 PM

Course Instructor

Prof. Zoran Kostic
website

Office Hours:
Thursdays, 11:30 AM - 1:00 PM, EST
Room 813 CEPSR

TA

TAs
Hanumesh Sirigeri, Yuxiang Chen, Mehmet Kerem Turkcan, Yi Luo, Yang Yu

TA Office Hours:
TBD
EE Lounge, 13th Floor, Mudd

Grading

TENTATIVE Grading

- 5 Homeworks: 5x8% =40%.
- Midterm: 25% Tuesday, March 22, 2016, in class.
- Final Exam: 30%
- Data Collection Project: 5

Homework submission will be through a repository hosted on BitBucket, using private setting.

About Homework #0

Homework #0 will help you ascertain the programming know how needed to solve the following 4 homeworks. If you have difficulties with this homework, please carefully evaluate whether you will have enough time to brush up your programming knowledge by the end of the change of program period, the last date to add/drop a course. You are strongly advised to drop the class if you cannot satisfactorily answer all questions of homework #0 by this date.

CourseWorks

- Course notes, book chapters and homeworks will be made available on CourseWorks.
- We strongly encourage everybody to enroll in Piazza for discussions.

Textbook

Deep Learning

by

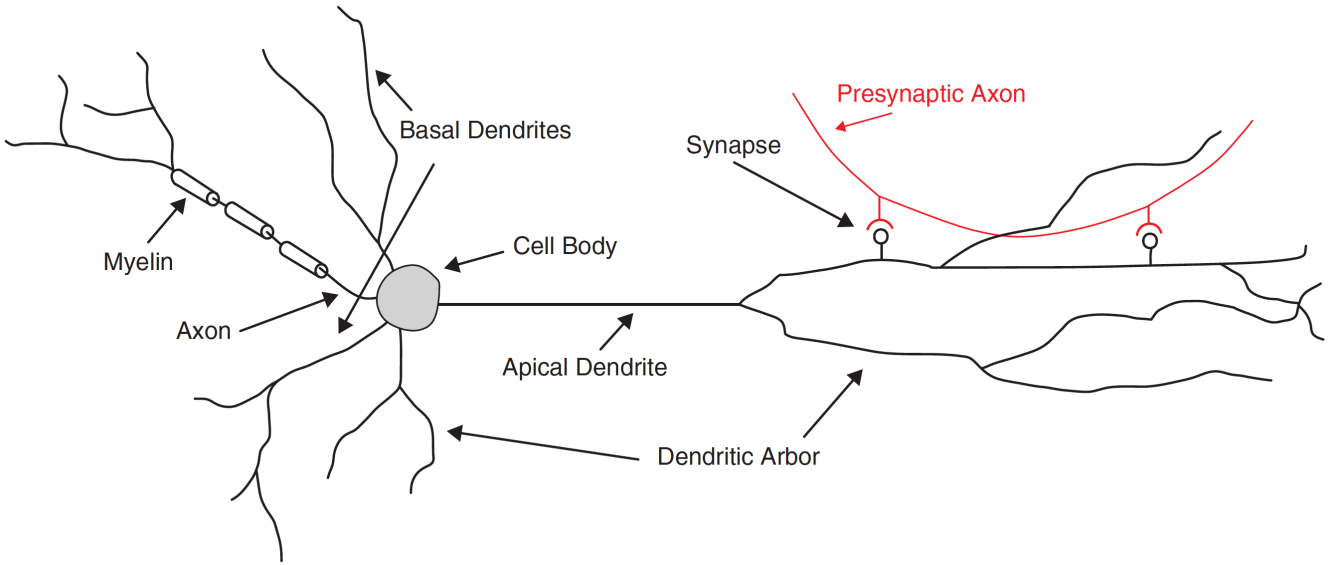
Ian Goodfellow, Yoshua Bengio and Aaron Courville

http://www.deeplearningbook.org/

Biological Neural Networks

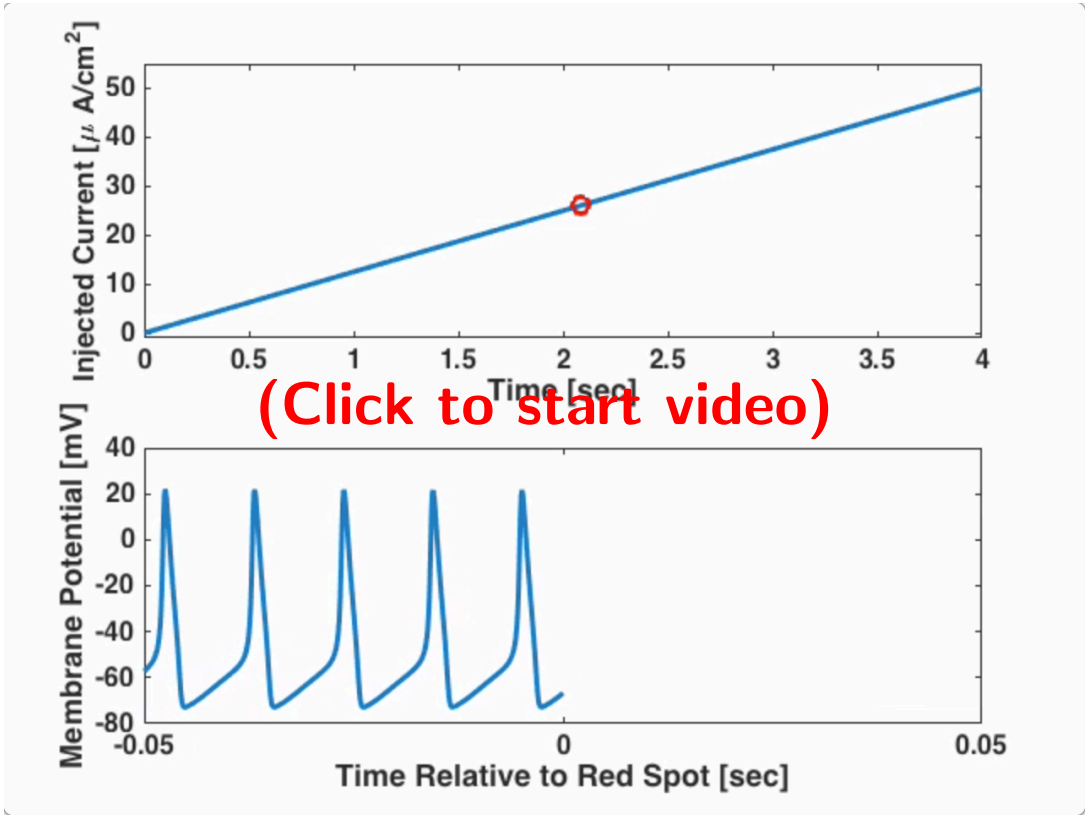
Neurons

A neuron is an electrically excitable cell that processes and transmits information through electrical and chemical signals.



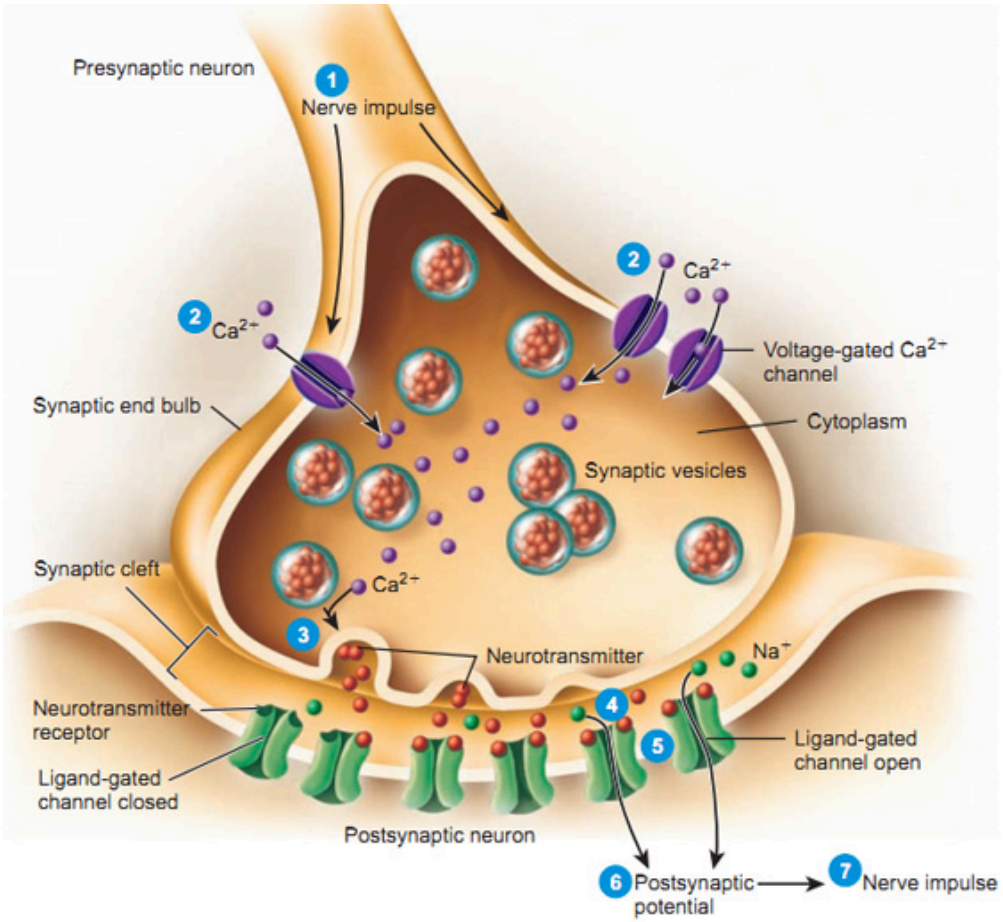
Biological Neural Networks

Neurons Generate Action Potentials when Excited



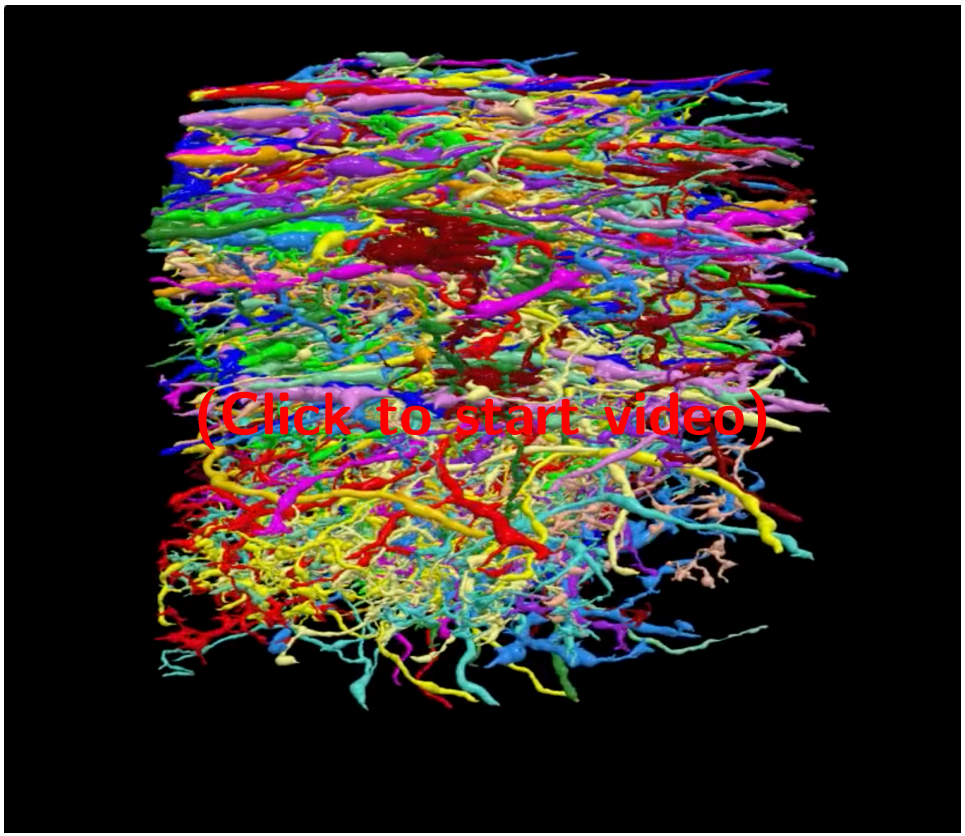
Biological Neural Networks

Signal Transmission between Neurons are through Synapses



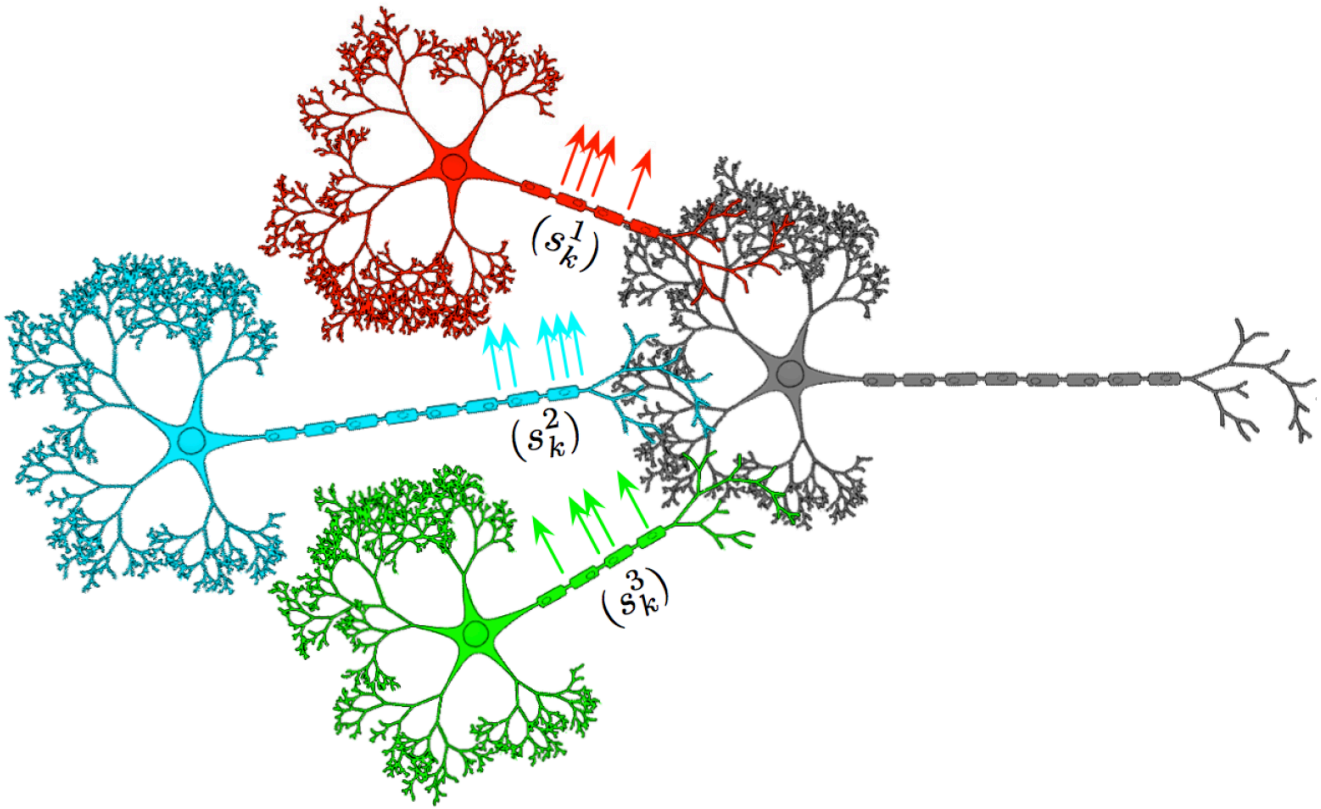
Biological Neural Networks

An Example of a Reconstructed Neural Network



Biological Neural Networks

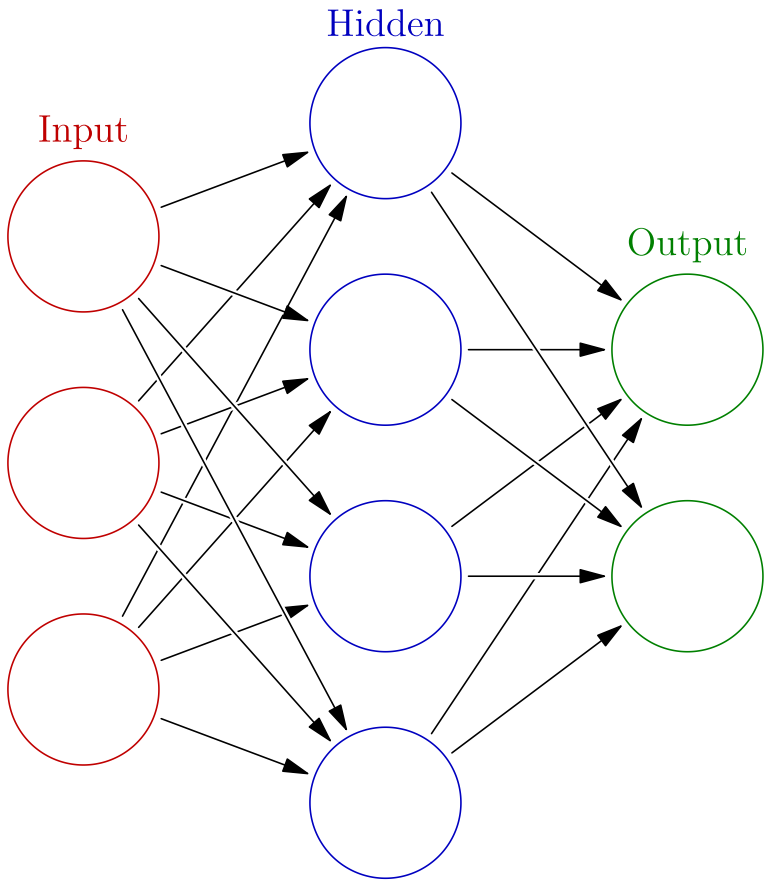
Inputs via Dendrites



Standard Artificial Neural Networks

A standard artificial neural network consists of

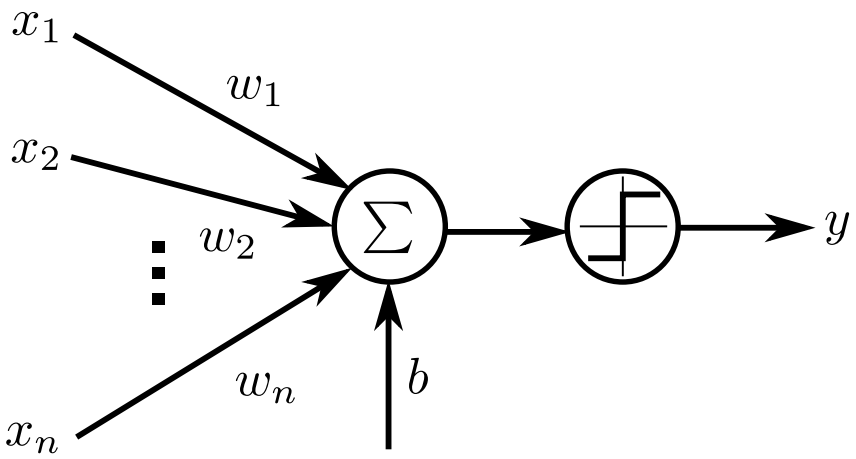
- many simple, connected "neurons",
- each neuron has an activation function
- input neurons get activated through sensors perceiving the environment
- other neurons are activated through weighted connections from previously activated neurons



Glosser.ca Wikipedia

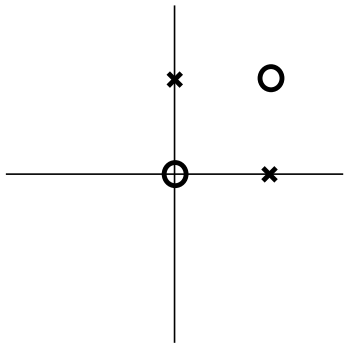
Standard Artificial Neural Networks

Perceptron



$$y(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{w}^T\mathbf{x} + b > 0 \\ -1, & \text{otherwise} \end{cases}$$

- Linear classifier
- Does not work on the following case (two classes are star and circle):

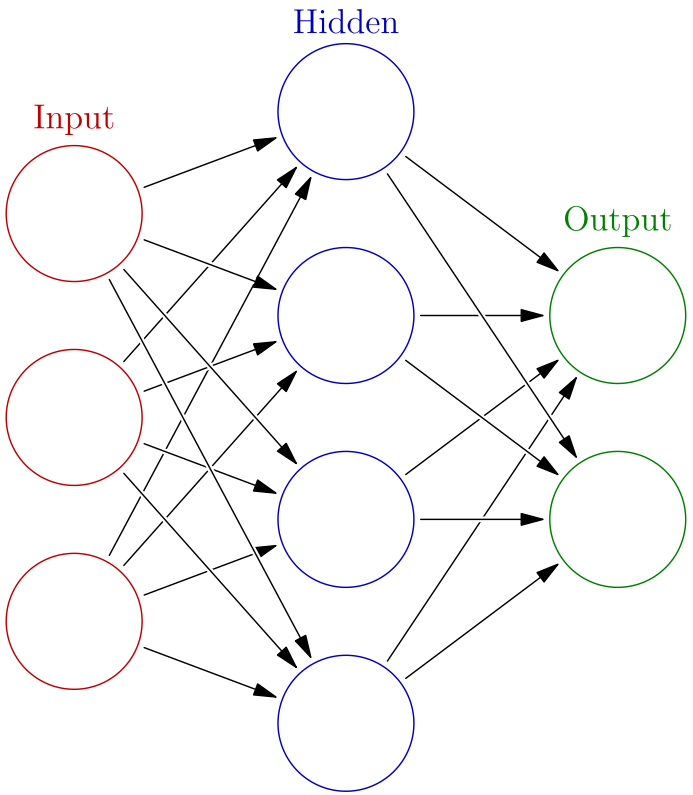


Deep Neural Networks
Multilayer Perceptron (MLP)

Universal approximation theorem:

A feed-forward network with a single hidden layer containing a finite number of neurons can approximate continuous functions on compact subsets of $\mathbb{R}^n$.

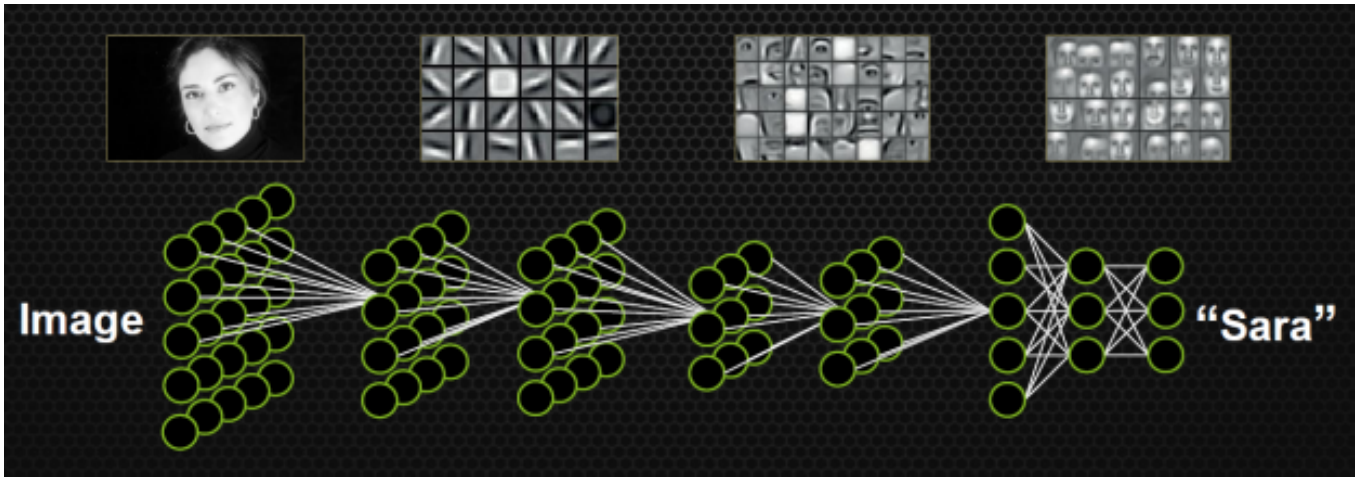
How to train a deep neural network becomes the central question.



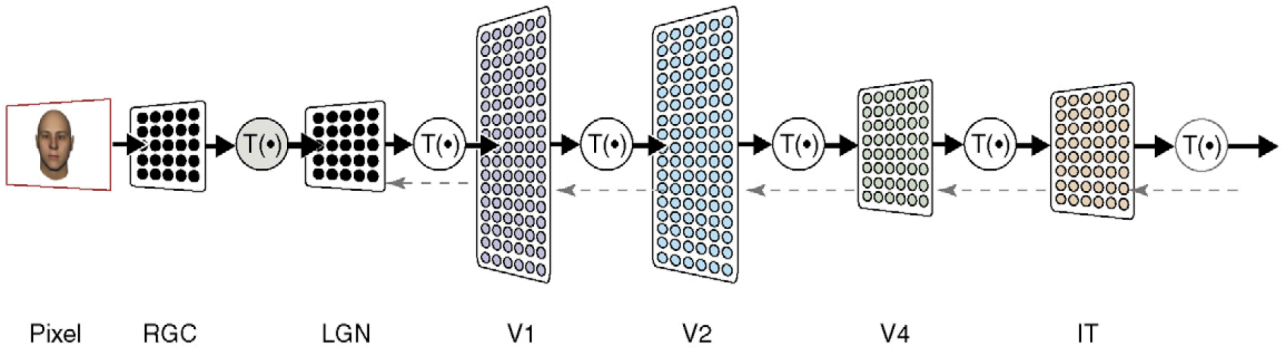
Deep Neural Networks

Neural Networks with Many Stages/Layers

Deep Neural Network

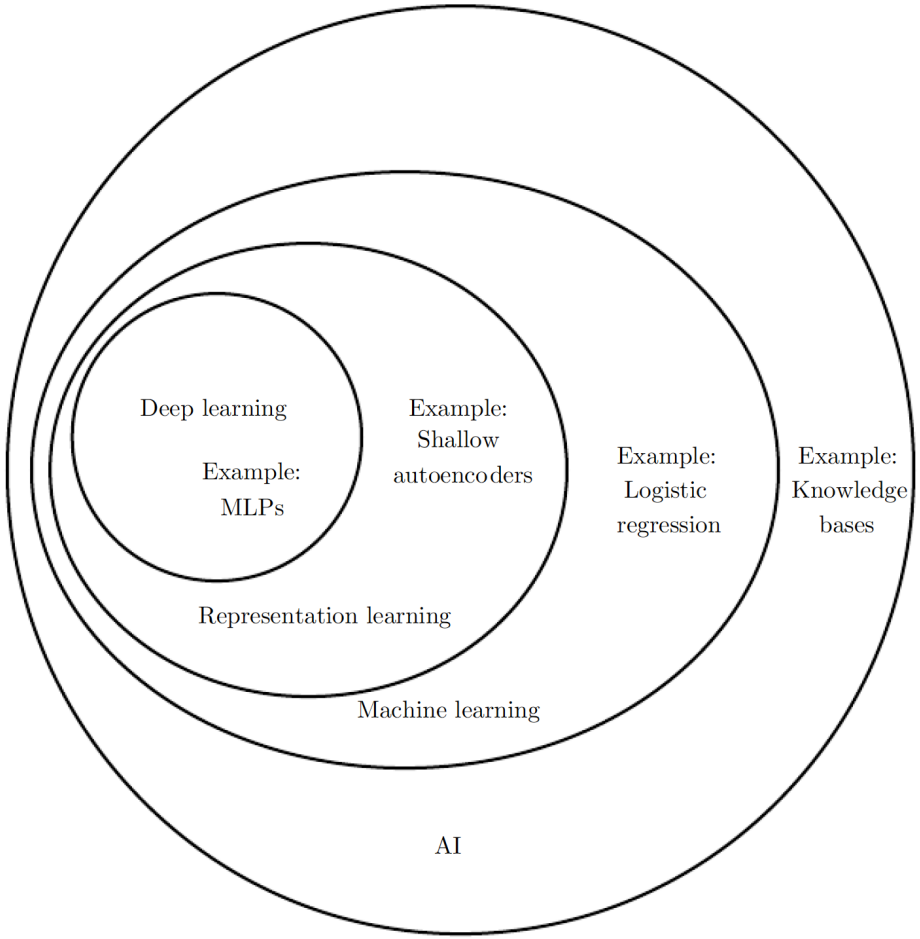


Ventral Visual Processing Stream



Deep Learning

A Venn Diagram of Relationship between AI disciplines



The Importance of Features

Hard-coded or Learned?

- Conventional machine-learning techniques were limited in their ability to process natural data in their raw form.
- Requires careful engineering and considerable domain expertise to design a feature extractor that transform the raw data into a suitable internal representation or feature vector from which the learning subsystem (e.g. a classifier) can detect or classify patterns in the input.

The Importance of Features

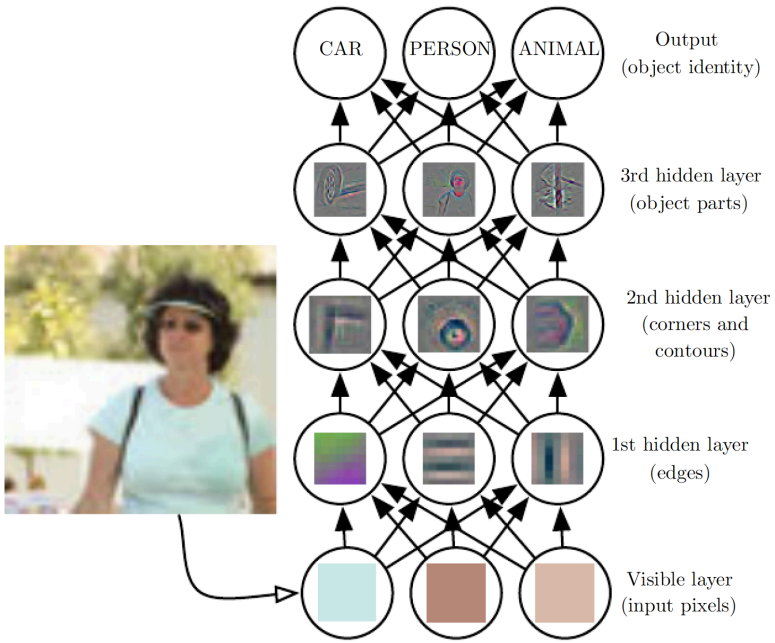
How to Learn What Features to Use?

Obtaining a representation may be as difficult as solving the original problem.

Representation Learning is a set of methods that allows a machine to be fed with raw data and to automatically discover the representation needed for detection or classification.

Deep Learning

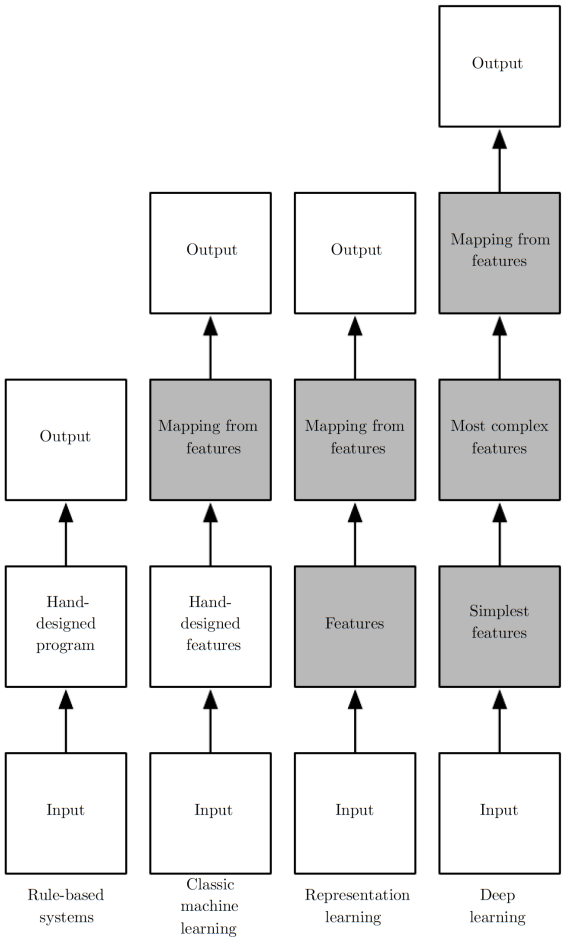
Deep Learning Allows the Computer to Build Complex Concepts out of Simpler Concepts



Deep-learning methods are representation-learning methods with multiple levels of representation, obtained by composing simple, but non-linear modules that each transform the representation at one level into a representation at a higher, slightly more abstract level.

Deep Learning

A Comparison of Different AI Disciplines



Deep Learning Has Had a Long and Rich History

- McCulloch-Pitts Neuron (1943), Perceptron (1960s)
- Back-propagation algorithm for training deep networks (1980s)
- Deeper networks than before can be trained and have record breaking results (since 2006), emphasizing the theoretical importance of depth.

Many Neuron Networks are Influenced by Neuroscience

But Deep Learning $\neq$ Brain

- Influence from the structure of mammalian visual system.
- Rectified linear activation function.
- One deep learning algorithm solves many different tasks.

It is not yet clear

- how memory works in the brain,
- how values/functions are represented,
- what algorithms biological neural systems employ for learning.

Recent Impact of Deep Learning

Adopted by the Technology Industry

- Google
- Microsoft
- Facebook
- IBM
- NVIDIA
- Netflix
- Baidu
- Car manufacturers (Tesla, BMW, Volvo, Ford self-driving cars)
- and many more

Recent Impact of Deep Learning

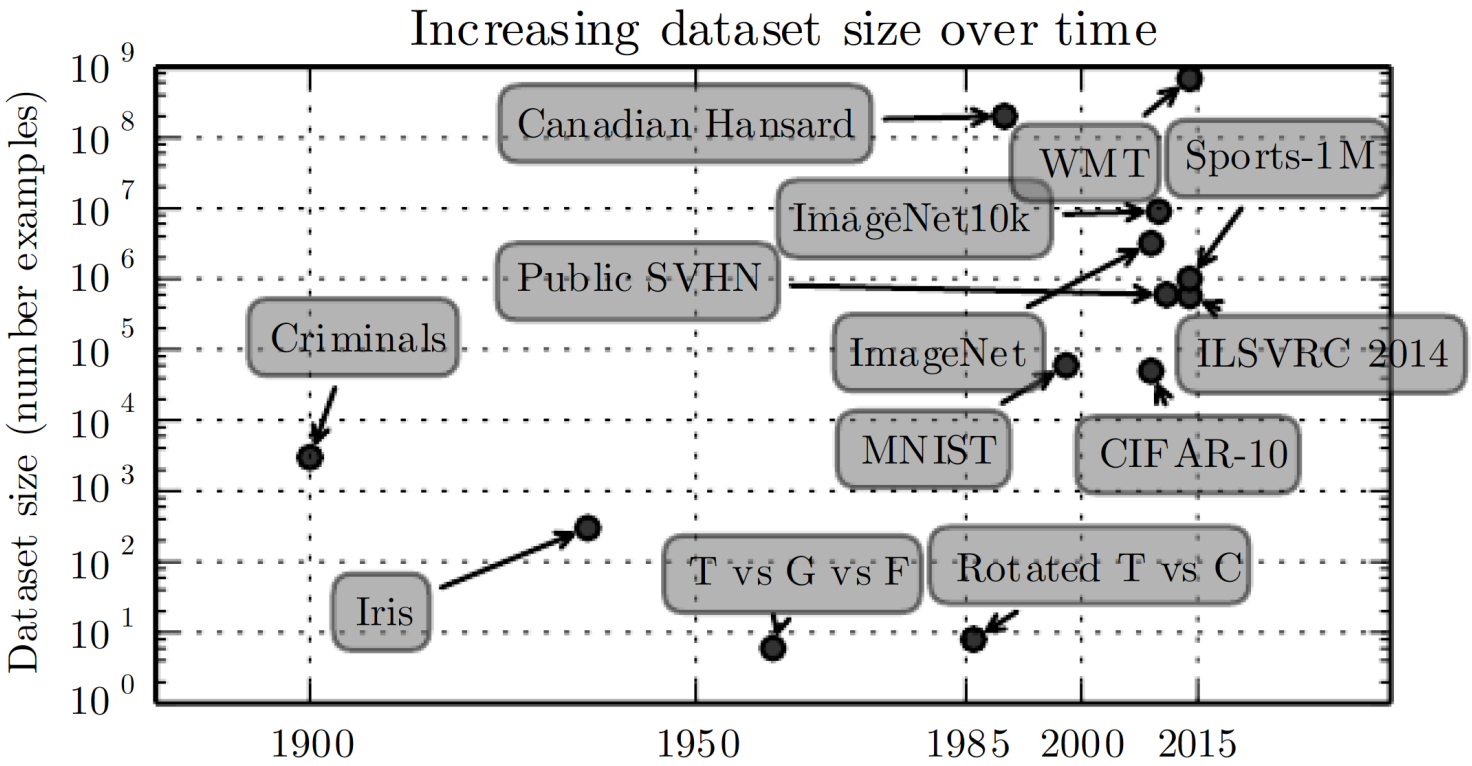
Deep Learning Algorithms Have Broken Many Records

- MNIST (Hand Written Digits)
- Traffic Sign Contest
- ImageNet (Visual Object Detection and Image Classification)
- etc.

Recent Popularity

Why did deep learning become popular again in the recent years?

Deep Learning Benefits from Increasing Dataset Sizes



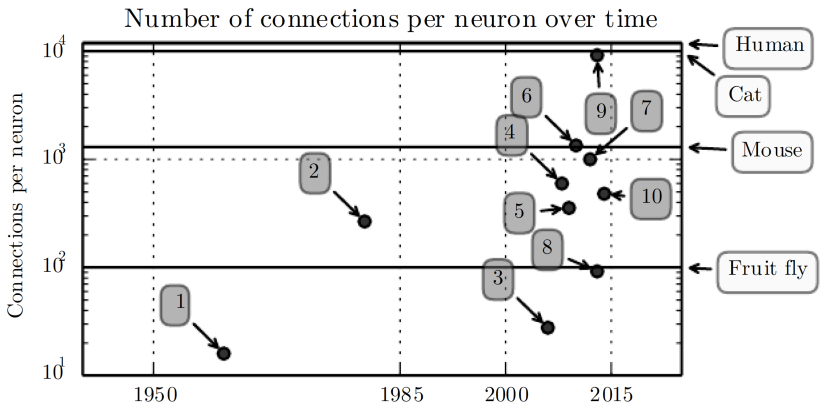
Deep Learning Benefits from Better Hardware

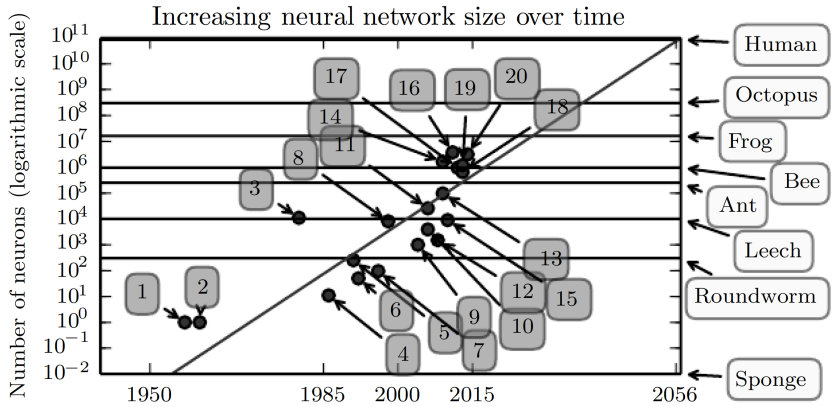
The Role of GPUs on Popularizing Deep Learning Algorithm

- 2010: a new MNIST record of 0.35% error rate. Made possible mainly through a GPU implementation of the Back Propagation algorithm that was up to 50 times faster than standard CPU versions.
- GPU based deep learning algorithms have won several pattern recognition competitions in the following years.
- Recent ImageNet records from Microsoft and Google are all GPU based.

Deep Learning Benefit from Better Hardware

Faster, Larger Scale Hardware Supports Increasing Model Size





Deep Learning in Science

A Useful Tool for Processing Massive Amount of Data

- Reconstructing neuron morphology from electro-microscopy raw data
- Analyzing particle accelerator data
- Predicting the activity of potential drug molecules
- Genetic research
- More to come?

Popular Software Packages/Libraries

- Theano (http://deeplearning.net/software/theano/)
- Caffe (http://caffe.berkeleyvision.org/)
- Torch (http://torch.ch/)
- Google TensorFlow (https://www.tensorflow.org/)
- cuDNN (https://developer.nvidia.com/cudnn)

Using Amazon EC2 GPU Instances

Setup Guide

Live setup guide covered:

- Jupyter notebook "Using Amazon Machine Image with Preloaded Theano Link"
- Sign up for EC2 and AWS Educate.
- Create a Group and User
- Create a Security Group
- Create a key pair

Using Amazon EC2 GPU Instances

Create an Instance

Live setup guide covered:

- Create an EBS Volume
- Launch an instance
- Using EBS volume in an instance

Connecting to Amazon EC2 Instance

Instruction for OS X and Linux

See also this link

Logging in:

1. Open a terminal and type

```
ssh -i path-to-your-key.pem ubuntu@ec2-DNS
```

Transfer files:

1. Open a terminal and type

```
sftp -i path-to-your-key.pem ubuntu@ec2-DNS
```

Connecting to Amazon EC2 Instance

Instruction for Windows

See also this link

Create Private Key:

1. In PuTTYgen, load the key you downloaded when you generate the key pair in EC2 console.
2. Save private key

Logging in:

1. Add your private key to Pageant.
2. Open PuTTY, enter DNS to host name, connect.
3. Type ubuntu for login name

Transfer files

1. Add your private key to Pageant.
2. Launch WinSCP, enter DNS in host name, enter ubuntu for user name
3. Click Login

Useful Tools

- tmux: terminal multiplexer.
  Allows you to create virtual terminal inside.
- text editor: vim/emacs or pick your favourite
- Jupyter notebook

Using Jupyter Notebooks

See also "Using Amazon Machine Image with Preloaded Theano Link"

Markdown Basics:

- https://help.github.com/articles/markdown-basics/
- Basic formatting
- Latex math

Using Git Repositories

Interactive tutorial: https://try.github.io

GUI tools for Mac and Windows:

- git https://git-scm.com/
- SourceTree https://www.sourcetreeapp.com/