Introduction to Machine Learning

Spring 2024

Welcome to CSE/STAT 416: Introduction to Machine Learning!

Registration

Do not email the course staff or instructor requesting an add-code for the course. The course staff do not have any add-codes. Please see the Registration FAQ for answers to common registration questions.

Announcements¶

▼ Mar 27

Welcome to CSE/STAT 416

Information about the class posted including:

- Links to the course website (here) and other resources
- Information about the class structure. No pre-class work for the first class.
- Office hours start Monday, April 1st.

See the full announcement on Ed!

This Week (at a glance)¶

Monday (05/27)

- Learning Reflection 8 Due

Wednesday (05/29)

- Course Review: Victory Lap

Thursday (05/30)

- Homework 7 Due

Calendar¶

Info

This is a rough sketch of the quarter and things are subject to change. We can accurately predict the past, but predicting the future is hard!

Lessons

Anything listed in the “Lesson” materials for a day should be done before attending class that day. We recommend doing all the slides before the “Pause and Think” slide. Each class session will start by reviewing what was in the

Lesson and then most time will be spent on working on practice problems in the Lessons. See the syllabus for more info!

Jump to Today | Expand all Below

	Topic	Homeworks	Learning Reflections
Module 0 - Introduction / Regression			
Wed 03/27	LES 00 Regression Note: Normally there is are pre-class materials that you should complete before attending class. For the first day there are none! You should complete the Checkpoint after class (due before the next class). Note (2): The CSE/STAT 416 Training links for today are extremely useful resources to make sure you are prepared for the class. The course will be in Python, so Resources 1 and 2 will be very helpful if you are less familiar with the language. Resource 3 covers the mathematical background we expect students to be comfortable with in the course. You are expected to complete these trainings, but they are not factored into your grade. We encourage you to work on them now before assignments are released. pre-class: None megathread in-class: pdf annotated pptx post-class: checkpoint resources: videos extra resources **Videos** • Recording **Extra resources** • [Schafer] AnIML - Introduction (draft) ◦ *Note: This is a draft book for the course. It is still a work in progress so it will still have some errors and TODOs left throughout. Feedback is welcome and encouraged!* • [Schafer] AnIML - Linear Regression (draft) • [Nguyen] Previous Quarter Lecture Notes ◦ *Note: These notes are from a different offering of the course and may use slightly different notation and terminology.* • [ESL] Section 1, 2.3.1 ◦ *Note: The Elements of Statistical Learning [ESL] is an advanced machine learning textbook that we link to as optional readings for students who want to learn more of the nitty-gritty details behind the model's derivations. You do not need to understand readings labeled as "Optional"* CSE/STAT 416 Training: 1) Python Practice, 2) `numpy` Practice, 3) Math/Probability/Statistics Practice		Out LR0 Due 11:59 pm
Thu 03/28	SEC 00 Course Infrastructure; Pandas resources: handout LES 01 Assessing Performance; Bias + Variance Tradeoff pre-class: lesson megathread in-class: pdf annotated pptx demo post-class: checkpoint resources: videos extra resources advanced resources **Videos** • Recording **Extra resources** • [Schafer] AnIML - Assessing Performance (draft) • [Nguyen] Lecture Notes from previous quarter **Advanced resources** • [ESL] Section 2.3.1, 7.1-7.4		
Fri 03/29			
Mon 04/01			
Tue 04/02		Out HW0 House Prices Due 11:59 pm	
Module 1 - Assessing Performance			

LES 02 Regularization: Ridge
pre-class: lesson megathread
in-class: pdf annotated pptx demo
post-class: checkpoint
resources: videos extra resources advanced resources

Videos

- Recording

Wed 04/03

Extra resources

- [Schafer] AnIML - Ridge Regression (draft)
- [Nguyen] Lecture Notes from previous quarter

Advanced resources

- [ESL] Section 3.1-3.2, 3.4.1
- [ESL] Section 7.1-7.4

Thu 04/04 SEC 01 Ridge and LASSO; Code
resources: handout

Out LR1 Due 11:59 pm

LES 03 Regularization: LASSO, Feature selection
pre-class: lesson megathread
in-class: pdf annotated pptx demo
post-class: checkpoint
resources: videos extra resources advanced resources

Videos

- Recording

Fri 04/05

Extra resources

- [Nguyen] Lecture notes from previous quarter
- [Schafer] AnIML - LASSO and Feature Selection

Advanced resources

- [ESL] Section 2.9, 5.5.2, 7.2
- [ESL] Section 3.4.2, 7.10

Mon 04/08

Out HW1 Ridge and LASSO Due 11:59 pm

Tue 04/09

Module 2 - Classification

LES 04 Classification
pre-class: lesson megathread
in-class: pdf annotated pptx
post-class: checkpoint
resources: videos extra resources advanced resources

Videos

- Recording

Wed 04/10

Extra resources

- [Nguyen] Lecture notes from previous quarter
- [Schafer] AnIML - Classification

Advanced resources

- [ESL] Section 1, 2.3.1, 4.1-4.2

Thu 04/11 SEC 02 Classification ; Logistic Regression
resources: handout

Out LR2 Due 11:59 pm

LES 05 MLE / Logistic Regression
pre-class: lesson megathread
in-class: pdf annotated pptx demo
post-class: checkpoint
resources: videos extra resources advanced resources

Videos

- Recording

Fri 04/12

Extra resources

- [Nguyen] Lecture notes from a previous quarter
- [Schafer] AnIML - Logistic Regression

Advanced resources

- [ESL] Section 4.4.1-4.4.4, 9.1.2, 7.5-7.6

Mon 04/15

Tue 04/16

Out HW2 Sentiment Analysis with Logistic Regression Due 11:59 pm

Module 3 - Societal Impact, Bias, and Fairness

LES 06 Bias and Fairness
pre-class: lesson megathread
in-class: pdf annotated pptx
post-class: checkpoint megathread
resources: videos extra resources advanced resources

Videos

- Recording

Wed 04/17

Extra resources

- [Schafer] AnIML - Bias and Fairness (7.1-7.3)

Advanced resources

- People and AI Research (PAIR, Google)
- ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT)
- A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle (Suresh & Guttag 2019)

Out LR3 Due 11:59 pm

Thu 04/18

SEC 03 Midterm Review
resources: handout

LES 07 Fairness and Tradeoffs ; Recap
pre-class: lesson megathread
in-class: pdf annotated pptx
post-class: checkpoint megathread
resources: videos extra resources advanced resources

Videos

- Recording

Fri 04/19

Extra resources

- [Schafer] AnIML - Fairness and Tradeoffs (7.3-7.6)

Advanced resources

- The Ethical Algorithm - Michael Kearns & Aaron Roth
- On the (im)possibility of fairness (Friedler et al. 2016)

Mon 04/22

Tue 04/23

Out EXAM Midterm Due 11:59 pm

Module 4 - Trees/Ensemble Methods

LES 08 Naive Bayes / Decision Trees
pre-class: lesson megathread
in-class: pdf annotated pptx
post-class: checkpoint megathread
resources: videos extra resources

Videos

Wed 04/24

- Recording

Extra resources

- [Schafer] AnIML - Naïve Bayes
- [Schafer] AnIML - Decision Trees
- [Nguyen] Lecture notes from a previous quarter

Thu 04/25

SEC 04 Trees and Ensemble Methods
resources: handout

LES 09 Ensemble Methods
pre-class: lesson megathread
in-class: pdf annotated pptx
post-class: checkpoint megathread
resources: videos extra resources advanced resources

Out LR4 Due 11:59 pm

Videos

- Recording
- Lecture Supplement

Fri 04/26

Extra resources

- [Schafer] AnIML - Ensemble Methods (coming soon)
- [Nguyen] Lecture notes from a previous quarter

Advanced resources

- Deriving AdaBoost
- Explaining AdaBoost (Schapire 2013)
- [ESL] Section 9.2.4, 10.1-10.10

Mon 04/29

Out HW3 Loan Safety with Decision Trees Due 11:59 pm

Tue 04/30

Module 5 - Deep Learning

LES 10 Neural Networks
pre-class: lesson
in-class: pdf annotated pptx
post-class: checkpoint
resources: videos extra resources advanced resources

Videos

- Recording

Wed 05/01

Extra resources

- [Schafer] AnIML - Neural Networks (coming soon)
- [Nguyen] Lecture notes from a previous quarter
- Playground

Advanced resources

- Neural Network Notes
- Backpropogation Algorithm
- Proof of Approximating and Function

Thu 05/02

Out LR5 Due 11:59 pm

SEC 05 Deep Learning
resources: handout

LES 11 Deep Learning; Convolutional Neural Networks
pre-class: lesson megathread
in-class: pdf annotated pptx
post-class: checkpoint
resources: videos extra resources advanced resources

Videos

- Recording

Fri 05/03

Extra resources

- [Schafer] AnIML - Convolutional Neural Networks (coming soon)
- [Nguyen] Lecture notes from a previous quarter
- What is a Convolutional Neural Network?

Advanced resources

- CS231n: Convolutional Neural Networks for Visual Recognition

Mon 05/06

Out HW4 Deep Learning with PyTorch Due 11:59 pm

Tue 05/07

Module 6 - Non-Parametric Methods

LES 12 Precision + Recall / kNN
pre-class: lesson megathread

in-class: pdf annotated pptx
post-class: checkpoint
resources: videos extra resources advanced resources

Videos

- Recording

Wed 05/08

Extra resources

- [Schafer] AnIML - Precision + Recall / kNN (coming soon)

Advanced resources

- word2Vec and biased embeddings
- Auto-Encoders

Out LR6 Due 11:59 pm

Thu 05/09 SEC 06 Kaggle Setup
resources: handout

LES 13 Kernel Methods; Locality Sensitive Hashing
pre-class: lesson megathread
in-class: pdf annotated pptx
post-class: checkpoint
resources: videos extra resources

Videos

Fri 05/10

- Recording

Extra resources

- [Schafer] AnIML - Kernel Methods (coming soon)
- [Schafer] AnIML - Locality Sensitive Hashing (coming soon)

Mon 05/13

Tue 05/14

Out HW5 ML Practice on Kaggle Due 11:59 pm

Module 7 - Clustering

LES 14 Clustering
pre-class: lesson megathread
in-class: pdf annotated pptx
post-class: checkpoint
resources: videos extra resources advanced resources

Videos

- Recording

Wed 05/15

Extra resources

- k-means visualization
- [Schafer] AnIML - k-means (coming soon)
- [Nguyen] Lecture Notes from previous quarter

Advanced resources

- [ESL] Section 13.2.1, 14.3.6, 14.3.11

Out LR7 Due 11:59 pm

Thu 05/16 SEC 07 Numpy ; Variable Encoding ; Clustering
resources: handout

LES 15 Hierarchical Clustering
pre-class: lesson megathread
in-class: pdf annotated pptx
post-class: checkpoint
resources: videos extra resources

Videos

Fri 05/17

- Recording

Extra resources

- [Schafer] AnIML - Other Clustering (coming soon)
- [Nguyen] Lecture Notes from previous quarter

Out HW6 k-means with Text Data

Mon 05/20

Tue

05/21		Due 11:59 pm	
Module 8 - Recommender Systems			
Wed 05/22	LES 16 PCA / Recommender Systems Intro pre-class: lesson megathread in-class: pdf annotated pptx post-class: checkpoint resources: videos **Videos** • Recording extra_resources: PCA Visualized, [Schafer] AnIML - Dimensionality Reduction (coming soon), [Schafer] AnIML - Recommender Systems 19 - 19.4, [Nguyen] Lecture Notes from previous quarter		Out LR8 Due 11:59 pm
Thu 05/23	SEC 08 PCA ; Recommender Systems resources: handout		
Fri 05/24	LES 17 Recommender Systems / Matrix Factorization pre-class: lesson megathread in-class: pdf annotated pptx post-class: checkpoint resources: videos **Videos** • Recording extra_resources: [Schafer] AnIML - Recommender Systems, [Nguyen] Lecture Notes from previous quarter	Out HW7 Tweet Topic Modelling Due 11:59 pm	
Mon 05/27			
Tue 05/28			
Module 9 - Course Wrap Up			
Wed 05/29	LES 18 Course Wrap Up; Next Steps; Generative AI pre-class: lesson megathread in-class: pdf annotated pptx post-class: checkpoint resources: videos extra resources		Out LR9 Due 11:59 pm
Thu 05/30	SEC 09 Final Review resources: handout		
Fri 05/31			
Mon 06/03		EXAM Final Exam At 6:00 PM	