Machine Learning applied to Planetary Sciences

PTYS 595B/495B

Leon Palafox

https://leonpalafox.github.io/MLClass/

Our path so far

- We started learning Linear Regression
  - Here we are trying to 'regress' against a single real number.
  - We can quantify the effect of different variables on the outcome.

- Logistic Regression
  - Our first classifier
  - Fairly good, and is somewhat fast

- SVMs
  - Probably the most popular ML technique.

For years, Neural Networks were the most popular ML technique.
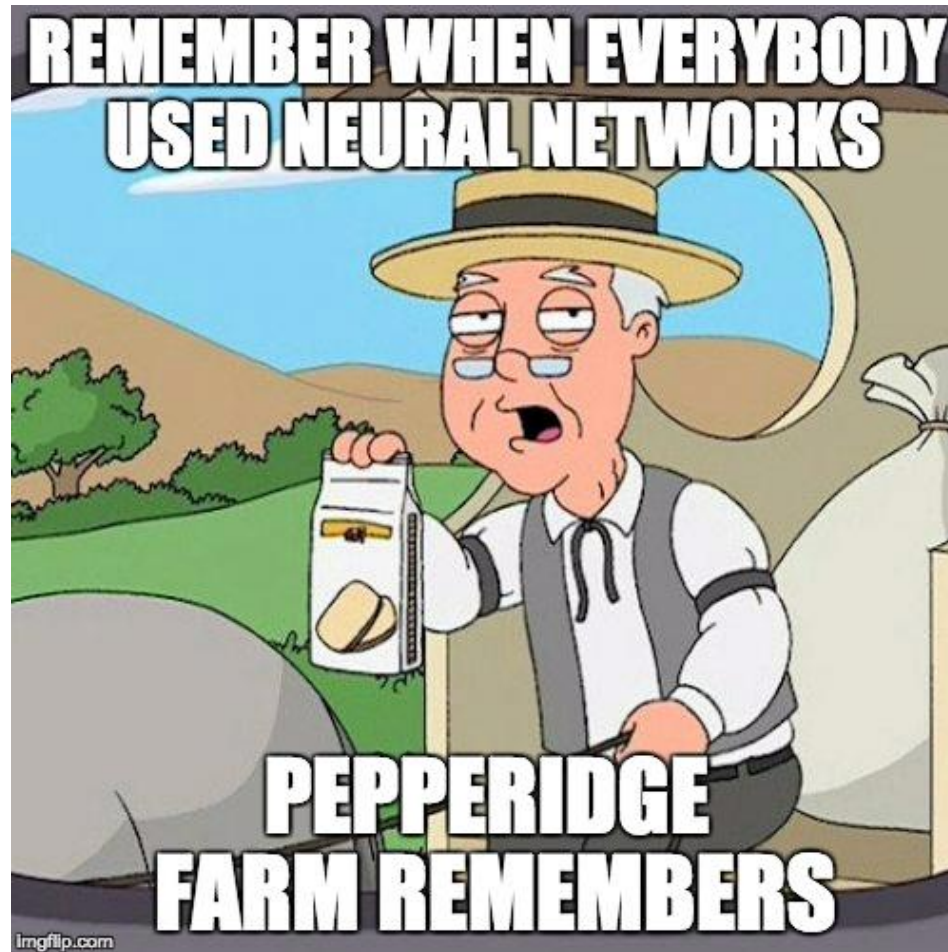
Background

- NNs have gone through a heavy rebranding thorough the years.

- In 1943, McCulloch and Pitts created the first model of an artificial neuron.

- By 1958, Rosenblatt had come up with the Perceptron, the cornerstone of modern NN.

- In 1986, Rumelhart started the connectionism euphoria.

Background

- Processing power was still an issue and until 2006, common NNs were researched by only small clusters of people.

- Training was expensive, and the results only marginally better (or worse) than SVMs or Logistic Regression.

- In 2006, Hinton and Bengio made huge discoveries on how to train NNs and they rebranded them as Deep Nets.

- During this time, Convolutional Neural Networks (CNN) had been a great tool for image pattern recognition.

Motivation

- Deep Nets and CNNs, are by today standards the best algorithm for Image Pattern Recognition.

- The three Big Kahunas of NNs and Deep Nets, Geoffrey Hinton, Yann LeCun and Yoshua Bengio are working actively with Google, Facebook and University of Toronto, respectively.

  - Bengio just got 90 million dollars to do Deep Learning research.

Motivation
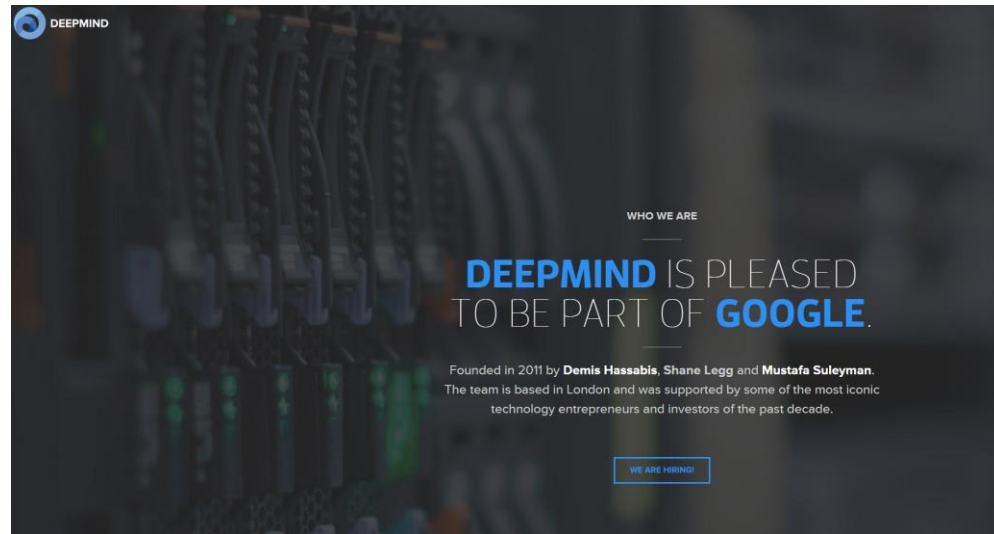
- In January Google bought DeepMind, a startup with no WebPage, no Product, a single NIPS (AI conference) Demo.

- They bought it for $500 million.

- Facebook was deeply interested as well.

READ SOMETHING THAT MEANS SOMETHING.
$1 A WEEK »

SUBSCRIBE
RENEW
GIVE A GIFT
NON-U.S. ORDERS

Sign in | Link your subscription | TNY Store

THE NEW YORKER

NOVEMBER 25, 2012

IS "DEEP LEARNING" A REVOLUTION IN ARTIFICIAL INTELLIGENCE?

BY GARY MARCUS

Tweet · +1 ·

NEWS DESK

Can a new technique known as deep learning revolutionize artificial intelligence, as yesterday's front-page article at the New York *Times* suggests? There is good reason to be excited about deep learning, a sophisticated "machine learning" algorithm that far exceeds many of its predecessors in its abilities to recognize syllables and images. But there's also good reason to be skeptical. While the *Times* reports that "advances in an artificial intelligence technology that can recognize patterns offer the possibility of machines that perform human activities like seeing, listening and thinking," deep learning takes us, at best, only a small step toward the creation of truly intelligent machines. Deep learning is important work, with immediate practical applications. But it's not as breathtaking as the front-page story in the New York *Times* seems to suggest.

THE WALL STREET JOURNAL. | JAPAN

日本リアルタイム
JAPAN REAL TIME

8:47 pm JST
Oct 1, 2014   JAPAN

NTT, Toyota Seek 'Deep Learning' Expertise

ARTICLE | COMMENTS

DEEP LEARNING · NTT · PREFERRED NETWORKS · TOYOTA MOTOR CORP

Email · Print · f 176 · ▼ 140 · g+ · in · ⊙ · B!

By TAKASHI MOCHIZUKI [CONNECT]

The New York Times

The New York Times

Science

Search All NYTimes.com   [Go]

WORLD · U.S. · N.Y. / REGION · BUSINESS · TECHNOLOGY · SCIENCE · HEALTH · SPORTS · OPINION · ARTS · STYLE · TRAVEL · JOBS · REAL ESTATE · AUTOS
ENVIRONMENT · SPACE & COSMOS

Scientists See Promise in Deep-Learning Programs

A voice recognition program translated a speech given by Richard F. Rashid, Microsoft's top scientist, into Mandarin Chinese.

Hao Zhang/The New York Times

By JOHN MARKOFF
Published: November 23, 2012

Using an artificial intelligence technique inspired by theories about how the brain recognizes patterns, technology companies are reporting startling gains in fields as diverse as computer vision, speech recognition and the identification of promising new molecules for designing drugs.

The advances have led to widespread enthusiasm among researchers who design software to perform human activities like seeing, listening and thinking. They offer the promise of machines that converse with humans and perform tasks like driving cars and working in factories, raising the specter of automated robots that could replace human workers.

The technology, called deep learning, has already been put to use in services like Apple's Siri virtual personal assistant, which is based on Nuance Communications' speech recognition service, and in Google's Street View, which uses machine vision to identify specific addresses.

But what is new in recent months is the growing speed and accuracy of deep-learning programs, often called artificial neural networks or just "neural nets" for their resemblance to the neural connections in the brain.

"There has been a number of stunning new results with deep-learning methods," said Yann LeCun, a computer scientist at New York University who did pioneering

A student team led by the computer scientist Geoffrey E. Hinton used
Keith Penner

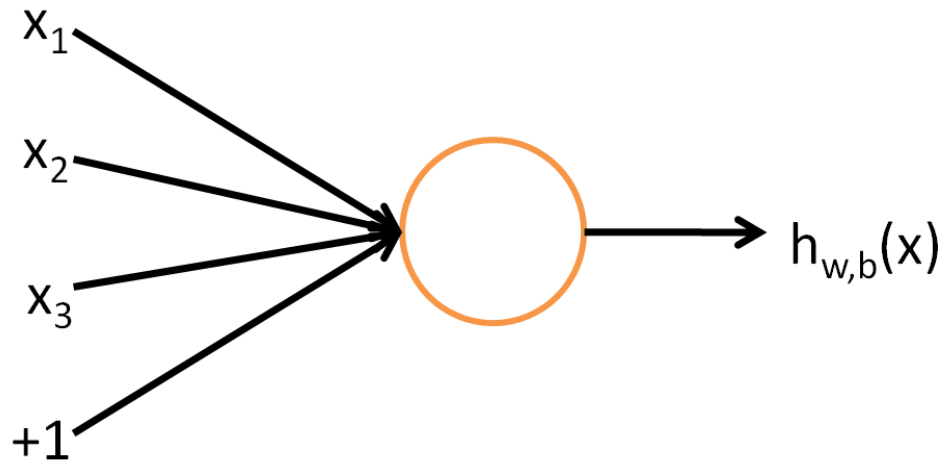# Perceptron

- The main idea is to mimic how a real neuron works. It has a nucleus and pathways that connect it to other equally simple units.

x_1
x_2
x_3
$+1$

$h_{w,b}(x)$

Cell body
Axon
Telodendria
Nucleus
Axon hillock
Synaptic terminals
Golgi apparatus
Endoplasmic reticulum
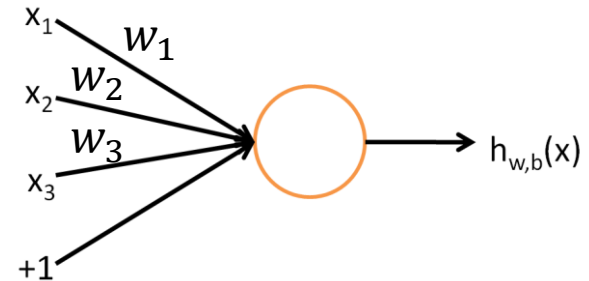Mitochondrion
Dendrite
Dendritic branches
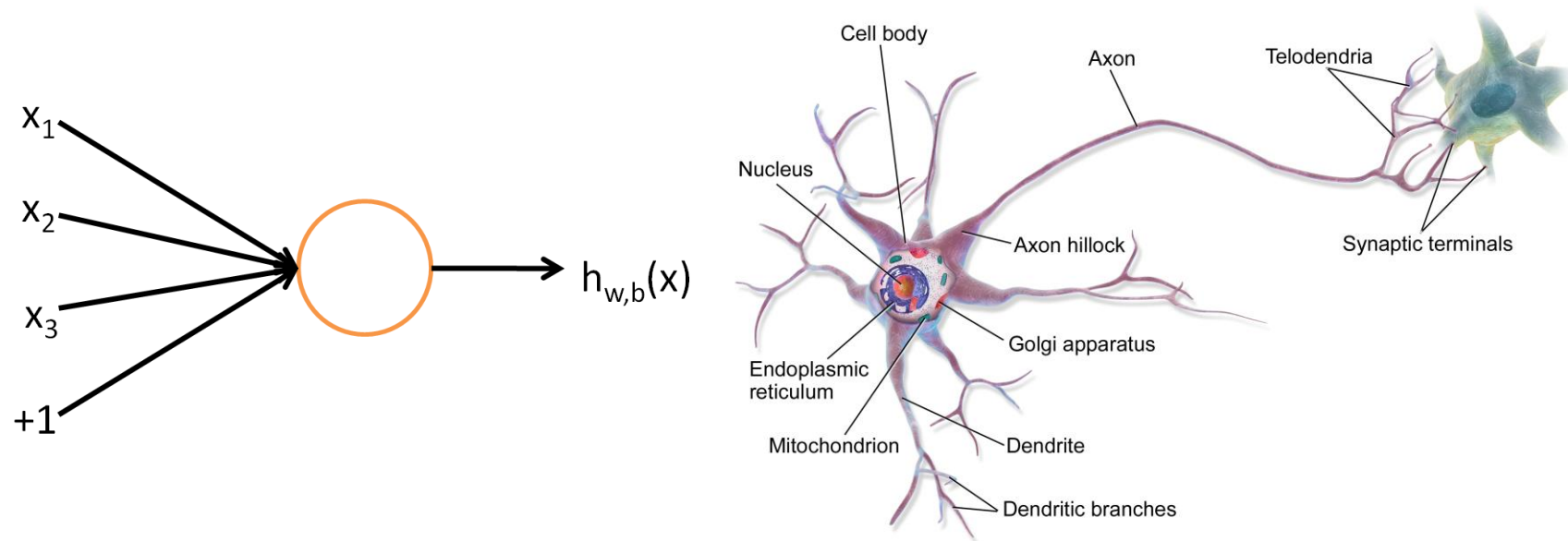
Perceptron

- Tries to mimic a real NN, since it has a nucleus that processes some inputs and give an output.

- $h_{w,b}(x)$ is a function of all the inputs, and is composed of two terms.

Perceptron

$$h_{w,b}(x) = f\left(\sum_{i=1}^{3} W_i x_i + b\right)$$

f is called the activation function, and it works as a way to discretize the outputs of the perceptron.
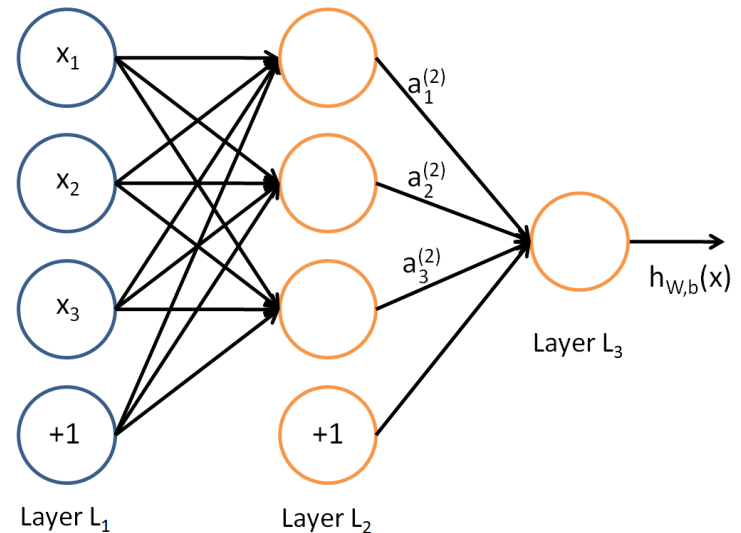
One of the most common activations functions is the sigmoid function:

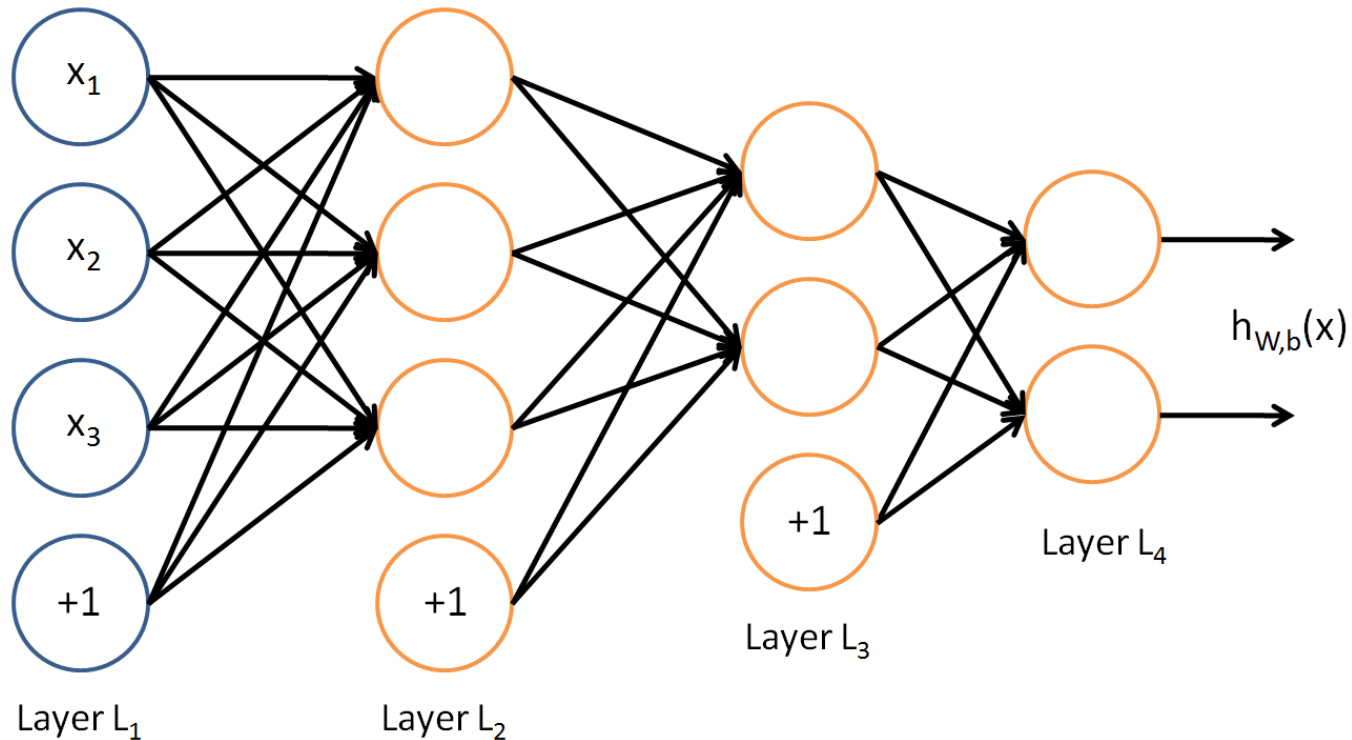$$f(z) = \frac{1}{1 + \exp(z)}$$

This looks very familiar

Neural Network

- Naturally, a NN is going to be a set of perceptrons interconnected within each other.

$$a_1^{(2)} = f(W_{11}^{(1)}x_1 + W_{12}^{(1)}x_2 + W_{13}^{(1)}x_3 + b_1^{(1)})$$
$$a_2^{(2)} = f(W_{21}^{(1)}x_1 + W_{22}^{(1)}x_2 + W_{23}^{(1)}x_3 + b_2^{(1)})$$
$$a_3^{(2)} = f(W_{31}^{(1)}x_1 + W_{32}^{(1)}x_2 + W_{33}^{(1)}x_3 + b_3^{(1)})$$
$$h_{W,b}(x) = a_1^{(3)} = f(W_{11}^{(2)}a_1^{(2)} + W_{12}^{(2)}a_2^{(2)} + W_{13}^{(2)}a_3^{(2)} + b_1^{(2)})$$
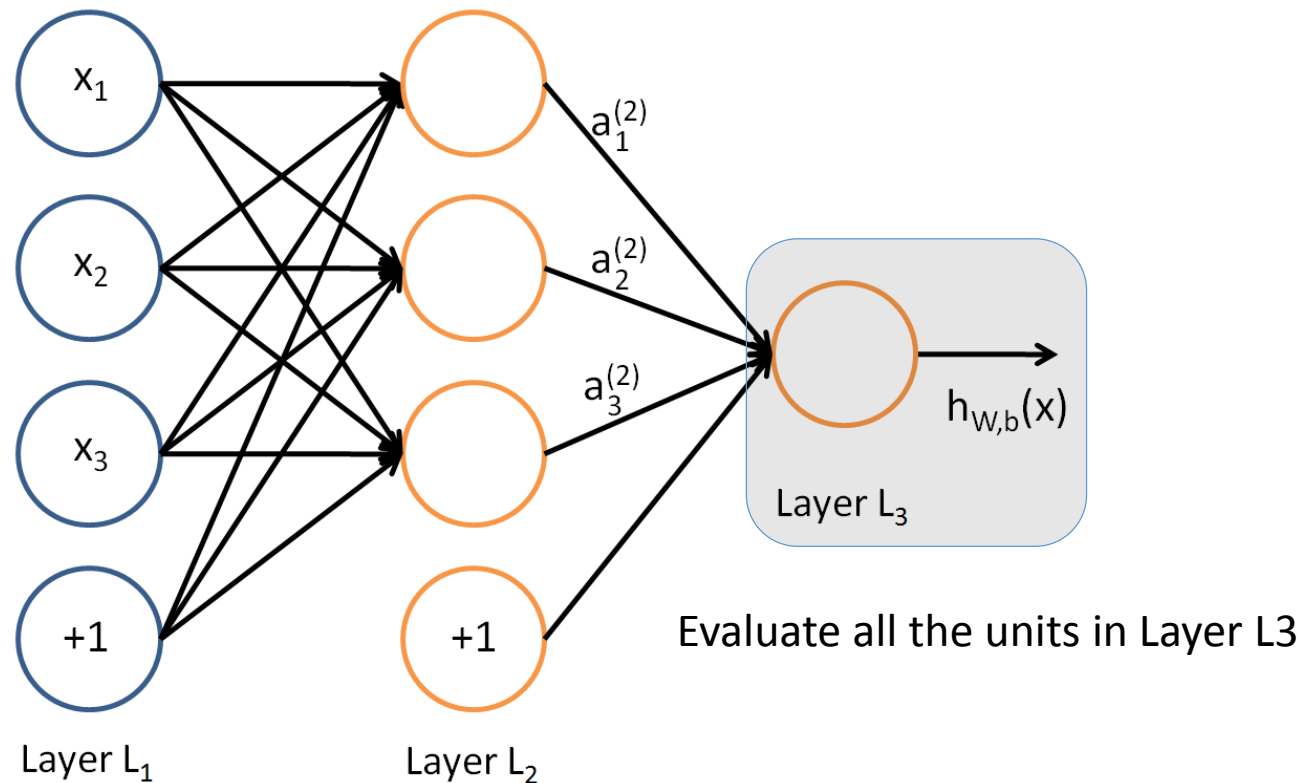
Neural Network

- We can add as many layers and outputs as we want, for example a two binary output allows us to classify in four classes.
- We also regularize NNs, since they can be also prone to overfitting.
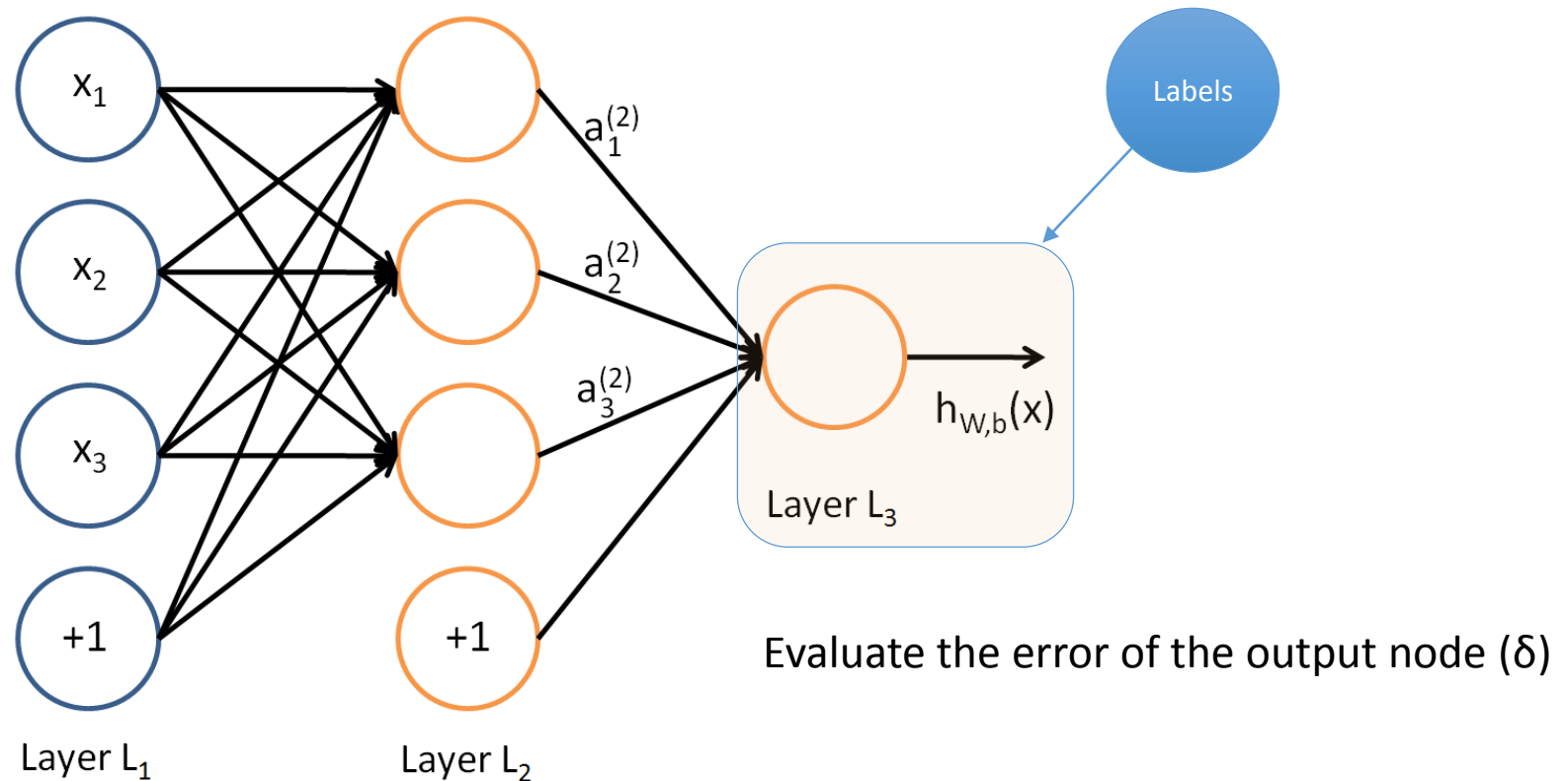
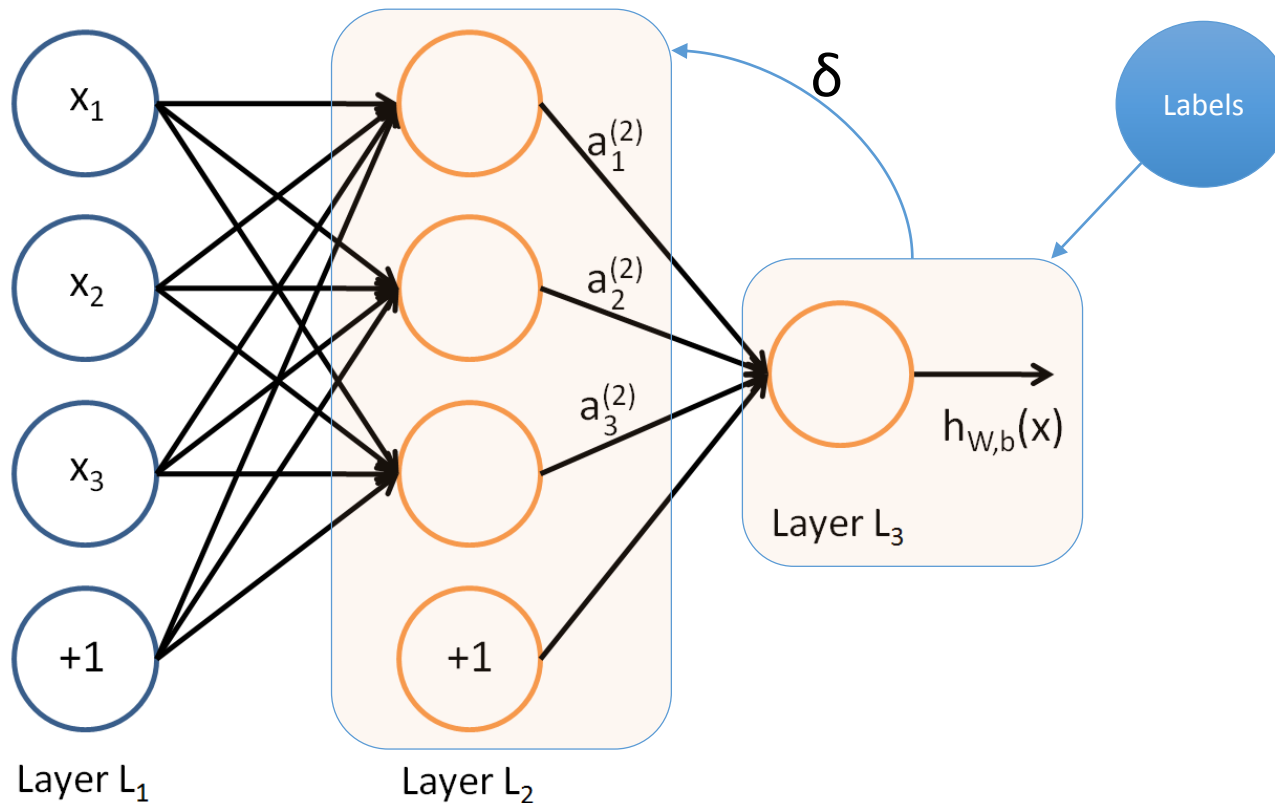# Feedforward-Backpropagation
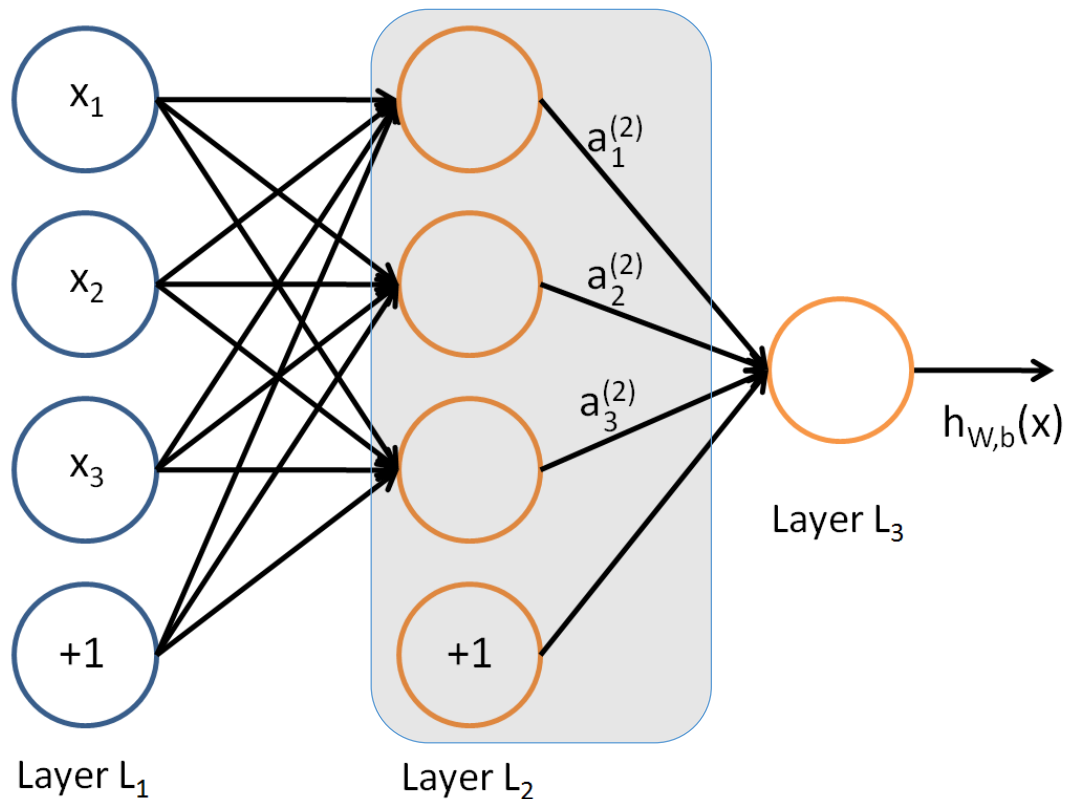
Evaluate all the units in Layer L2

Feedforward-Backpropagation

Layer L_1

Layer L_2

$a_1^{(2)}$

$a_2^{(2)}$

$a_3^{(2)}$

$h_{W,b}(x)$

Layer L_3

Evaluate all the units in Layer L3

Feedforward-Backpropagation

Evaluate the error of the output node (δ)

Feedforward-Backpropagation

Evaluate the errors of the middle layer nodes (δ)

Basic Intuitions behind NNs

- Instead of using a single classifier, we are using a bundle of them.

- Each classifier, via different weights will emphasize different parts of the input signal.

- For example, in an image, that could be shadows, or shapes.

TensorFlow!

- http://playground.tensorflow.org