CSCI-567: Machine Learning

Fall 2019

Course Description:

This course provides students with an in-depth introduction to the theory and practical algorithms for machine learning from a variety of perspectives. It covers some of the main models and algorithms for regression, classification, clustering and Markov decision processes. Topics includes linear and logistic regression, regularization, probabilistic (Bayesian) inference, SVMs and kernel methods, ANNs, clustering, and dimensionality reduction. The course uses the Python programming language and assumes in addition familiarity with linear algebra, probability theory, and multivariate calculus. This course is designed to give graduate-level students a thorough grounding in the methodologies, technologies, mathematics and algorithms currently needed by people who apply machine learning to a whole host of applications.

Learning Objectives:

- Understanding a wide variety of learning algorithms.
- Develop skills to apply learning algorithms to solving practical problems.
- Understanding how to perform evaluation of learning algorithms and model selection.
- Implement in code common ML algorithms (as assessed by the homeworks).

Optional Textbooks:

Machine Learning: A Probabilistic Perspective, by Kevin Murphy

Machine Learning, by Tom Mitchell

Prerequisites:

Students in the class are expected to have a reasonable degree of mathematical sophistication, and to be familiar with the basic knowledge of linear/matrix algebra, multivariate calculus, probability and statistics. Undergraduate classes in these subjects should be sufficient. Students are also expected to have knowledge of basic algorithm design techniques (greedy, dynamic programming, randomized algorithms, linear programming, approximation algorithms) and basic data structures. Programming in Python is required.

Review Materials:

Algorithms: https://store.cognella.com/82372-1b-002

Linear Algebra: http://viterbi-web.usc.edu/~adamchik/567/review-linalg.pdf

Probability: http://viterbi-web.usc.edu/~adamchik/567/review-prob.pdf

Python Tutorial: http://cs231n.github.io/python-numpy-tutorial/

Theory Assignments:

- There will be four written theory assignments.
- The assignments should be submitted electronically via Desire2learn.
- Theory assignments must be neatly written or typed.
- You may work in groups of 2-3. However, each person should hand-in their own writeup.
- Collaboration should be limited to talking about the problems, so that your writeup is written entirely by you and not copied from your partner.
- There are NO late days for assignments.

Programming Assignments:

- There will be four (or five) programming assignments.
- Programming assignments should be submitted electronically to Vocareum.
- They are auto-graded.
- You may submit your implementation several times. The limit is set differently for each assignment.
- Collaboration should be limited to talking about the problems.
- Each assignment will be checked for code plagiarism.
- There are NO late days for assignments.

Exams:

- There will be two midterm exams.
- No makeup exams will be provided.
- If you skip the exam, you may be eligible for an IN grade for the course. The incomplete grade has to be completed within one year. However, in order to get an IN you have to have a valid cause. Please read the University policy on IN grade for more details.
- The exam solutions and grading rubric will always be posted.

- There will be a regrading session for each exam where you can discuss grading errors. A regrade is allowed only when there are clear and obvious grading errors. Grading errors are simple mistakes made on the part of the graders, and not differences in interpretation of a question or answer.

Grading:

Final grades for the course will be determined by a curve. We will compute the letter grade cutoffs by setting the mean score to be equal a B.

Artifact	Weight	Date
Theory assignments	**16%**	
Programming assignments	**20%**	
First midterm exam	**32%**	Oct. 15
Second midterm exam	**32%**	Dec. 6 and 12

Academic Integrity:

The USC Student Conduct Code prohibits plagiarism. All USC students are responsible for reading and following the Student Conduct Code, which appears on https://policy.usc.edu/files/2018/07/SCampus-2018-19.pdf.

In this course we encourage students to study together. This includes discussing general strategies to be used on individual assignments. However, all work submitted for the class is to be done individually. Some examples of what is not allowed by the conduct code: copying all or part of someone else's work (by hand or by looking at others' files, either secretly or if shown), and submitting it as your own; giving another student in the class a copy of your assignment solution; consulting with another student during an exam. If you have questions about what is allowed, please discuss it with the instructor.

Schedule:

This schedule is meant as an outline. Depending on progress, material may be added or removed. Each lecture (and exam) is 2hrs and 20 mins long.

Date	Topics Covered
Aug. 26 - 30	Lecture 1: Course Overview, kNN, Cross-validation, Leave-one-out
Sep. 2 - 6	Lecture 2: Decision Tree, Naive Bayes, Entropy and Gini impurity, Reduced-Error Pruning
Sep. 9 - 13	Lecture 3: Linear Regression, Residual Sum of Squares, Nonlinear basis, Regularization
Sep. 16 - 20	Lecture 4: Perceptron, Logistic Regression, Gradient Descent, Surrogate Losses, Multiclass Classification
Sep. 23 - 27	Lecture 5: Neural Networks, Backpropagation, Preventing overfitting, CNN
Oct. 1 - 4	Lecture 6: Convolutional Neural Networks, Kernels, Mercer theorem
Oct. 7 - 11	Review for exam (Tue/Thu)
Oct. 14 - 18	Exam – I (Tue, Oct 15, 5-7:20pm). No lecture on Thu
Oct. 21 - 25	Lecture 7: Support Vector Machines, Linear Programming, Lagrangian Duality, KKT conditions, Dual SVM
Oct. 28 - 31	Lecture 8: Dimensionality Reduction, Principal Component Analysis, Boosting, AdaBoost, Bagging, Random Forest
Nov. 4 - 8	Lecture 9: K-means clustering, Gaussian Mixture Models
Nov. 11 - 15	Lecture 10: EM algorithm, Kernel Density Estimation
Nov. 18 - 22	Lecture 11: Hidden Markov Models, Viterbi algorithm, Baum-Welch algorithm
Nov. 25 - 29	Lecture 12: Reinforcement Learning, Multi-Armed Bandits, Markov Decision Processes
Dec. 2 - 6	Review for exam (Tue/Thu), Exam – II (Fri, Dec 6, 5-7:20pm), ALT (Dec, 12 4:30-6:30pm)

Office Hours:

Mon	Tue	Wed	Thu
Chaoyang He 9:45 am – 11:45 am RTH 323		Ke: 11 am – 1 pm MCB 1st floor Hall, west side	Liyu Chen 10 am – 12 pm MCB 1st floor Hall, west side
Jeremy 12 pm – 2 pm MCB 1st floor Hall, west side	Ke: 12 pm – 2 pm MCB 1st floor Hall, west side	Jeremy 1 pm – 3 pm MCB 1st floor Hall, west side	CP 12 pm – 2 pm MCB 1st floor Hall, west side
Chaoyang He 2 pm – 4 pm RTH 323	Victor 3 pm – 4:45 pm SAL 242		Victor 3 pm – 4:45 pm SAL 242

Programming Assignments:

Assignment	Content	Out	Due
PA1	kNN, Decision trees	Sep. 1	Sep 22
PA2	Regressions	Sep 22	Oct 13
PA3	Neural Networks	Oct 13	Nov 3
PA4	HMM & PCA	Nov 3	Nov 24

Theory Assignments:

Assignment	Content	Out	Due
TA1	kNN, Decision trees, Naive Bayes (lectures 1 - 2)	Sep. 6	Sep 20
TA2	Regressions, NN (lectures 3 - 6)	Sep 20	Oct 9
TA3	SVM, Boosting, PCA (lectures 7 - 9)	Oct 24	Nov 8
TA4	Clustering, GMM, HMM (lectures 10 - 12)	Nov 8	Nov 25