Introduction to the course

why is machine learning so prevalent today?

It turns out that machine learning is a field that had grown out of the field of AI, or artificial intelligence. We wanted to build intelligent machines and it turns out that there are a few basic things that we could program a machine to do such as how to find the shortest path from A to B.

But for the most part we just did not know how to write AI programs to do the more interesting things such as web search or photo tagging or email anti-spam. There was a realization that the only way to do these things was to have a machine learn to do it by itself. So, machine learning was developed as a new capability for computers and today it touches many segments of industry and basic science.

Machine Learning:

- Grew out of work in AI
- New capability for computers

Examples:

- Database mining Large datasets from growth of automation/web.
  - *E.g., Web click data, medical records, biology, engineering*

    Click data: Tons of Silicon Valley companies are today collecting web click data, also called clickstream data, and are trying to use machine learning algorithms to mine this data to understand the users better and to serve the users better.

    Medical records: With the advent of automation, we now have electronic medical records, so if we can turn medical records into medical knowledge, then we can start to understand disease better.

    Computational biology: Biologists are collecting lots of data about gene sequences, DNA sequences, and so on, and machines running algorithms are giving us a much better understanding of the human genome, and what it means to be human.

    Engineer: In all fields of engineering, we have larger and larger, and larger and larger data sets, that we're trying to understand using learning algorithms.
- Applications can't program by hand
  - *E.g., Autonomus helicopter, handwriting recognition, most of Natural Language Processing (NLP), Computer Vision.*
- Self-customizing programs
  - *E.g., Amazon, Netflix product recommendations*
- Understanding human learning (barin, realAI).

What is Machine Learning?

Machine Learning definition

Two definitions of Machine Learning are offered. Arthur Samuel described it as: "the field of study that gives computers the ability to learn without being explicitly programmed." This is an older, informal definition.

Tom Mitchell provides a more modern definition: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."

Example: playing checkers.

E = the experience of playing many games of checkers

T = the task of playing checkers.

P = the probability that the program will win the next game.

In general, any machine learning problem can be assigned to one of two broad classifications:

Supervised learning and Unsupervised learning.

Video Question: Suppose your email program watches which emails you do or do not mark as spam, and based on that learns how to better filter spam. What is the task T in this setting?

> (T) Classifying emails as spam or not spam.

- (E) Watching you label emails as spam or not spam.
- (P) The number (or fraction) of emails correctly classified as spam/not spam.
- None of the above - this is not a machine learning problem.

Machine Learning algorithms:

- Supervised learning
- Unsupervised learning

Others: Reinforcement learning, recommender systems.

Also talk about: Practical advice for applying learning algorithms

In supervised learning, the idea is we're going to teach the computer how to do something. Whereas in unsupervised learning, we're going to let it learn by itself.