S&DS 365 / 665

Intermediate Machine Learning

Course Overview

Wednesday, January 26

Yale

Outline

- Overview of course
- Topics
- Syllabus and logistics

Course objectives

Gain a solid understanding of concepts and methods of modern machine learning

- Become comfortable with the core ideas (and notation!)
- Gain an understanding of:
  - How and why methods work (and sometimes don't!)
  - Which models are appropriate for a given problem
  - How to adapt and extend methods when needed
- Get close to the research frontier

What does “Intermediate” imply?

- A second course in machine learning
- Assume familiar with things like PCA, bias/variance, basics of neural nets
- Have played around with some ML methods on some data sets
- Previous exposure to Python helpful
- More on this later...

What's the difference between 265 and 365?

- Little overlap in topics
- IML will be more technical/mathematical than iML
- We'll do some theory (but not for theory's sake!)
- The courses have similar organization (lectures, demos, assignments...)
- IML will be required for S&DS majors (BS)
- More on differences later...

$265 - 365 = -100$

What is Machine Learning?

The study of algorithms and statistical models to develop computer programs that improve with experience.

What is Machine Learning?

Machine Learning is closely aligned with Statistics, but with a focus on computation, scalability, prediction, representation, and complex problems

- Speech recognition
- Machine translation
- Object recognition and scene classification
- Autonomous driving...

Subproblems of these and other complex problems are concrete, statistical estimation and inference problems that can be studied in isolation.

AI vs. ML

Machine learning focuses on making predictions and inferences from data.

AI combines machine learning components into a larger system that includes a decision making component.

An AI system exhibits a behavior, resulting from the collective decisions that are made.

Machine Learning and AI

The focus in AI courses in the computer science curriculum used to be very different:

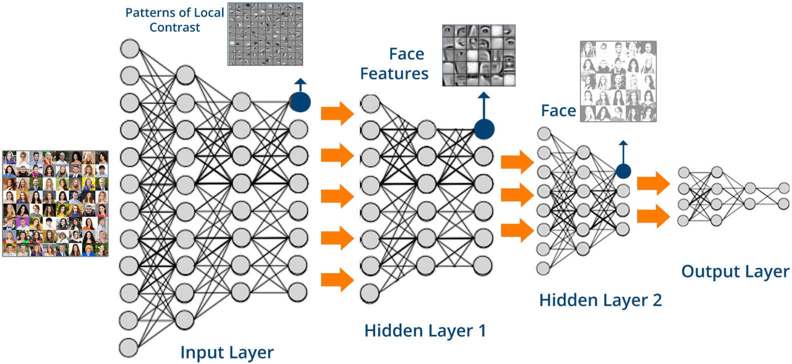
- Logic
- Search
- Games

Now much more “core” ML material than can be fit in a one semester course.

Machine learning frameworks

- Supervised learning
- Unsupervised (and semi-supervised) learning
- Reinforcement learning
- {Representation learning}

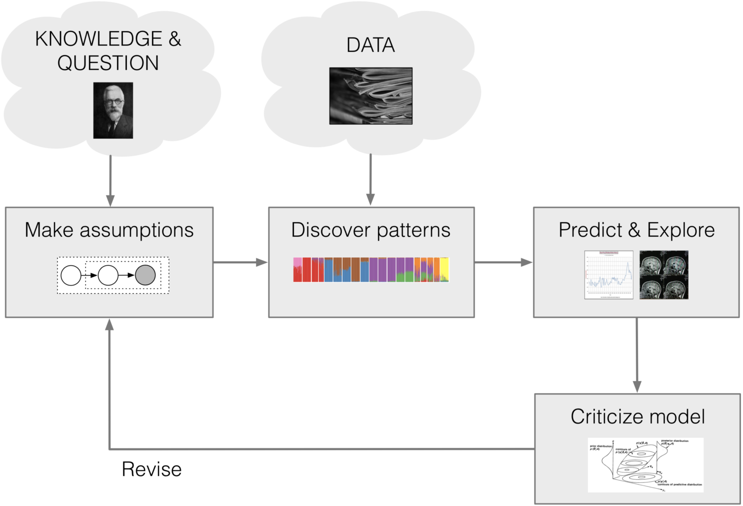
Deep learning is a type of machine learning



- Heuristics motivated from simplified view of the brain
- A particular form of nonlinear classification/regression/density estimation

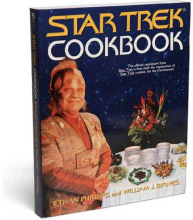
This course

- Part 1: Supervised learning
- Part 2: Unsupervised learning
- Part 3: Reinforcement learning
- Part 4: Sequence learning



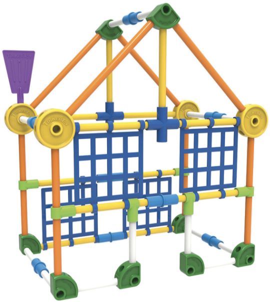
Ack: Dave Blei

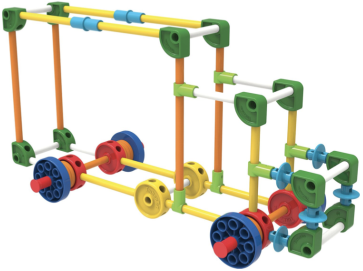
Machine learning *modus operandi* (1/2)



Increasingly powerful computing platforms (Keras, Tensorflow, Pytorch) implementing a few standard architectures.

Machine learning *modus operandi* (2/2)





More challenging to design and implement. Computation typically not "off the shelf"

Logistics

IML Team

- Instructor: John Lafferty
- Teaching Fellows

  Javid Dadashkarimi (PhD student, Computer Science)
  Wendy Luo (PhD student, Biomedical Engineering)
  Tianhao Wang (PhD Student, Statistics and Data Science)

- ULAs

  Henry Smith
  Sarah Zhao
  Malcolm Tang

- Office hours posted on Canvas; OH by Zoom (at least initially)

Course materials

Materials posted to `http://interml.ydata123.org`; sometimes to Canvas

Please use Ed Discussion for any questions about lectures, homework, etc.

Please use email for questions about grading and grades.

Syllabus

Intermediate Machine Learning is a second course in machine learning at the advanced undergraduate or beginning graduate level. The course assumes familiarity with the basic ideas and techniques in machine learning, for example as covered in S&DS 265. The course treats methods together with mathematical frameworks that provide intuition and justifications for how and when the methods work. Assignments give students hands-on experience with machine learning techniques, to build the skills needed to adapt approaches to new problems.

Syllabus

Topics include nonparametric regression and classification, kernel methods, risk bounds, nonparametric Bayesian approaches, graphical models, attention and language models, generative models, sparsity and manifolds, and reinforcement learning. Programming is central to the course, and is based on the Python programming language and Jupyter notebooks.

Prerequisites

- Background in probability and statistics, at the level of S&DS 242 (Theory of Statistics)
- Familiarity with the core ideas from linear algebra, for example through Math 222 (Linear Algebra with Applications)
- Computational skills at the level of S&DS 265 (Introductory Machine Learning) or CPSC 200 (Introduction to Information Systems)
- Previous familiarity with Python is recommended

Installing Jupyter

- A beginner's guide to installing Python and Jupyter on your computer is here: `https://bit.ly/22KVCfsV`
- See installation guide on course Canvas site: `Files > Getting Started`
- Use Python 3.x version

Assessment

- Assignments (50%)
- Mid-semester exam (20%)
- Final exam: 20%
- Quizzes (10%)

Assignments

- Four assignments total
- Roughly every 2 weeks
- Due at 11:59pm on a Wednesday
- Late assignments not accepted
- Submitted using Gradescope
- Mix of concepts, problem solving and data analysis
- Prepared using Python notebooks

Collaboration

Collaboration on homework assignments with fellow students is encouraged. However, such collaboration should be clearly acknowledged, by listing the names of the students with whom you have had any discussions concerning the problem. *You may not share written work or code—after discussing a problem with others, the solution must be written by yourself.*

Quizzes

- Four quizzes total
- Taken online (Canvas)
- Short, 10-20 minutes
- Assess understanding of essentials

Week	Dates	Topics	Lecture Materials	Assignments & Exams
1	Jan 26, 28	Course overview		
2	Jan 31, Feb 2	Smoothing and kernels		
3	Feb 7, 9	Neural networks for classification		Feb 9: Assn 1 out
4	Feb 14, 16	Risk bounds and generalization error		Feb 16: Quiz 1
5	Feb 21, 23	Nonparametric Bayes		Feb 23: Assn 1 in; Assn 2 out
6	Feb 28, Mar 2	Approximate inference		Mar 2: Quiz 2
7	Mar 7, 9	Approaches to generative modeling		Mar 9: Assn 2 in
8	Mar 14, 16	Structure learning		Mar 16: Midterm exam

9	Mar 28, 30	Deep reinforcement learning		Mar 30: Assn 3 out
10	Apr 4, 6	Policy gradient methods		Apr 6: Quiz 3
11	Apr 11, 13	Sequential and sequence-to-sequence models		Apr 13: Assn 3 in; Assn 4 out
12	Apr 18, 20	Attention and language models		Apr 20: Quiz 4
13	Apr 25, 27	Special topic: Machine learning in neuroscience		Apr 27: Assn 4 in
	May 7	Final exam, 2pm location TBD		Registrar: final exam schedule

Exams

- Midterm: Wednesday, March 16, in class
- Final: Saturday, May 7 at 2pm

http://catalog.yale.edu/ycps/final-examination-schedules/

Auditing

- Auditors are welcome!
- Full access to Canvas
- Just expected to regularly attend class

Topics covered

- Part 1: Supervised learning
- Part 2: Unsupervised learning
- Part 3: Reinforcement learning
- Part 4: Sequence learning

Topics covered

- *Part 1: Supervised learning*
  - Sparse regression
  - Smoothing and kernels
  - Convolutional neural networks
  - Risk bounds and generalization error
- Part 2: Unsupervised learning
- Part 3: Reinforcement learning
- Part 4: Sequence learning

Topics covered

- Part 1: Supervised learning
- *Part 2: Unsupervised learning*
  - Nonparametric Bayes
  - Approximate inference
  - Approaches to generative models
  - Structure learning
- Part 3: Reinforcement learning
- Part 4: Sequence learning

Topics covered

- Part 1: Supervised learning
- Part 2: Unsupervised learning
- *Part 3: Reinforcement learning*
  - Deep Q-Learning
  - Policy gradient methods
- Part 4: Sequence learning

Topics covered

- Part 1: Supervised learning
- Part 2: Unsupervised learning
- Part 3: Reinforcement learning
- *Part 4: Sequence learning*
  - Classical techniques (Kalman filters, HMMs)
  - LSTMs and GRUs
  - Attention and language models

References

- "The Elements of Statistical Learning: Data Mining, Inference, and Prediction," by T. Hastie, R. Tibshirani, and J. Friedman,

  http://www-stat.stanford.edu/~tibs/ElemStatLearn/
- "Probabilistic Machine Learning: An Introduction," by K. Murphy, MIT Press, https://probml.github.io/pml-book/book1.html
- "Pattern Recognition and Machine Learning," C. Bishop, Springer, https://www.microsoft.com/en-us/research/uploads/prod/2006/01/

  Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf

Questions?