

Lecture 0: Introduction

Agenda

- Course overview
- Why machine learning?
- What this course is (not) about

Overview

This course serves as an introduction to a **variety** of machine learning topics both from a **theoretical** and **applied** perspective.

Note that terms in **bold**:

- *Variety*: We will introduce a diverse collection of machine learning methods for solving real-world problems.
- *Theory*: For every method we study, we will emphasize understanding its fundamental properties: correctness, computational efficiency, potential ways to improve it, etc.
- *Applications*: We will illustrate the efficacy of machine learning methods in the context of how they impact applications related to specific domains, with an emphasis on applications from electrical and computer engineering.

The science of machine learning

Why focus on the *principles* of data processing? And why now?

As of today, ML (and more broadly, “data science”) is an off-the-charts hot topic. Virtually every scientific, engineering, and social discipline is undergoing a major foundational shift towards more “analytics” oriented.

There are several reasons why this is the case; the foremost being that computing devices and sensors have pervasively been ingrained our daily personal and professional lives, and that it is easier and cheaper to acquire and process data than ever before. Hence, it is common to hear and read pithy newspaper headlines about data analytics, such as this:

“Why data is the new oil.” (Fortune Magazine, June 2016)

Of course, qualitative claims of this scale must always be taken with a pinch of salt.

It is undoubtedly true that a renewed focus on data-driven analysis and decision making has had considerable impact on a lot of fields. It is also equally true that important questions remain unanswered, and that much of the ML tools used by practitioners are deployed as “black-boxes”.

The only way to place ML on a solid footing is to build it bottom-up from the principles upwards; in other words, ask the same foundational questions, as you would for any other scientific discipline.

What this course is (not) about

This course addresses ML from a few different angles (while not exclusively dwelling upon any single one of them):

- data engineering
- algorithms
- statistics
- optimization
- applications

Prior knowledge/experience in any of the above topics will be a plus, but not necessary; the course will be fairly self-contained.

Learning outcomes

At the end, we will have:

- become familiar with several commonly used methods for solving ML problems such as:
 - retrieval
 - classification
 - inference
 - visualization
 - planning
- implement these methods in Python code, and
- obtained hands-on experience for applying these methods on real-world datasets.