

CS 780/880 Introduction to Machine Learning

Spring 2017

Website: http://cs.unh.edu/~mpetrik/teaching/intro_ml_17/

Overview

This class will teach you how to use machine learning to understand data and make predictions. The class will focus on understanding the fundamental concepts and algorithms that underlie modern machine learning and data science algorithms.

Contact Information

Marek Petrik	
Email:	mailto:mpetrik@cs.unh.edu
Office:	Kingsbury W233
Office hours:	Thursday: 2pm - 3:30pm
Discussion forum:	https://piazza.com/unh/spring2017/cs780cs880

The fastest way to get your question about course material answered is to post it on piazza (see the link above).

Textbooks

ISL James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning [online pdf]
ESL Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning. Springer Series in Statistics (2nd ed.) [online pdf]

Other references

DL Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. [online pdf on github]
LA Strang, G. Introduction to Linear Algebra. (2016)
CO Boyd, S., & Vandenberghe, L. (2004). Convex Optimization. [online pdf]
RL Sutton, R. S., & Barto, A. (2012). Reinforcement learning. 2nd edition [online pdf draft]

Evaluation

The weights of individual components in computing the grade are as follows:

50%	6 Assignments
15%	Midterm exam
30%	Final exam
15%	Class project

Note that the weights sum to 110%. Each assignment will have slightly different questions for the graduate and undergraduate version of the class.

Disability If you are registered with the student disability office, please let me know right away so that I can provide proper accommodation.

Academic honesty

- University Academic Honesty Policy:
<https://www.unh.edu/student-life/handbook/academic/academic-honesty>
- Tutorial on Plagiarism:
<http://cola.unh.edu/plagiarism-tutorial-0>

Assignments

General policies

- Assignments are due on the day indicated **by 12:40pm**
- Assignments should be turned in as a PDF on myCourses, or printed and turned in at class; if other methods fail, email to <mailto:mpetrik@cs.unh.edu> with Subject that contains the string [CS780880HW]
- There is no credit for assignments turned in late; This is so we can discuss solutions in timely manner.
- Collaboration when solving assignments is *encouraged* but all of the writing must be yours. You may collaborate via Piazza or in person.
- Please do not copy solutions from online sources; it defeats the purpose of the assignments and you will not do well in the exams.

Programming Assignments The class will involve hands-on data analysis using machine learning methods. The recommended language for programming assignments is R which is an excellent tool for statistical analysis and machine learning. No prior knowledge of R is needed or expected; the book and lectures will include a gentle introduction to the language.

Exams

The final exam will be take-home over a period of several days. The exam will require about 3 hours to complete. No collaboration on the exams.

Tentative Syllabus

Please check for updates on the website. The dates on the class website are authoritative

Date	Day	Topic	Reading	Assignment
Jan 24 Jan 26	Tue Thu	<i>Snow day</i> Statistical learning and R language	ISL 1,2	
Jan 31 Feb 02	Tue Thu	Linear regression basics <i>No class</i>	ISL 3.1-2	
Feb 07 Feb 09	Tue Thu	Linear regression advanced Classification and logistic regression	ISL 3.3-6 ISL 4.1-3	
Feb 14 Mar 16	Tue Thu	Classification, naive Bayes and LDA Linear algebra for machine learning: review	ISL 4.4-6 LA 3,4,6	1
Feb 21 Feb 23	Tue Thu	Linear algebra and optimization Overfitting and resampling methods	CO 2,3 ISL 5.1	2 project description, data
Feb 28 Mar 02	Tue Thu	Cross-validation and bootstrapping Linear model selection, priors	ISL 5.2 ISL 6.1	evaluation
Mar 07 Mar 09	Tue Thu	Midterm review Midterm exam ; material until 2/23		3
Mar 14 Mar 16	Tue Thu	<i>Spring break, no class</i> <i>Spring break, no class</i>		
Mar 21 Mar 23	Tue Thu	Linear model selection and regularization Building nonlinear features	ISL 6.2 ISL 7.1-3	lit. review and method
Mar 28 Mar 30	Tue Thu	Nearest neighbor methods and GAMs Tree-based methods and boosting	ISL 7.4-7 ISL 8	4
Apr 04 Apr 06	Tue Thu	Support vector machines and other techniques Unsupervised learning, PCA	ISL 9 ISL 10	preliminary results
Apr 11 Apr 13	Tue Thu	Reinforcement learning Neural networks and deep learning	RL 1,3,4 DL	5
Apr 18 Apr 20	Tue Thu	Neural networks and deep learning Big data and machine learning	DL	
Apr 25 Apr 27	Tue Thu	Machine learning in practice Project presentations		6 final report
May 02 May 04	Tue Thu	Guest speaker Final exam review		
May 11-17	?	Final exam		