

ISTA 421 (undergraduate) and INFO 521 (graduate)

Introduction to Machine Learning

Fall 2019

Course Description:

Machine learning is the discipline that develops and studies algorithms that can adjust their parameters (i.e., *learn*) based on data (experience). This course will introduce the fundamentals of modern machine learning. We will describe how to implement several practical methods for pattern recognition (regression and classification), clustering, and dimensionality reduction, as well as methods for training and evaluating learned models.

Prerequisites:

Calculus I (derivatives) and II (integrals), Linear Algebra;
Programming experience (at least 3 courses; the language used in the course is Python 3);
Introductory statistics and probability (ISTA 311 recommended; ISTA 116 or comparable is sufficient).

Units: 3

Course Location and Times:

Harvill, Room 115 (south building)
Monday and Wednesday, 5:00 PM to 6:15 PM

Instructor Information:

Clayton T. Morrison, PhD
Associate Professor
School of Information
Office: Harvill 437A
Office Phone: 520-621-6609
Email: claytonm@email.arizona.edu
Office Hours: M/W 4-5pm, Harvill 437A. Contact by email to set up an appointment

Course Objectives:

The objective of this course is to introduce the core methods used in modern machine learning, gain experience implementing these methods as algorithms in a programming language (Python), and use and evaluate them on data.

Textbooks and Resources:

- Simon Rogers and Mark Girolami (2016). *A First Course in Machine Learning, Second Edition*. Chapman & Hall / CRC Press.
This is available for free through the UA Library; see course website for direct link.
Be sure to use the *Second Edition*.
- Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani (2013). *Introduction to Statistical Learning, with Applications in R*. Springer.
Available online for free: <http://www-bcf.usc.edu/~gareth/ISL/>

Additional readings will be provided in class and through the D2L website.

We will reference some material from the following, all available for free online:

- Hal Daumé III (2019). *A Course in Machine Learning*. Online: <http://ciml.info>
- Trevor Hastie, Robert Tibshirani and Jerome Friedman (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Second Edition).
Online: <http://www-stat.stanford.edu/~tibs/ElemStatLearn/download.html>

Other resources, recommended for additional, complementary, and more in-depth coverage, but not required for this course:

- David Barber (2012). *Bayesian Reasoning and Machine Learning*.
[Available for free online: <http://www.cs.ucl.ac.uk/staff/d.barber/brml/>]
- A. Blum, J. Hopcroft and R. Kannan (2018). *Foundations of Data Science*.
Online: <https://www.cs.cornell.edu/jeh/book%20no%20so;utions%20March%202019.pdf>
- I. Goodfellow, Y. Bengio and A. Courville (2016). *Deep Learning*. [On the foundations of modern neural networks. Available for free online:
<http://www.deeplearningbook.org/>]
- C. Rasmussen and C. Williams (2006). *Gaussian Processes for Machine Learning*.
[Available for free online: <http://www.gaussianprocess.org/gpml/>]
- Y. S. Abu-Mostafa, M. Magdon-Ismael and H.-T. Lin (2012). *Learning from Data*. AMLBook. [Very good introductory book with more of a focus on learning theory.]
- Kevin P. Murphy (2012). *Machine Learning: A Probabilistic Perspective*. The MIT Press. [Very comprehensive, a good reference; but beware of numerous errors in the first four printings – read with the errata handy (errata available from Kevin’s book site: <http://www.cs.ubc.ca/~murphyk/MLbook/>).]
- Christopher Bishop (2006). *Pattern Recognition and Machine Learning*. Springer.
[More advanced, Bayesian perspective.]

Grade policies:

Grade categories and proportions:

Grading will be based on programming exercises and written assignments (Assignments), one in-class midterm (Midterm) and a final assignment (Final Assignment). **Points** will be assigned to each according to the amount of effort expected. Graduate students will receive additional and alternate problems, and therefore need to earn more points total for many of the assignments; this reflects the expected higher level of competency and preparation of graduate students.

The total points in each category are scaled to contribute to the final **cumulative grade percentage** value according to the following distribution (the same distribution scaling is used for both undergraduates or graduates):

Assignments: 70%

(Written Homework & Programming Exercises, assigned approximately every other week.)

Midterm: 10%

Final Assignment: 20%

Letter grades are then assigned according to the final cumulative grade percentage:

A: $90\% \leq$ cumulative grade

B: $80\% \leq$ cumulative grade $< 90\%$

C: $70\% \leq$ cumulative grade $< 80\%$

D: $60\% \leq$ cumulative grade $< 70\%$

Late Work Policy:

Assignments are due by the date/time specified.

No examinations may be taken after the examination date.

Assignments will generally be submitted as PDF documents and code files that are committed and pushed to the course GitHub Classroom repository. We *may* also use the course D2L assignment turn-in for some assignments.

In case of emergencies affecting turning work in on time, you ***must*** contact the instructor *immediately*.

Classroom Behavior:

Please be considerate: Please disable your mobile device ringer or sound notifications. If you get an urgent call, please quietly leave the lecture hall to conduct it. Screens are distracting! If you want to use a laptop or tablet for taking notes, please sit toward the side and back of the class in order to not disrupt other students.

Asking Questions: During class, feel free to interrupt with questions whenever they occur to you. The instructor may ask you to hold off on your question for a few moments.

Answering Questions: We frequently ask questions of the class during lectures to judge the level of understanding. Some students really like answering questions, sometimes to the point of discouraging anyone else from answering. If you are an eager answerer, pace yourself; let someone else answer an easy one once in a while, and save the hard ones for yourself.

Course Topics

Introduction

The Linear Model

- Linear Regression and Linear Least Mean Squares

- Linear classification and the Perceptron

- Extending the Linear Model

Learning Theory and Model Evaluation

- Bias-Variance Tradeoff

- Cross Validation

Probabilistic methods

- Review of Probability

- Maximum Likelihood

- The Bayesian approach

 - Priors, Marginal Likelihood, Hyperparameters

Optimization and Approximation Methods

- Gradient methods (gradient descent, Newton's method)

- Laplace estimation

- Sampling, MCMC, Metropolis-Hastings

Classification

- Logistic Regression

- Bayesian Classification, Naive Bayes

- Nearest Neighbors

- Support Vector Machines

Neural Networks

- Perceptron, Backpropagation

- Autoencoders and "Deep" architectures

Clustering

- K-means

- Mixture Models

- Expectation Maximization

Projection Methods

- Principal Components Analysis

- Latent Variable Models

Additional Topics (depending on time)

- Nonparametric Bayesian methods

 - Gaussian Processes

 - Topic Modeling

- Ensembles, Boosting and Random Forests

UNIVERSITY POLICIES

Honors Credit:

Students wishing to contract this course for Honors Credit should email me to set up an appointment to discuss the terms of the contract and to sign the Honors Course Contract Request Form.

The Honors College website: <https://www.honors.arizona.edu/>

Missed Classes (Absence):

The UA's policy concerning Class Attendance and Administrative Drops is available at: <http://catalog.arizona.edu/policy/class-attendance-participation-and-administrative-drop>

All holidays or special events observed by organized religions will be honored for those students who show affiliation with such religions. Absences pre-approved by the UA Dean of Students office will be honored. See:

<http://policy.arizona.edu/human-resources/religious-accommodation-policy>

No matter the reason for missing class, the student is always responsible for any missed material.

Final Exams:

Final Exam Regulations and Information:

<https://www.registrar.arizona.edu/courses/final-examination-regulations-and-information> Final Exam Schedule:

<http://www.registrar.arizona.edu/students/courses/final-exams>

Audit Policy:

<http://catalog.arizona.edu/policy/audit-policy>

Accessibility and Accommodations:

It is the University's goal that learning experiences be as accessible as possible. If you anticipate or experience physical or academic barriers based on disability, please let me know immediately so that we can discuss options. You are also welcome to contact Disability Resources (520-621-3268) to establish reasonable accommodations. DRC online: <http://drc.arizona.edu/>

Please be aware that the accessible table and chairs in this room should remain available for students for whom standard classroom seating is not usable.

Disability Resource Center. 1221 E. Lowell St. Tucson, Az 85721 Tel: 520-621-3268 (TTY). Fax: 520-621-9423. uadrc@email.arizona.edu

UA SALT Center: <http://www.salt.arizona.edu>

Student Code of Academic Integrity (a.k.a., Cheating):

Students are encouraged to share intellectual views and discuss freely the principles and applications of course materials. However, graded work/exercises must be the product of independent effort unless otherwise instructed. Students are expected to adhere to the UA Code of Academic Integrity as described in the UA General Catalog. See:

<http://deanofstudents.arizona.edu/policies-and-codes/code-academic-integrity>

The University Libraries have some excellent tips for avoiding plagiarism available at: <http://new.library.arizona.edu/research/citing/plagiarism>

*Selling class notes and/or other course materials to other students or to a third party for resale is **not permitted** without the instructor's express written consent.* Violations to this and other course rules are subject to the Code of Academic Integrity and may result in course sanctions. Additionally, students who use D2L or UA email to sell or buy these copyrighted materials are subject to Code of Conduct Violations for misuse of student email addresses. This conduct may also constitute copyright infringement.

Additional Information on conduct:

The Arizona Board of Regents list of Prohibited Conduct (pdf):
<https://public.azregents.edu/Policy%20Manual/5-303-Prohibited%20Conduct.pdf>

The Arizona Board of Regents Student Code of Conduct (pdf):
<https://public.azregents.edu/Policy%20Manual/5-308-Student%20Code%20of%20Conduct.pdf>

Policies Against Threatening Behavior:

The Student Code of Conduct (5-308.F.11) dictates that no person or organization may interfere with or threaten University-sponsored classroom activities. The following links provides details about the policy:

<http://policy.arizona.edu/education-and-student-affairs/threatening-behavior-students>

<http://policy.arizona.edu/education-and-student-affairs/disruptive-behavior-instructional-setting>

Non-Discrimination and Anti-Harassment Policy:

<http://policy.arizona.edu/human-resources/nondiscrimination-and-anti-harassment-policy>

Statements of Inclusion:

Elective name and pronoun usage: This course supports elective gender pronoun use and self-identification; rosters indicating such choices will be updated throughout

the semester, upon student request. As the course includes in-class discussion, it is vitally important for us to create an educational environment of inclusion and mutual respect.

Inclusive Excellence is a fundamental part of the University of Arizona's strategic plan and culture. As part of this initiative, the institution embraces and practices diversity and inclusiveness. These values are expected, respected and welcomed in this course.

Office of Diversity: <http://diversity.arizona.edu/>

Health: <https://www.health.arizona.edu/>

Counseling: <http://www.health.arizona.edu/counseling-and-psych-services>

Additional Student Resources:

UA Academic policies and procedures are available at:

<http://policy.arizona.edu/>

UA Non-discrimination and Anti-harassment policy:

<http://policy.arizona.edu/human-resources/nondiscrimination-and-anti-harassment-policy>

Student Assistance and Advisory information is available at:

<http://deanofstudents.arizona.edu/student-assistance/students/student-assistance>

Confidentiality of Student Records:

<http://registrar.arizona.edu/personal-information/family-educational-rights-and-privacy-act-1974-ferpa>

Required or Special Materials: None.

Notice of Potentially Offensive Material: None.

Miscellaneous University Policies:

(1) On Dropping Classes:

If you find yourself thinking about dropping this (or any other) class, first make sure that that's what you really want to do. Chatting with the instructor or your academic advisor may help.

Undergraduate Administrative Drop Deadlines:

Last Day to Admin Drop without a W: 9/8/2019

Last Day to Admin withdraw: 11/3/2019

Graduate Administrative Drop Deadlines:

Last Day to Admin Drop without a W: 9/22/2019

Last Day to Admin withdraw: 11/3/2019

University of Arizona Class Attendance and Administrative Drop Policy:

<http://catalog.arizona.edu/policy/class-attendance-participation-and-administrative-drop>

(2) Grades of 'Incomplete':

Office of Curriculum and Registration Grading Policy Manual:

<https://www.registrar.arizona.edu/grades/incomplete-i-grade>

UA General Catalog's Grades and the Grading System:

<http://catalog.arizona.edu/policy/grades-and-grading-system>

The university's course catalog contains all of the details about incompletes, but here's the key sentence: "The grade of I may be awarded only at the end of a term, when all but a minor portion of the course work has been satisfactorily completed."

To qualify for an incomplete, a student must have maintained a passing grade for the class until the term is nearly complete, and then, due to an unusual and substantiated cause beyond the student's control, the student is unable to complete the class work. In short, you *cannot* get an "I" just because you aren't happy with your grade.

“Subject to Change” Statement:

The instructor reserves the right to change with advance notice where appropriate the content of the course. This right does not apply to the grade, absence or University policies.