

Computer Science 4033/5033: Machine Learning

Instructor: Dr. McGovern

Fall 2010

1 Course Overview

By taking this course, you will learn about current techniques that enable machines to learn and adapt their behavior over time and to new situations. You will take a small step into creating intelligent machines for the future by choosing and completing your own semester-long machine learning project. The human beings will never be taken out of the loop entirely (unless we create true AI!) and you will also gain experience at teamwork and at presenting your work professionally through the project and homework.

2 Learning Objectives

The general/overall learning objectives are:

- Understand the basic techniques in ML
- Select the ML solution best-suited for a novel domain and justify your choice
- Synthesize one area of machine learning in depth and apply it to a novel application
- Implement and evaluate the effectiveness of your ML method applied to a novel application
- Communicate ideas clearly to a variety of audiences both in oral and written form
- Function effectively in a team

The specific topics we will cover include:

Supervised learning

- Regression

- Neural nets
- Kernel methods and Support Vector Machines

Reinforcement learning

- The RL problem
- Dynamic Programming
- Temporal Difference Learning
- Function Approximation

General techniques/concepts

- Ensemble methods
- Overfitting
- Expectation Maximization (EM)

Graphical models

- Bayesian Networks: using them, inference, building/creating them, learning them automatically both from observed and unobserved evidence

The specific objectives for these detailed topics are:

- Describe the differences between supervised learning, unsupervised learning, and semi-supervised learning
- Build a regression model for linear and non-linear data
- Derive appropriate backup equation for a neural net given its structure and hidden units
- Implement and evaluate neural nets
- Explain why the kernel trick makes support vector machines solvable
- Apply kernel methods to classification task using a model other than SVMs
- Describe the basic premise behind reinforcement learning and why it works
- Design an RL agent for a novel non-gridworld task and explain why you think your design is appropriate
- Apply techniques for function approximation (such as neural networks) to real-world RL agent design
- Explain why many weak predictors used in an ensemble fashion can create a strong predictor
- Apply appropriate ensemble methods to novel problems

- Recognize when a technique is overfitting to a problem
- Design and apply appropriate correction techniques to minimize overfitting
- Apply EM to mixture-model task
- Derive EM solution for a problem where the underlying learning method is not one we discussed in class (e.g. not Bayes networks or mixture models).
- Explain why EM works to someone outside CS
- Describe basic rules of probability that form the foundation of Bayesian Networks
- Build a compact Bayesian Network for a problem given a description of all the factors and influences involved
- Correctly use a Bayesian Network for inference (answering a query given some evidence)
- Explain how a Bayesian Network can learn conditional probability distributions for unobserved data

3 General Information

Class time: Tues/Thursday 12 noon - 1:15pm

Irregular exam: Poster session Dec 10 time TBA (probably 3-5pm)

Class location: Carson 123

Prerequisites: MATH 4753 or ENGR 3293 or IE 3293 or MATH 4743 or permission of the instructor. Prior programming experience is assumed.

Required materials: 1) The textbook will be the 3rd edition of *Artificial Intelligence: A Modern Approach* by Russell and Norvig. 2) *Reinforcement Learning* by Sutton and Barto. This book is available for free online at:
<http://www.cs.ualberta.ca/~sutton/book/the-book.html>
or you can purchase it through MIT Press.

Instructor: Dr. McGovern

- *Office:* Devon 251
- *Phone:* 325-5427 (email preferred over voicemail)
- *URLs for class:*
<http://learn.ou.edu>
http://www.cs.ou.edu/~amy/courses/cs5033_fall2010/index.html
- *Email:* amcgovern@ou.edu

- *Office hours:* Tuesday 1:30-3, Thursday 10-11:30. Also by appointment. Additional appointments for office hours are possible and you just need to email me. Also available irregularly via AIM at *dramymcgovern*. Please note open door policy on my door and stop by if the door is open.

4 Evaluation

You will be learning and practicing many aspects of machine learning. What you get out of a course will depend on what you put into it! In order to give you a fair grade at the end of the semester, I will evaluate you on a combination of your project (60%) and homework (40%). Participating in class is one of the best ways to learn so please ask questions and attend class.

Undergraduates: Students taking 4033 will shorter homework assignments and a smaller project. The grade cutoff lines are not required to be the same for the two courses.

Grade questions: To maintain fairness in grading, please bring any disagreements to me within a week of the item being returned.

Online Grade Summary: Desire2Learn has a grade book that I will use to store all of your grades. It is your responsibility to verify that the grades on D2L are correct. If an error is found, bring the document to me and I will correct Desire2Learn.

Borderline grades: Borderline final grades will be decided by your class participation which means that being an active participant in class can push you over a grade boundary.

Final Examination: Because this class contains a semester project which is presented at the poster session, there will be no final examination.

Due dates: To be fair to everyone and to minimize disruption to class, homeworks and projects are due at the beginning of class, 12 noon on the day listed in the schedule. Assignments will drop 10% of the grade per day that it is late with no assignment being accepted beyond 72 hours after the original due date. In addition, you have one “slack” day to spend however you choose during the semester. This will entitle you to turn in **one** assignment *or* project checkpoint up to one day late with no penalty. Keep in mind that you only get one of these so use it wisely.

Projects: Your final project will be due the last week of classes. Per university policy, you may turn this project in prior to pre-finals week if you have completed the project.

5 Course Policies

The following set of rules will help keep us all on the same page all semester and help to ensure fair treatment for all students.

Academic Misconduct: Academic misconduct hurts everyone but particularly the student who does not learn the material. All work submitted for an individual grade should be the work of that single individual and not his/her friends. It is fine to ask a fellow student for help as long as that help does not consist of copying any computer code, or solutions to other assignments. Students working on joint projects may certainly help one another and are expected to share code within the project group. However, they may not share beyond the group.

1. Collaboration is encouraged for homework and projects. For the projects, you will work within your groups. For the homework, you may form study groups so long as each homework is in your own words. Write your study partners' names on your homework when you turn it in.
2. Do not show another student (or group) a copy of your projects or homework before the submission deadline. The penalties for permitting your work to be copied are the same as the penalties for copying someone else's work.
3. Make sure that your computer account is properly protected. Use a good password, and do not give your friends access to your account or your computer system. Do not leave printouts or thumb drives around a laboratory where others might access them.

Upon the first documented occurrence of academic misconduct, I will report it to the Campus Judicial Coordinator. The procedure to be followed is documented in the University of Oklahoma Academic Misconduct Code¹. In the unlikely event that I elect to admonish the student, the appeals process is described in <http://www.ou.edu/provost/integrity-rights/>.

Project code: Your project code and writeups must be written exclusively by you or your group. **Use of any downloaded code or code taken from a book (whether documented or undocumented) is considered academic misconduct and will be treated as such.** Exceptions from this policy (such as a project that builds on an existing open-source project) may be granted but you **MUST** speak with me first.

Classroom Conduct: Disruptions of class will not be permitted. Examples of disruptive behavior include:

¹<http://www.ou.edu/studentcode>

- Allowing a cell phone or pager to repeatedly beep audibly.
- Playing music or computer games during class in such a way that they are visible or audible to other class members.
- Exhibiting erratic or irrational behavior.
- Behavior that distracts the class from the subject matter or discussion.
- Making physical or verbal threats to a faculty member, teaching assistant, or class member.
- Refusal to comply with faculty direction.

In the case of disruptive behavior, I may ask that you leave the classroom and may charge you with a violation of the Student Code of Responsibilities and Conduct.

Class Web Page: Login to the Desire2Learn website using your 4+4 (first four letters of your last name followed by the last four digits of your student number), using your standard OU password. If you have difficulty logging in, call 325-HELP. This software provides a number of useful features, including a list of assignments and announcements, an electronic mailing list, newsgroups, and grade book. All handouts are available from Desire2Learn. You should check the site daily. When I update the site, I will post an announcement telling you what has been added and where it is located. You are responsible for things posted on the site with a 24 hour delay.

Class evaluations: The College of Engineering utilizes student ratings as one of the bases for evaluating the teaching effectiveness of each of its faculty members. The results of these forms are important data used in the process of awarding tenure, making promotions, and giving salary increases. In addition, the faculty uses these forms to improve their own teaching effectiveness. The original request for the use of these forms came from students, and it is students who eventually benefit most from their use. Please take this task seriously and respond as honestly and precisely as possible, both to the machine-scored items and to the open-ended questions

Class Email Alias: Urgent announcements will be sent through email. It is your responsibility to:

- Have your university supplied email account properly forwarded to the location where you read email.
- Make sure that your email address in Desire2Learn is correct, and forwards email to the place where you read it. I'll send out a test message during the first week

of class. If you do not receive this message, it is your responsibility to get the problem resolved immediately.

- Have your email program set up properly so that replying to your email will work correctly the first time. You can send email to yourself and reply to yourself to test this.

If you need assistance in accomplishing any of these tasks, contact 325-HELP.

Newsgroups and Email: The newsgroup on Desire2Learn should be the primary method of communication, outside of class. This allows everyone in the class to benefit from the answer to your question. If you email me a question of general interest, I may post your question and my answer to the newsgroup. Matters of personal interest should be directed to email instead of to the newsgroup, e.g. informing me of an extended personal illness. Posting guidelines for the newsgroup are available on Desire2Learn.

Religious Holidays: It is the policy of the University to excuse the absences of students that result from religious observances and to provide without penalty for the rescheduling of examinations and additional required classwork that may fall on religious holidays.

Incompletes: The grade of I is intended for the rare circumstance when a student who has been successful in a class has an unexpected event occur shortly before the end of the class. I will not consider giving a student a grade of I unless the following three conditions have been met. 1. It is within two weeks of the end of the semester. 2. The student has a grade of C or better in the class. 3. The reason that the student cannot complete the class is properly documented and compelling.

Accommodation of Disabilities: The University of Oklahoma is committed to providing reasonable accommodation for all students with disabilities. Students with disabilities who require accommodations in this course are requested to speak with the professor as early in the semester as possible. Students with disabilities must be registered with the Office of Disability Services prior to receiving accommodations in this course. The Office of Disability Services is located in Goddard Health Center, Suite 166, phone 405/325-3852 or TDD only 405/325-4173.