CS-UY 4563: Introduction to Machine Learning

Overview

This course provides a hands on approach to machine learning and statistical pattern recognition. The course describes fundamental algorithms for linear regression, classification, model selection, support vector machines, neural networks, dimensionality reduction and clustering. The course includes computer exercises on real and synthetic data using current software tools. A number of applications are demonstrated. Students should have competency in computer programming

- Prof: Linda Sellie
    - Zoom office hours: Thursdays 2:00 - 4:00, and by appointment. The zoom link is https://nyu.zoom.us/j/95664494993
- TAs: TBA
    - office hours: TBA
- Texts:
    - *Machine Learning Refined: Foundations, Algorithms, and Applications* 2nd edition by Jeramy Watt, Reza Borhani, and Aggelos Katsaggelos.
    - *http://cs229.stanford.edu/syllabus.html*
    - *Hands-On Machine Learning with Scikit-Learn and TensorFlow* 2nd edition by Aurlien Gron. On line resources https://github.com/ageron/handson-ml2
- Resources:
    - More resources will be provided during the semester
- Supplementary texts and resources
    - Mathematical Background: https://mml-book.com, or https://davetang.org/file/Singular_Value_Decomposition_Tutorial.pdf, or the appendix of our textbook *Machine Learning Refined: Foundations, Algorithms, and Applications* has a nice review in the appendix
    - Installing python (you will need to do this before the first assignment): http://jupyter-notebook-beginner-guide.readthedocs.io/en/latest/index.html
    - Python tutorial: https://docs.python.org/3/tutorial/
- Grading:
    - Midterm 1: 25%, Midterm 2: 25%; Final project: 20%; class participation, homework & quiz: 30%,
    - Programming assignments (Labs) will involve python-based exercises (most of the labs were modified versions of Prof. Sundeep Rangan's labs).
    - Final project is done in groups of two.
    - Midterm exams and quiz are closed book.

- Pre-requisites:
    - One of: MA-UY 2224 (Data Analysis); MA-UY 2222 (Data Analysis 2) or EE-UY 2223 (Probability) or equivalent.
    - Undergraduate probability and linear algebra (MA-UY 2034, MA-UY 2034G, MA-UY 3044 or MA-UY 3054)
    - Programming experience is essential. CS-UY 1134 (Data Structures and Algorithms)

Tentative Outline

- Introduction
    - Course logistics. Examples of machine learning problems used today. Formulate machine learning problems (identify task, data, objectives). Classify ML problems as supervised vs. unsupervised, regression vs. classification.
    - Read: Chapter 1 in *Machine Learning Refined: Foundations, Algorithms, and Applications*
    - Additional resources: Chapter 1 in *Hands-On Machine Learning with Scikit-Learn and TensorFlow*
- Simple regression and multiple Linear regression
    - Least squares formula, Gradient Descent, Normal Equations Method
    - Read: Sections 1 - 3 in http://cs229.stanford.edu/notes2020spring/cs229-notes1.pdf (Note that we will first learn simple regression and then learn multiple linear regression)
    - Intuition: Chapter 2.1-2.5 & 2.7 in *Machine Learning Refined: Foundations, Algorithms, and Applications*
    - Background:

      Appendix *B1-B4, B9-B9.1* in *Machine Learning Refined: Foundations, Algorithms, and Applications*

      https://betterexplained.com/articles/vector-calculus-understanding-the-gradient/

      https://en.wikipedia.org/wiki/Normal_distribution
    - Additional resources:

      Chapter 5.1-5.3 in *Machine Learning Refined: Foundations, Algorithms, and Applications*

      Pages 111-123 and page 128-130 in *Hands-On Machine Learning with Scikit-Learn and TensorFlow*
- Model selection and regularization: Identify the order in a multiple linear regression model
    - Understanding underfitting and overfitting with polynomials; irreducible error; bias and variance tradeoff; cross validation; regularization techniques
- Logistic Regression
    - Gradient descent
- Support vector machines (SVMs)

    - Maximum margin; duality; kernel methods
- Neural networks
    - Formulation; back propagation
- Convolutional and deep networks
    - Convolutional layers; pooling layers
- PCA
    - Dimensionality reduction
- Clustering and K-Means
    - K-means; mixture models; EM methods
- TBA
- Final project presentations

Moses Center Statement of Disability:

If you are student with a disability who is requesting accommodations, please con- tact New York University's Moses Center for Students with Disabilities (CSD) at 212-998-4980 or mosescsd@nyu.edu. You must be registered with CSD to receive accommodations.
Information about CSD can be found at www.nyu.edu/csd. It is located at 726 Broadway on the 2nd floor.

NYU School of Engineering Policies and Procedures on Academic Misconduct

Introduction:

The School of Engineering encourages academic excellence in an environment that promotes honesty, integrity, and fairness, and students at the School of Engineering are expected to exhibit those qualities in their academic work. It is through the process of submitting their own work and receiving honest feedback on that work that students may progress academically. Any act of academic dishonesty is seen as an attack upon the School and will not be tolerated. Furthermore, those who breach the School's rules on academic integrity will be sanctioned under this Policy. Students are responsible for familiarizing themselves with the School's Policy on Academic Misconduct.

Definition:

Academic dishonesty may include misrepresentation, deception, dishonesty, or any act of falsification committed by a student to influence a grade or other academic evaluation. Academic dishonesty also includes intentionally damaging the academic work of others or assisting other students in acts of

dishonesty. Common examples of academically dishonest behavior include, but are not limited to, the following:

1. Cheating: intentionally using or attempting to use unauthorized notes, books, electronic media, or electronic communications in an exam; talking with fellow students or looking at another person's work during an exam; submitting work prepared in advance for an in-class examination; having someone take an exam for you or taking an exam for someone else; violating other rules governing the administration of examinations.

2. Fabrication: including but not limited to, falsifying experimental data and/or citations.

3. Plagiarism: intentionally or knowingly representing the words or ideas of another as one's own in any academic exercise; failure to attribute direct quotations, paraphrases, or borrowed facts or information.
4. Unauthorized collaboration: working together on work that was meant to be done individually.

5. Duplicating work: presenting for grading the same work for more than one project or in more than one class, unless express and prior permission has been received from the course instructor(s) or research adviser involved.

6. Forgery: altering any academic document, including, but not limited to, academic records, admissions materials, or medical excuses.

See http://engineering.nyu.edu/files/SACCofC9-14-15.pdf for the Code of Conduct, which can be found here: http://engineering.nyu.edu/life/student-affairs/ community- standards- procedures

NYU School of Engineering Policies and Procedures on Excused Absences

– complete policy here \url{https://engineering.nyu.edu/campus-and-community/student-life/office-student-affairs/policies#chapter-id-30199

1. Introduction: An absence can be excused if you have missed no more than 10 days of school. If an illness or special circumstance has caused you to miss more than two weeks of school, please refer to the section labeled Medical Leave of Absence.

2. Students may request special accommodations for an absence to be excused in the following cases:
    - Medical reasons
    - Death in immediate family
    - Personal qualified emergencies (documentation must be provided)
    - Religious Expression or Practice

Deanna Rayment, (deanna.rayment@nyu.edu), is the Coordinator of Student Advocacy, Compliance and Student Affairs and handles excused absences. She is located in 5 MTC, LC240C and can assist you should it become necessary.

NYU School of Engineering Academic Calendar -- complete list https://www.nyu.edu/registrar/calendars/university-academic-calendar.html#1198

The last day of the final exam period is Dec. 21st. Final exam dates for undergraduate courses will not be determined until later in the semester. Final exams for graduate courses will be held on the last day of class during December 15-21. If you have two final exams at the same time, report the conflict to your professors as soon as possible. Do not make any travel plans until the exam schedule is finalized.

Also, please pay attention to notable dates such as Add/Drop, Withdrawal, etc. For confirmation of dates or further information, please contact Susana: sgarcia@nyu.edu