CSCE 5215: Machine Learning

Fall 2015

Course Information & Syllabus

Instructor: Cornelia Caragea
Office: F228 Discovery Park
Email: ccaragea@unt.edu
WWW: http://www.cse.unt.edu/∼ccaragea/
Class website: http://www.cse.unt.edu/∼ccaragea/cse5215.html

Teaching Assistant: Aniket Sakinala (AniketSakinala@my.unt.edu)

Lectures: Tue, Thr 2:30pm-3:45pm, Room NTDP B142
Office Hours: Cornelia: Tue & Thr 1:00pm - 2:00pm or by appointment, F228 Discovery Park

Course Objective: The course objectives are to understand machine learning algorithms and identify challenging problems on the Web, learn how to apply machine learning algorithms to these problems, and how to use the existing tools and design new ones. The course will cover both traditional and newly developed machine learning algorithms and their Web applications. Examples of topics include: classification and clustering, kernel methods and Support Vector Machines, Bayesian learning methods, semi-supervised learning techniques, and feature selection, construction, and clustering.

Course Work and Evaluation: There will be one exam for the course. Students will be evaluated based on the exam, homework assignments, reading assignments and paper presentations, and a class project. Students will be required to write summaries and praises/critiques for the assigned readings. Critical thinking will be an important criterion used for evaluating the assignments. Students are encouraged to attend every lecture and to participate in class discussion.

Assignments are due by 11:59pm on the due date. Assignments may be turned in up to 3 days late, with a penalty of 10% for each day late. No credit will be given after 3 days. There will be no final exam for this class. The final is replaced by the project. The grading criterion is shown below:

Section	Weight
Homework	10%
Midterms	40%
Paper Presentation	15%
Project	30%
Class Participation	5%

Prerequisites: Basic knowledge on probability and statistics, basic knowledge on calculus, good understanding of data structures and algorithms. Background in machine learning is not required.

Targeted audience: *Graduate and undergraduate students from Computer Science and related areas.*

Attendance: Attendance is essential and thus is expected.

Required textbooks:

- Pattern Recognition and Machine Learning, Christopher Bishop.

Other Recommended textbooks:

- Machine Learning, Tom Mitchell.
- The Elements of Statistical Learning: Data Mining, Inference and Prediction, Trevor Hastie, Robert Tibshirani, Jerome Friedman (available online at: http://statweb.stanford.edu/∼tibs/ElemStatLearn/).

Topics: The tentative topics are as follows:

Linear Regression
Linear Classification
Generative classification (Naive Bayes)
Practical Issues in Machine Learning
Ensemble Methods
Neural Networks
Instance-based Learning
Decision Tree Learning
Kernel Methods / Support Vector Machines
Clustering / K-Means Clustering
Expectation Maximization, Mixture Models
Semi-supervised Learning
Topic Models

Americans with Disabilities Act: We cooperate with the Office of Disability Accommodation to make reasonable accommodations for qualified students (cf. Americans with Disabilities Act and Section 504, Rehabilitation Act) with disabilities. If you have not registered with ODA, we encourage you to do so. If you have a disability for which you require accommodation, please discuss your needs with the instructor or submit a written Accommodation Request on or before the fourth class day.