CS 8850: Advanced Machine Learning **Fall 2017**

Syllabus

INSTRUCTOR: DANIEL L. PIMENTEL-ALARCÓN

Introduction

Machine learning is essentially *estimation* with computers. When you train a neural network to classify images of dogs and cats, you are essentially trying to *estimate* a function f composed of cascaded inner products and non-linearities, such that given a new image $\mathbf{X}$

$$f(\mathbf{X}) = \begin{cases} 0 & \text{if } \mathbf{X} \text{ contains a dog,} \\ 1 & \text{if } \mathbf{X} \text{ contains a cat.} \end{cases}$$

A support vector machine would aim to *estimate* regions $\mathcal{D}$ and $\mathcal{C}$ (e.g., limited by a line) such that $\mathbf{X} \in \mathcal{D}$ (e.g., above a certain line) if $\mathbf{X}$ contains a dog, and $\mathbf{X} \in \mathcal{C}$ (e.g., below a certain line), if $\mathbf{X}$ contains a cat.

k-means clustering would *estimate* the centers of the dogs and cats images, and then decide whether $\mathbf{X}$ contains one or the other based on proximity. Mixture regression would *estimate* these centers as subspaces.

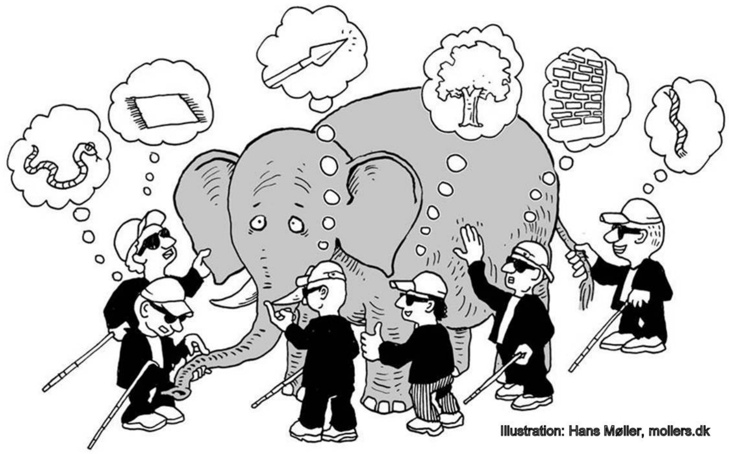
Feature selection and dimensionality reduction would aim to *estimate* a subset of components or characteristics in $\mathbf{X}$ that determine whether it contains a dog or a cat.

You get the picture. This course will get to the core of machine learning: estimation. This includes probabilistic modeling, hypothesis testing, maximum likelihood, bayesian inference, classification, regression, trending models, and state-of-the-art algorithms.

The goal is to gain a deep understanding of the fundamentals of machine learning. At the end, you will have insights on how machine learning works, why, when a method may be better than other, and how to adapt or tailor methods for a particular application. But more importantly, your clearer understanding of machine learning will allow you to develop new theory and methods. Who knows, maybe you'll discover a theorem answering a long open question in deep learning. Maybe you'll invent an ultimate machine learning algorithm with a fancy, catchy name that becomes the next support vector machine or neural network ;)

Instructor

Daniel L. Pimentel-Alarcón
email: pimentel@gsu.edu
office: 25 Park Place, Suite 742
website: `http://danielpimentel.github.io`



As we move along the course, we will dive into some math. We will see some equations, some theorems, some algorithms, etc. Knowing your math is important. It is important to understand the steps of an equation or the proof of a lemma. Each of these is an important **piece to solve a problem**. Don't loose track of the problem! It is equally important (or perhaps even more so) to **understand the big picture** and how the pieces come together.

Lectures

Monday, Wednesday
10:00am-11:45am
Aderhold Learning Center 330

Office Hours

Monday, 8:45am-9:45am,
Wednesday 11:45am-12:45pm.

Prerequisites: basic background in

- Linear Algebra
- Probability
- Statistics

Grading

- 20% Homework
- 20% Project
- 10% Midterm exam
- 50% Final exam

Topics to cover

- Probabilistic modeling
- Hypotheses testing
- Maximum likelihood
- Bayesian inference
- Classification
- Regression
- Trending models
- State-of-the-art algorithms

Lecture notes

The content of this course can be found in the lecture notes:

`https://danielpimentel.github.io/teaching.html`

Hint: skim through the lecture notes before the lecture. This gives you an edge, because:

- You can take your time to digest the main ideas and understand them better.
- You can ask your doubts in the lecture.
- Homework and exams will appear easier, and you will be able to do them much faster.

Homeworks

Homeworks are due Wednesdays at the **beginning** of lecture, and will be posted at

`https://danielpimentel.github.io/teaching.html`

Additional resources

- Statistical Signal Processing (Scharf).
- Fundamentals of Statistical Signal Processing (Kay).
- Elements of Statistical Learning (Hastie, Tibshirani, and Friedman).
- Pattern Recognition and Machine Learning (Bishop).
- A Probabilistic Theory of Pattern Recognition (Devroye, Gyorfi, and Lugosi).
- Elements of Information Theory (Cover and Thomas).
- Statistical Inference (George Casella and Roger L. Berger).

Remarks

- The course syllabus provides a general plan for the course; deviations may be necessary.
- Your constructive assessment of this course plays an indispensable role in shaping education at Georgia State. Upon completing the course, please take the time to fill out the online course evaluation.
- Students who wish to request accommodation for a disability may do so by registering with the Office of Disability Services. Students may only be accommodated upon issuance by the Office of Disability Services of a signed Accommodation Plan and are responsible for providing a copy of that plan to instructors of all classes in which accommodations are sought.
- Students are strongly encouraged to work together on homework assignments, but each student must submit his or her own writeup. Plagiarism of material written by classmates, book or article authors, or web posters is prohibited. Students must work independently on exams. Academic integrity will be strictly enforced.