S&DS 365 / 665

**Intermediate Machine Learning**

# Course Overview

Wednesday, August 31

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, maximum likelihood, basics of neural nets
- Have experimented with basic ML methods on some data sets
- Previous exposure to Python
- More on this later...

# What's the difference between 265 and 365?

- Little overlap in topics
- IML is 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 is required for S&DS majors (BS); iML aims to be accessible to broad cross-section of Yale students
- More on differences later...

$265 - 365 = -100$

# IML Team

- Instructor: John Lafferty
- Teaching Fellows
  Louis Deschuttere (Masters student, S&DS)
  Chris Xu (PhD student, S&DS)
  Sophia Zhu (Masters student, S&DS)
- ULAs
  Benjamin Christensen
  Jacques Morris
  Eric Sun
- Course Manager
  Siddhartha Chatterjee

# 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.

# 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:

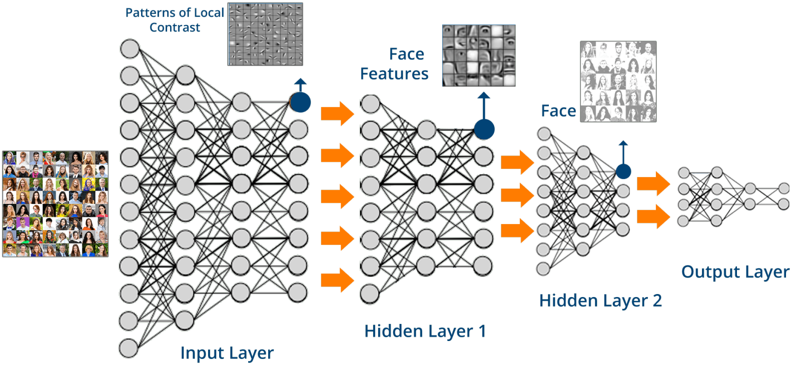
- Logic
- Search
- Games

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

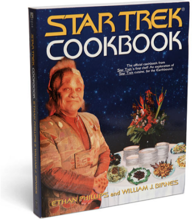



- Heuristics motivated from simplified view of the brain
- A particular form of nonlinear classification/regression/density estimation

# This course

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

# Machine learning *modus operandi* (1/2)

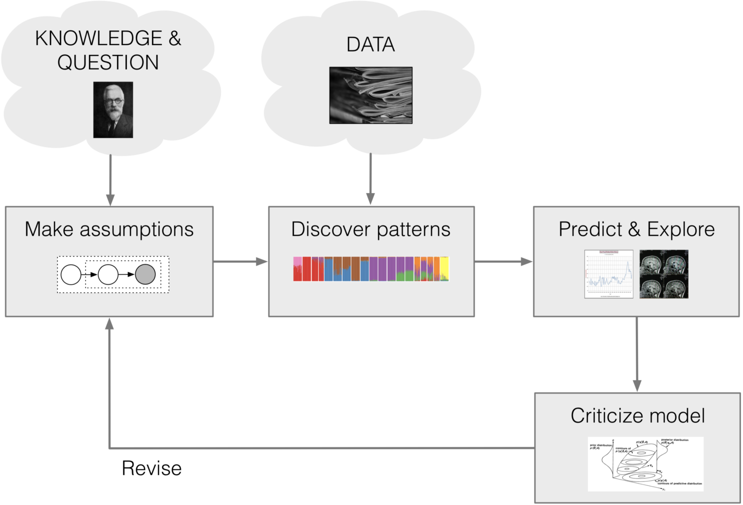




Increasingly powerful computing platforms (Keras, Tensorflow, Pytorch) implementing a few standard architectures.

# **Machine learning** *modus operandi* **(2/2)**

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




Ack: Dave Blei

# Logistics

# IML Team

- Instructor: John Lafferty
- Teaching Fellows
  Louis Deschuttere (Masters student, S&DS)
  Chris Xu (PhD student, Biomedical Engineering)
  Sophia Zhu (Masters student, S&DS)
- ULAs
  Benjamin Christensen
  Jacques Morris
  Eric Sun
- Course Manager
  Siddhartha Chatterjee

# 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

- 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 due on a Wednesday at midnight. Late assignments not accepted.

# Assignments

- Five 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

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

`http://interml.ydata123.org`

| Week | Dates | Topics | Demos & Tutorials | Lecture Slides | Readings & Notes | Assignments & Exams |
|---|---|---|---|---|---|---|
| 1 | Aug 31, Sep 2 | Course overview | CO Python elements<br>CO Pandas and regression<br>CO Lasso example | Aug 31: Course overview<br>Sep 2: Sparse regression | PML Section 11.4 | |
| 2 | Sep 7 | Smoothing and kernels | CO Smoothing example<br>CO Using different kernels | Sep 7: Smoothing | PML Sections 16.3, 17.1 | Sep 7: Quiz 1 |
| 3 | Sep 12, 14 | Density estimation and risk bounds | CO Mercer kernels<br>CO Density estimation demo | Sep 12: Mercer Kernels<br>Sep 14: Density estimation | Notes on Mercer kernels<br>Bias-variance tradeoff for density estimation | Sep 14: CO Assn 1 out |
| 4 | Sep 19, 21 | Neural networks for classification | TensorFlow playground<br>CO Convolution demo | Sep 19: Neural networks<br>Sep 21: Convolutional neural networks | PML Sections 13.1, 13.2<br>Notes on backpropagation | Sep 21: Quiz 2 |
| 5 | Sep 26, 28 | Nonparametric Bayes | CO Parametric Bayes<br>CO Gaussian processes | Sep 26: Gaussian processes<br>Sep 28: Gaussian processes continued | PML Section 17.2<br>Notes on Bayesian inference<br>Notes on nonparametric Bayes | Sep 28: Assn 1 in; CO Assn 2 out |
| 6 | Oct 3, 5 | Gibbs sampling and approximate inference | CO Gibbs sampling | Oct 3: Gibbs sampling<br>Oct 5: Introduction to approximate inference | Notes on Gibbs sampling | Oct 5: Quiz 3 |
| 7 | Oct 10, 12 | Variational inference | CO Variational autoencoders | Oct 10: Variational inference and VAEs<br>Oct 12: VAEs and review | PML Section 20.3<br>Notes on variational inference | Oct 12: Assn 2 in<br>CO Assn 3 out |

| | | | | | | |
|---|---|---|---|---|---|---|
| 8 | Oct 17 | Midterm | | | Practice midterm | Oct 17: Midterm exam |
| 9 | Oct 24, 26 | Graphs and structure learning | Graphical lasso demo<br>Graph neural networks | Oct 24: Sparsity and graphs<br>Oct 26: Discrete data and graph neural nets | Notes on graphs and structure learning<br>PML Section 23.4 | |
| 10 | Oct 31, Nov 2 | Deep reinforcement learning | Q-learning demo | Oct 31: Reinforcement learning<br>Nov 2: Deep reinforcement learning | Sutton and Barto, Section 6.5 | Nov 2: Assn 3 in<br>Assn 4 out |
| 11 | Nov 7, 9 | Policy gradient methods | DQN demo<br>Policy gradients demo<br>Actor-critic demo | Nov 7: Policy gradient methods<br>Nov 9: Actor-critic methods | Sutton and Barto, Section 13.1-13.3, 13.5 | Nov 9: Quiz 4 |
| 12 | Nov 14, 16 | Sequential and sequence-to-sequence models | RNN demo: Fakespeare<br>TensorFlow: Text generation | Nov 14: HMMs and RNNs<br>Nov 16: RNNs, GRUs, LSTMs, and all that | PML Chapter 15 | Nov 16: Assn 4 in<br>Assn 5 out |
| 13 | Nov 21, 23 | No class, Thanksgiving break | | | | |
| 14 | Nov 28, 30 | Attention and language models | GPT-3 demo<br>Codex demo | Nov 28: Sequence-to-sequence models<br>Nov 30: Attention | PML Sections 15.4, 15.5 | Nov 30: Quiz 5 |
| 15 | Dec 5, 7 | Transformers | Transformer demo | Dec 5: Attention and transformers<br>Dec 7: Course wrap up | | Dec 7: Assn 5 in |

# Exams

- Midterm: Monday, October 17, in class
- Final: Tuesday Dec. 20, 7:00pm

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

# Auditing

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

# Questions on logistics?

# 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
  - Actor-Critic approaches
- 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
  - Transformers

# 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

**Questions?**

# Questions?

*Remember: Class on Friday!*