

INF 552: Machine Learning for Data Informatics

Units : 4
Location : ZHS 163

Instructor : Ion Muslea
Office Hours : by appointment
Contact Info : imuslea.inf.552@gmail.com

Teaching Assistant: TBA

Course Description and Learning Objectives

The learning objectives for students in this course are:

- acquire a *hands-on understanding of applying machine learning in the real world*
- broadly understand major algorithms used in Machine Learning
- understand supervised and unsupervised learning techniques
- understand Bayesian decision theory and nonparametric methods
- understand decision trees, dimensionality reduction, clustering, and kernel machines
- understand Bayesian estimation, hidden Markov models, and graphical models

This is a foundational course with primary application in data analytics. It is intended to be accessible to students with technical backgrounds as well as to students with less technical backgrounds. The textbook for INF-552 is “Introduction to Machine Learning”, 3rd edition, MIT Press, 2015, by Ethem Alpaydin.

Recommended Preparation

A basic mathematical background in probability, statistics, and linear algebra, as well as basic programming skills and a basic understanding of engineering principles are strongly encouraged.

Description and Assessment of Assignments

Grading will be based on students’ understanding of lecture material, the thoroughness of their projects, and their ability to explore related areas. Students can work in groups of 2-3, but should mention their individual contributions.

Grading Breakdown

- Project: 70%
- Quizzes: 10% (scope: reading materials for during **last 2 weeks**)
- Final exam: 20% (**December 6**, 4:30-6:30 pm)

Course Project

Course project: the purpose of the class project is for you to acquire the hands-on experience of identifying a data problem and applying machine learning approaches to solve it. Students are encouraged to identify unique applications for machine learning and develop novel approaches.

Working as a group is permitted if the project is large enough to justify it; a team can consist of up to 3 persons.

Project Timeline (changes might apply):

- Week 1-4: Identifying team members and project topics
- Week 5: **9/20** Draft proposal due (team member, topics and milestones)
- Week 7: **10/4** Proposal due
- Week 11: **11/1** Mid-term report (data description, preliminary results)
- Week 14: **11/22** Project presentation
- Week 15: **11/29** Project presentation
- **Week 15: 11/29** Final report (task & model description, major discovery, lessons learned)

Sample projects: google “*ideas machine learning class projects*”

- <http://www.hlt.utdallas.edu/~vgogate/ml/2012s/projects.html>
 - <http://cs229.stanford.edu/projects2011.html>
 - <http://www.cs.cmu.edu/~guyton/Class/10701/projects.html>
 - 4 more links from original page
- <https://elitedatascience.com/machine-learning-projects-for-beginners>
- <http://web.cs.hacettepe.edu.tr/~ilyas/Courses/BIL712/possibleprojecttopics.pdf>
- ...

Bootstrap your project work by using the tools you are most proficient with, such as:

- Java & Weka
- Python & ML libraries
- Matlab & ML libraries
- C++ & MLC++
- *etc*

Grading breakdown of the course project:

- Proposal: 10%
- Mid-term report: 25%
- Final report: 35%
- Presentation: 30%

Assignment Submission Policy

Students are required to submit the assignments on USC’s BlackBoard portal. One submission per team is sufficient.

Grading Timeline

Feedback and grades will be available within 1-2 weeks after assignment due dates.

Course Schedule

TBA

Academic Conduct

Plagiarism—presenting someone else’s ideas as your own, either verbatim or recast in your own words—is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in SCampus in Part B, Section 11, “Behavior Violating University Standards” <https://policy.usc.edu/student/scampus/part-b>. Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on scientific misconduct <http://policy.usc.edu/scientific-misconduct>.

Support Systems

A number of USC’s schools provide support for students who need help with scholarly writing. Check with your advisor or program staff to find out more. Students whose primary language is not English should check with the American Language Institute <http://ali.usc.edu>, which sponsors courses and workshops specifically for international graduate students. The Office of Disability Services and Programs <http://dsp.usc.edu> provides certification for students with disabilities and helps arrange the relevant accommodations. If an officially declared emergency makes travel to campus infeasible, USC Emergency Information <http://emergency.usc.edu> will provide safety and other updates, including ways in which instruction will be continued by means of BlackBoard, teleconferencing, and other technology.