# CSE591 (Spring 2007): Machine Learning and Applications

## General Course Information

- Instructor: Dr. Jieping Ye
- Office: BY 568
- Email: jieping.ye@asu.edu
- Meeting Times: TTH 4:40PM–5:55PM
- Office Hours: TTH 10:00AM–11:45AM
- Location: BYAC 270
- Prerequisite: Basics of linear algebra, algorithm design and analysis.
- Web: `http://www.public.asu.edu/~jye02/CLASSES/Spring-2007/`
- Course Textbook: No textbook is required for this course.

## Catalog Description

Clustering, regression, classification, semi-supervised learning, feature reduction, manifold learning, ranking, and kernel learning.

## Objective

An in-depth understanding of some of the most important machine learning techniques and their applications in bioinformatics, computer vision, information retrieval, and other domains.

## Reference Books

- Introduction to Machine Learning, Ethem Alpaydin, 2004.
- Semi-Supervised Learning, Olivier Chapelle, Bernhard Schölkopf, and Alexander Zien, 2006.
- Pattern Recognition and Machine Learning, Christopher M. Bishop, 2006.
- Kernel Methods for Pattern Analysis. John Shawe-Taylor and Nello Cristianini, 2004.
- The Elements of Statistical Learning: Data Mining, Inference, and Prediction. T. Hastie, R. Tibshirani, and J. Friedman, 2001.
- Pattern Classification. Richard Duda, Peter Hart, and David Stork. 2nd edition, 2000.
- Kernel Methods in Computational Biology. Bernhard Schölkopf, Koji Tsuda, and Jean-Philippe Vert, editors.
- Introduction to Data Mining. Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. 2005.

- Gaussian Processes for Machine Learning, Carl Edward Rasmussen and Christopher K.I. Williams, 2006.

## Grading

- Homework (3): 30%. There will be three homeworks assigned.
- Project (1): 40%. Two to three students form a group to carry out a small research project. It can be an implementation and a comparative study of existing methods, a review of a specific topic, or the development of a new idea. Project proposal is due during the fourth week of classes. The intermediate project report is due during the twelfth week of classes. The final project report is due during the last week of classes. There will be a 10-minute final project presentation during the last week of classes.
- Exam (1): 20%. There will be one open-book exam during the second half of the semester.
- Class participation: 10%. Students are required to attend the lecture and participate in the class discussion.

## Tentative Class Schedule

| Week | Topic |
|---|---|
| 1 | Introduction, Basics of Probability, Linear Algebra, and Statistics |
| 2 | Clustering: K-means Clustering, Hierarchical Clustering, Spectral Clustering |
| 3 | Regression: Ridge Regression, Gaussian Process View of Regression |
| 4 | Classification: Nearest Neighbor, Naive Bayes, Linear Discriminant Analysis |
| 5 | Classification: Logistic Regression, Support Vector Machines |
| 6 | Semi-Supervised Clustering |
| 7 | Semi-Supervised Classification |
| 8 | Feature Reduction: Principal Component Analysis, Canonical Correlation Analysis |
| 9 | Spring Break |
| 10 | Feature Reduction: Partial Least Squares |
| 11 | Manifold Learning: Multidimensional Scaling, Isomap, Locally Linear Embedding |
| 12 | Manifold Learning: Laplacian Eigenmaps, Nystrom's method |
| 13 | Ranking: PageRank, HITS |
| 14 | Kernel Learning: Basics |
| 15 | Kernel Learning: Advanced Topics |
| 16 | Final Project Presentation |

## ASU Policies on Academic Integrity

Violations of the University Academic Integrity policy will not be ignored. Penalties include reduced or no credit for submitted work, a failing grade in the class, a note on your official transcript that shows you were punished for cheating, suspension, expulsion and revocation of already awarded degrees. The university requires that should I implement any of these penalties, I must report the matter to the Dean's office. The university academic integrity policy can be found at `http://library.west.asu.edu/refguides/integrity/asu-policies.html`