MIS 385N, Advanced Data Mining / Web Analytics

Spring 2020, TTH 11:00 AM - 12:30 PM, Room: GSB 3.130

Instructor: Dr. Mingyuan Zhou, Associate Professor

- Office: CBA 6.458 (east side of the building that faces the entrance of Gregory Gym)
- Phone: 512-232-6763
- Email: `mingyuan.zhou@mccombs.utexas.edu`
- Website: http://mingyuanzhou.github.io/
- Office Hours: Tuesday 2:00-3:00 PM
- Teaching Assistant: Yuguang Yue <yuguang@utexas.edu>
  - 3rd year PhD student @ Statistics and Data Sciences
  - TA Office hours: Monday and Wednesday, 3 to 4 pm
  - Location: CBA 4.304 (TA space G)

Course Description:

This graduate course will introduce core Machine Learning and Data Mining (MLDM) algorithms and demonstrate how they can be used in a wide variety of real-world applications. We will learn various MLDM algorithms from both the viewpoint of optimization and that of probabilistic modeling. Example topics include classification, discrete choice analysis, regression, recommendation systems, matrix factorization, manifold learning, (social) network analysis, information retrieval, sequential decision making, and topic modeling. Several recent progress in MLDM (e.g., deep neural networks, deep reinforcement learning, deep generative models) and their example applications will also be discussed.

Materials:

- Recommended but not required:

  [1] Pattern Recognition and Machine Learning, by Christopher Bishop.

  [2] Machine Learning: a Probabilistic Perspective, by Kevin Murphy.
  Python demo code: https://github.com/probml/pyprobml

  [3] Deep Learning, by Ian Goodfellow, Yoshua Bengio, and Aaron Courville.

  [4] The Elements of Statistical Learning, by Trevor Hastie, Robert Tibshirani, and Jerome Friedman.

  [5] Hands-on Machine Learning with Scikit-Learn & TensorFlow, by Aurélien Géron, https://github.com/ageron/handson-ml

[6] Dive Into Deep Learning, by Aston Zhang, Zachary C. Lipton, Mu Li, and Alexander J. Smola.

[7] scikit-learn: Machine Learning in Python (https://scikit-learn.org/stable/)

- Recommended software:

  Python, R, or Matlab

Grading:

Homework (50%)
Midterm Exam (25%)
Final Project (25%)

Tentative Course Schedule:

This schedule represents my current plans and objectives. As we go through the semester, those plans may need to change to enhance the class learning opportunity. Such changes, communicated clearly, are not unusual and should be expected.

Weeks 1-3 (January 21, 23, 28, 30, February 4, 6)
Introduction to Machine Learning, Supervised Learning
Weeks 4-5 (February 11, 13, 18, 20)
Unsupervised Learning
Week 6-7 (February 25, 27, March 3, 5)
Recommendation systems, collaboration filtering, matrix factorization
Week 8-9 (March 10, 12, 24, 26)
Manifold learning, word embeddings, social network analysis
Week 10-11 (March 31, April 2, 7, 9)
Information retrieval, text analysis, topic model
Week 12-15 (April 14, 21, 23, 28, 30, May 5, 7)
Introduction to deep learning, deep reinforcement learning, and deep generative models
April 16
Midterm Exam

Quantitative Reasoning Flag: This course carries the Quantitative Reasoning flag. Quantitative Reasoning courses are designed to equip you with skills that are necessary for understanding the types of quantitative arguments you will regularly encounter in your adult and professional life. You should therefore expect a substantial portion of your grade to come from your use of quantitative skills to analyze real-world problems.

University of Texas Honor Code: The core values of The University of Texas at Austin are learning, discovery, freedom, leadership, individual opportunity, and responsibility. Each member of the university is expected to uphold these values through integrity, honesty, trust, fairness, and respect toward peers and community.

Academic Integrity: Each student in this course is expected to abide by the University of Texas Honor Code. Any work submitted by a student in this course for academic credit will be the student's own work.

You are encouraged to study together and to discuss information and concepts covered in lecture and the sections with other students. You can give "consulting" help to or receive "consulting" help from such students. However, this permissible cooperation should never involve one student having possession of a copy of all or part of work done by someone else, in the form of an e-mail, an e-mail attachment file, a diskette, or a hard copy.

Should copying occur, both the student who copied work from another student and the student who gave material to be copied will both automatically receive a zero for the assignment. Penalty for violation of this Code can also be extended to include failure of the course and University disciplinary action.

During examinations, you must do your own work. Talking or discussion is not permitted during the examinations, nor may you compare papers, copy from others, or collaborate in any way. Any collaborative behavior during the examinations will result in failure of the exam, and may lead to failure of the course and University disciplinary action.

Students with Disabilities: Students with disabilities may request appropriate academic accommodations from the Division of Diversity and Community Engagement, Services for Students with Disabilities, 512-471-6259, http://www.utexas.edu/diversity/ddce/ssd/.

Religious Holy Days: By UT Austin policy, you must notify me of your pending absence at least fourteen days prior to the date of observance of a religious holy day. If you must miss a class, an examination, a work assignment, or a project in order to observe a religious holy day, you will be given an opportunity to complete the missed work within a reasonable time after the absence.

Policy on Academic Dishonesty: I have no tolerance for acts of academic dishonesty. Such acts damage the reputation of the school and the degree and demean the honest efforts of the majority of students. The minimum penalty for an act of academic dishonesty will be a zero for that assignment or exam.

The responsibilities for both students and faculty with regard to the Honor System are described on on the final pages of this syllabus. As the instructor for this course, I agree to observe all the faculty responsibilities described therein. As a Texas MSITM student, you agree to observe all of the student responsibilities of the Honor Code. If the application of the Honor System to this class and its assignments is unclear in any way, it is your responsibility to ask me for clarification.

Campus Safety: Please note the following recommendations regarding emergency evacuation, provided by the Office of Campus Safety and Security, 512-471-5767, http://www.utexas.edu/safety:

- Occupants of buildings on The University of Texas at Austin campus are required to evacuate buildings when a fire alarm is activated. Alarm activation or announcement

requires exiting and assembling outside.

- Familiarize yourself with all exit doors of each classroom and building you may occupy. Remember that the nearest exit door may not be the one you used when entering the building.
- Students requiring assistance in evacuation should inform the instructor in writing during the first week of class.
- In the event of an evacuation, follow the instruction of faculty or class instructors.
- Do not re-enter a building unless given instructions by the following: Austin Fire Department, The University of Texas at Austin Police Department, or Fire Prevention Services office.
- Behavior Concerns Advice Line (BCAL): 512-232-5050
- Further information regarding emergency evacuation routes and emergency procedures can be found at: http://www.utexas.edu/emergency.