

CS 6316 Machine Learning

Introduction

Yangfeng Ji

Information and Language Processing Lab
Department of Computer Science
University of Virginia



▶ Instructor:

- ▶ Yangfeng Ji
- ▶ Office: Rice 510

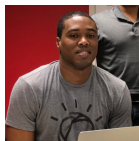
▶ Teaching assistants



Stephanie Schoch



Wanyu Du



Dane Williamson

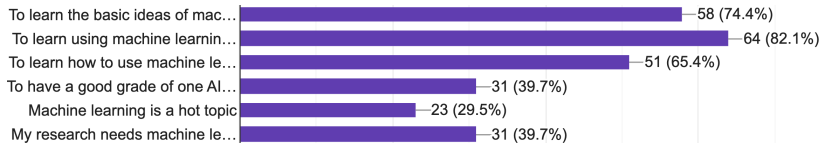
We will have one more TA for this class.

- ▶ Calculus and Linear Algebra
Multivariable derivatives, matrix-vector notations and operations; singular value decomposition, etc.

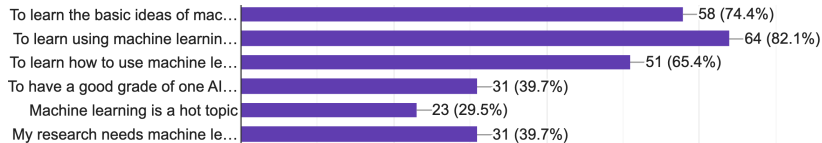
- ▶ Calculus and Linear Algebra
Multivariable derivatives, matrix-vector notations and operations; singular value decomposition, etc.
- ▶ Probability and Statistics
Mean and variance, multinomial distribution, conditional dependence, maximum likelihood estimation, Bayes theorem, etc.

- ▶ Calculus and Linear Algebra
Multivariable derivatives, matrix-vector notations and operations; singular value decomposition, etc.
- ▶ Probability and Statistics
Mean and variance, multinomial distribution, conditional dependence, maximum likelihood estimation, Bayes theorem, etc.
- ▶ Proficiency in Python
Some machine learning packages, e.g., sklearn and PyTorch/Tensorflow

The survey results (by Jan. 19, 4 PM)



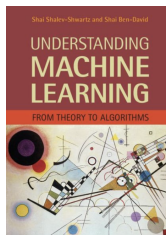
The survey results (by Jan. 19, 4 PM)



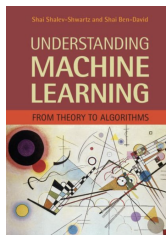
Top three answers

- ▶ To learn using machine learning to solve real-world problems (64 votes)
- ▶ To learn the basic ideas of machine learning (58 votes)
- ▶ To learn how to use machine learning packages (Sklearn, PyTorch, etc.) (51 votes)

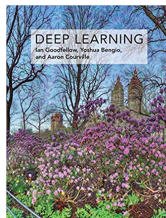
- ▶ Shalev-Shwartz and Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. 2014



- ▶ Shalev-Shwartz and Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. 2014



- ▶ Goodfellow, Bengio, and Courville. *Deep Learning*. 2016



This course will cover the basic materials on the following topics

1. Introduction to learning theory
2. Linear classification and regression
3. Model selection and validation
4. Support vector machines and kernel methods
5. Optimization methods
6. Neural networks and deep learning

The following topics will **not** be the emphasis of this course

- ▶ Statistical modeling
 - ▶ e.g., parameter estimation, Bayesian statistics, graphical models

The following topics will **not** be the emphasis of this course

- ▶ Statistical modeling
 - ▶ e.g., parameter estimation, Bayesian statistics, graphical models
- ▶ Machine learning engineering
 - ▶ e.g., how to implement a image classifier from end to end
 - ▶ although, we will provide some demo code for illustration purposes

The following topics will **not** be the emphasis of this course

- ▶ Statistical modeling
 - ▶ e.g., parameter estimation, Bayesian statistics, graphical models
- ▶ Machine learning engineering
 - ▶ e.g., how to implement a image classifier from end to end
 - ▶ although, we will provide some demo code for illustration purposes
- ▶ Advanced topics in machine learning
 - ▶ e.g., Reinforcement learning, active learning, semi-supervised learning, online learning
 - ▶ CS 6501 / 4501 *Interpretable Machine Learning* by Hanjie Chen and Yangfeng Ji

For students looking for additional reading materials

- ▶ Bishop. Pattern Recognition and Machine Learning. 2006
- ▶ Murphy. Machine Learning: A Probabilistic Perspective. 2012
- ▶ Mohri, Rostamizadeh, and Talwalkar. Foundations of Machine Learning. 2nd Edition. 2018
- ▶ Hastie, Tibshirani, and Friedman. The Elements of Statistical Learning (2nd Edition). 2009

Homework and Grading Policy

- ▶ Homeworks (72%)
 - ▶ Four homeworks, each of them worth 18%
- ▶ Final project (25%)
 - ▶ Project proposal: 8%
 - ▶ Final project presentation: 7%
 - ▶ Final project report: 10%
- ▶ Class attendance (3%): we will take attendance at three randomly-selected lectures. Each is worth 1%

Grading Policy

The final grade is threshold-based instead of percentage-based

Point range	Letter grade
[99 100]	A+
[94 99)	A
[90 94)	A-
[88 90)	B+
[83 88)	B
[80 83)	B-
[74 80)	C+
[67 74)	C
[60 67)	C-

Late Penalty

- ▶ Homework submission will be accepted up to 72 hours late, with **20% deduction** per 24 hours on the points as a penalty
- ▶ Submission will **not** be accepted if more than 72 hours late
- ▶ Make sure not submit wrong files
 - ▶ it is students responsibility to make sure they submit the right and complete files for each homework
- ▶ It is usually better if students just turn in what they have in time

Violation of the Honor Code

Plagiarism, examples are

- ▶ in a homework submission, copying answers from others directly (including, some minor changes)
- ▶ in a report, copying texts from a published paper (including, some minor changes)
- ▶ in a code, using someone else's functions/implementations without acknowledging the contribution

- ▶ Course webpage

<http://yangfengji.net/uva-ml-course/>

which contains all the information you need about this course.

- ▶ Collab

- ▶ For releasing homeworks and grading

- ▶ Campuswire

- ▶ For announcement, online QA, discussion, etc.

- ▶ Using the following link and code “0709”

<https://campuswire.com/p/GE924A494>

- ▶ Or, let our TAs know if you have received the invitation link

Now, let's have some fun!

Warning: you will see lots of mathematical notations.