PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

Machine Learning Techniques

DATASCI 420

Lesson 09-1 Deep Learning I: Overview

Agenda

- Conceptual overview of deep learning
- How it works-When to use it?
- Types of DNNs and frameworks
- Deep Recurrent NNs
- Deep Convolutional NNs
- Use case

Why is Deep Learning Suddenly So Popular?

DL models have been around for a long time

- Fukushima (1980) – Neo-Cognitron
- LeCun (1989) – Convolutional Neural Network

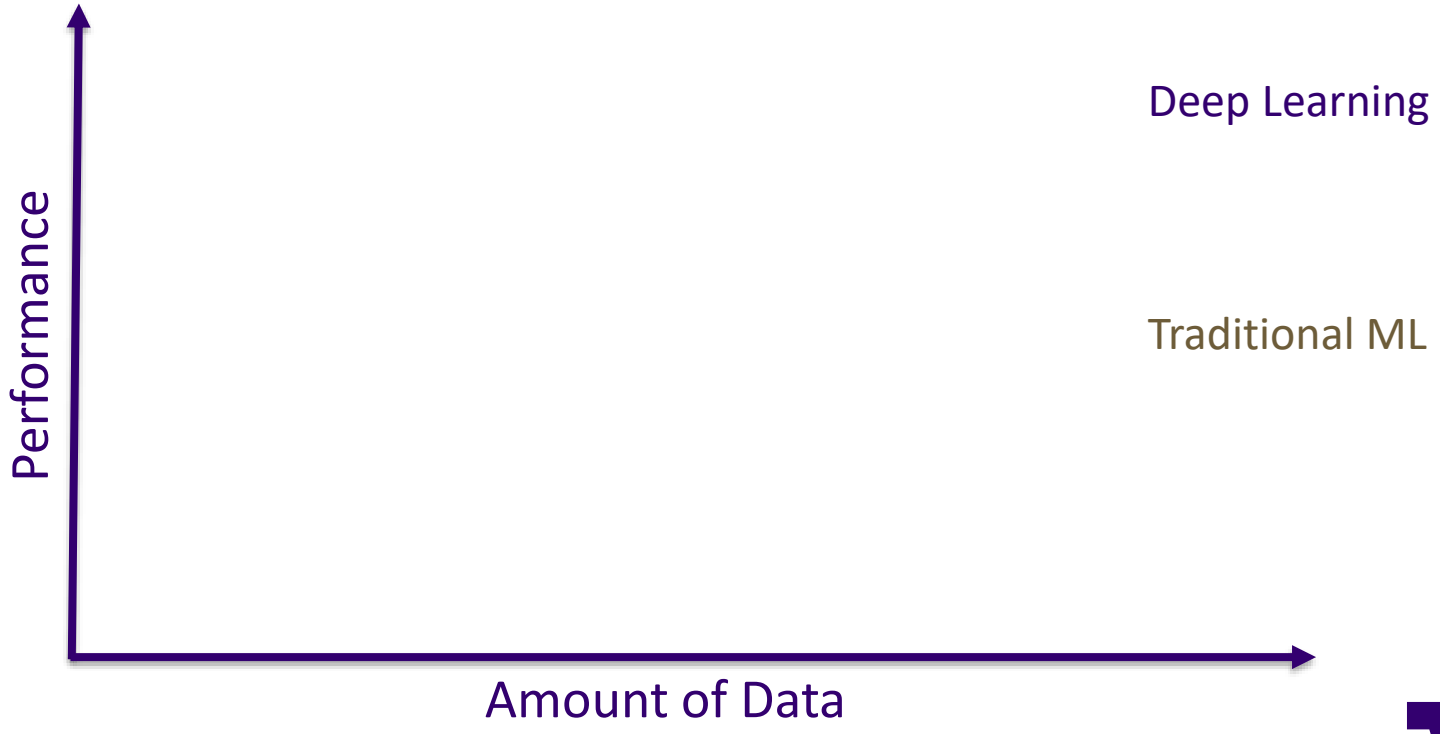
DL popularity grew recently due to:

- Availability of Big Data
- More powerful GPUs
- Runaway success in machine learning competitions

Motivation: Why Go Deep

- Deep representations may map more closely to the problem space
  - Allows for non-local generalization
  - Comprehensibility
- Deep architectures can be representationally efficient for certain types of applications
  - Fewer computational units are required for similar or better capability
- Multiple levels of latent variables allow combinatorial sharing of statistical strength
- Deep architectures work well (vision, audio, NLP, etc.)

When to use deep learning



Traditional Approach

Feature extractors, required:

- Expert knowledge
- Time-consuming hand-tuning
- In industrial applications, this is 70-80% of the time
- Sometimes are problem specific

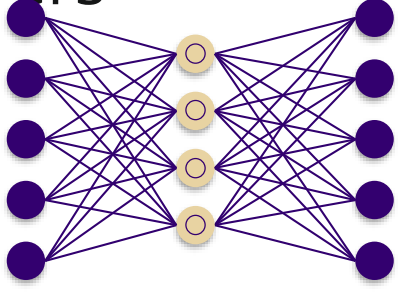
But, what if we could learn feature extractors?

Restricted Boltzmann Overview

- RB machines are useful for dimensionality reduction, classification, regression, collaborative filtering, feature learning and topic modeling
- An RB machine is a shallow, two-layer neural net (visible and hidden) which are the building blocks of deep-belief networks
- Differs from the neural networks:
  - Connections going both ways (forward *and* backward)
  - Activation function for each output node (as opposed to summing them)
  - Generative, as opposed to discriminative

Autoencoders

- Symmetric neural network—where all output units are connected back to the input units
- The middle layer is a “compressed” representation of the data where latent features can be learnt
- It works similarly to PCA, except that it is able, depending on the depth and the activation function, to learn complex non-linear relationships between the input and output layers



Summary

- Conceptual overview of deep learning
  - What it is
  - Why it is important
  - When to use it
- We compared deep learning to traditional machine learning approaches
- Discussed two types of neural networks, used for latent feature identification