Introduction to Deep Learning

School of Information Studies
Syracuse University

What is Deep Learning?

Deep learning is a subfield of machine learning

- Builds on Representation Learning to automatically learn good features/representations
- Deep Learning algorithm learns multiple levels of feature representations in increasing levels of complexity or abstraction

Deep learning can

- Not only automatically learn good features
- But do so by using vast amounts of unlabeled data

Overview material adapted from:

RS - Richard Socher, Stanford Course Notes, Deep Learning for NLP, 2016 and

MS - Manning and Socher, tutorial notes on Deep Learning, NAACL 2013.

School of Information Studies
Syracuse University

Current Machine Learning

Most current machine learning works well because **humans design good input features**

- Example: features for finding named entities or organization names (Finkel, 2010)

Machine learning solves an optimization problem over the feature space that learns weights for the features from labeled data, in order to make good predictions on new data

Typical classification tasks require human annotated labeled data

Feature	NER
Current Word	✓
Previous Word	✓
Next Word	✓
Current Word Character n-gram	all
Current POS Tag	✓
Surrounding POS Tag Sequence	✓
Current Word Shape	✓
Surrounding Word Shape Sequence	✓
Presence of Word in Left Window	size 4
Presence of Word in Right Window	size 4

Diagram RS

School of Information Studies
Syracuse University
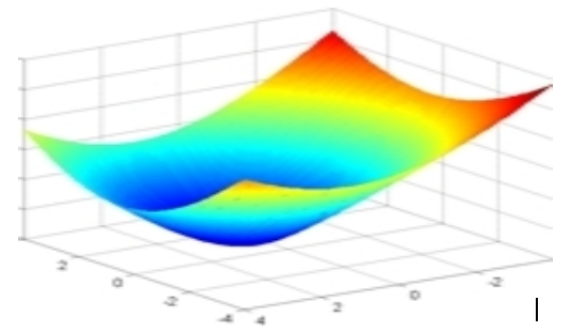
Deep Learning Architecture

Most commonly used architecture uses various types of multi-layer neural networks, such as Belief NN

Output layer

Here predicting a supervised target

Hidden layers

These learn more abstract representations as you head up

Input layer

Raw sensory inputs (roughly)

$\mathbf{h}^4$

$\mathbf{h}^3$

$\mathbf{h}^2$

$\mathbf{h}^1$

$\mathbf{x}$
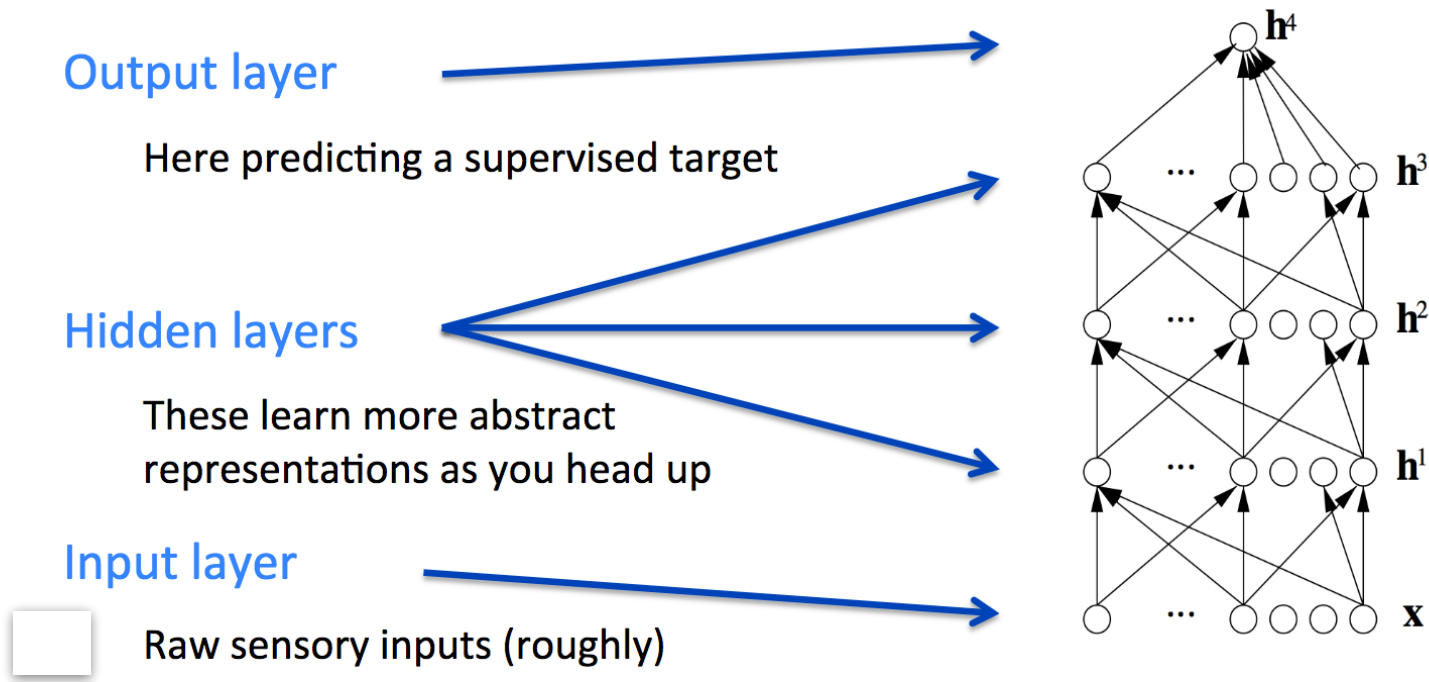
Diagram MS

School of Information Studies
Syracuse University

Single Neuron

A single neuron is a computational unit with an activation function (f). It takes inputs (3) plus a bias term (b) and gives an output

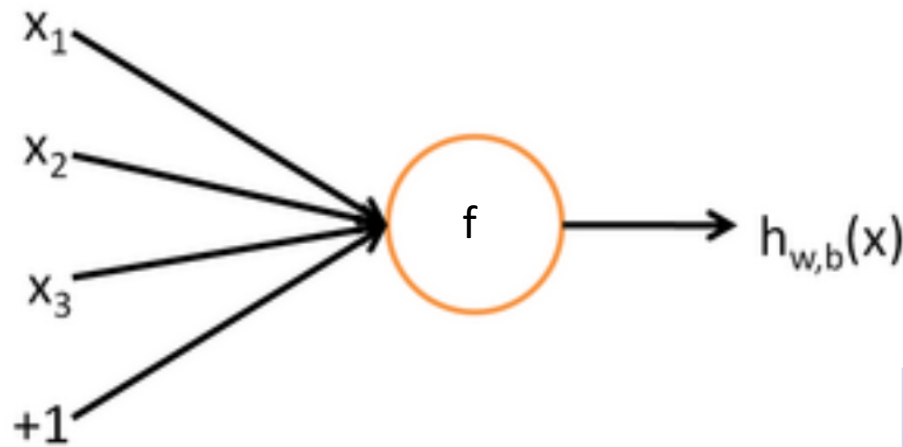- Expressed as a result h depending on weights w and b

Diagram MS

School of Information Studies
Syracuse University

Single Layer Neural Network (NN)

Can also solve current machine learning problems

If we feed a vector of inputs (the features) into a bunch of neurons then we get a vector of outputs, which can be combined according to the task we are trying to solve (the objective function h)
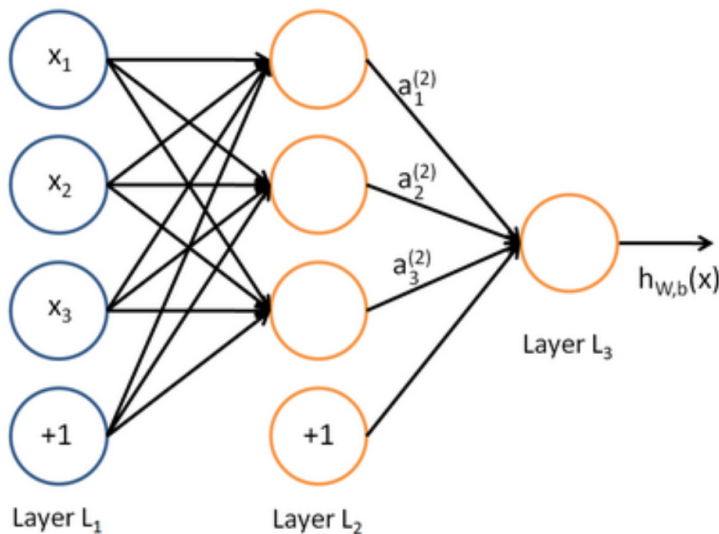

Diagram MS

Training the network learns the weights:
- run the network to predict an output,
- compare the output with the desired (gold) result,
- run back through the network adjusting the weights to reduce the error,
- and iterate.

The weights (and bias) gives the model to compute the predicted output for future data.

School of Information Studies
Syracuse University

Deep Learning NN

We can keep going and add multiple layers

And we can revise our learning algorithm to also learn representations of the input X

- Several algorithms for how to do the "feed forward" and "back propagation" in an efficient way
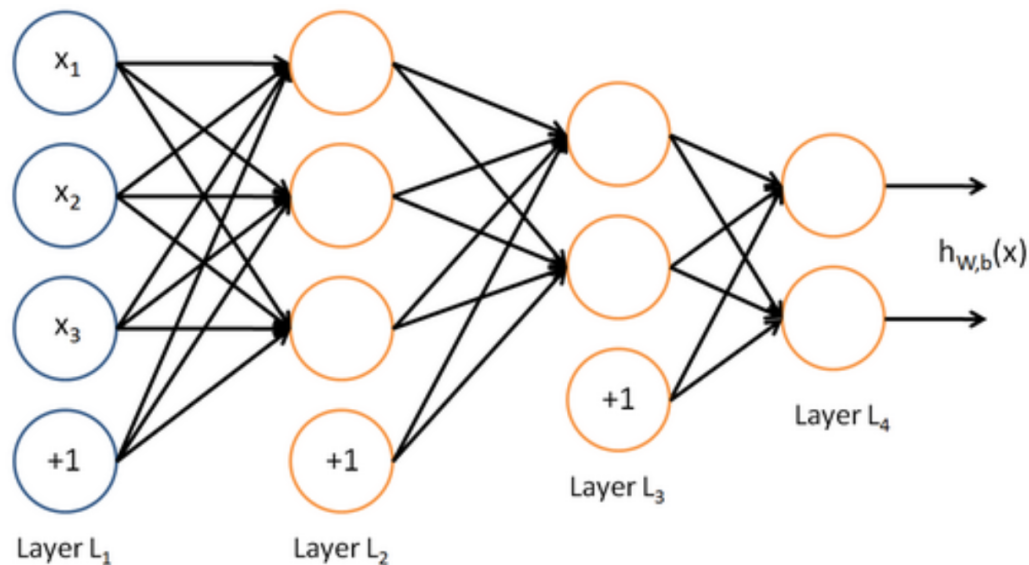

Diagram MS

School of Information Studies
Syracuse University

Reasons for Deep Learning

Break the bottleneck of manually designed features to automatically learn them

- Easy to adapt and to use

Can use large amounts of unsupervised data (e.g. raw text) to learn features and then use supervised data (with labels, like positive and negative) to learn a task

Deep learning ideas have been known but only recently outperforming other techniques

- Benefit from lots of data
- Multi-core machines with faster processors
- New models and algorithms