EIN 4905/EIN 6905 DATA ANALYSIS AND DATA MINING IN SYSTEMS ENGINEERING

Spring 2015
M 1st-2nd period and W 3rd period LIT0121

Catalog Description
EIN 4905 Honors/EIN 6905 provide an insight into the theory background and applications of supervised and unsupervised learning algorithms. Selected topics include Decision Trees, Bayesian Networks, Support Vector Machines, K-Means clustering, Biclustering and Principle Component Analysis. In addition, we will cover material on recent emerging topics such as Robust Data Mining and Massive Data Sets.

Prerequisites or Co-requisites
DMOR (or equivalent) or advanced undergraduates with Matrix Computations ESI 4327C, Statistics STA 4322, and Operations Research ESI 4312

Course Objectives
1. To gain an understanding and appreciation of the principles and methodologies relevant to data mining and data analysis.
2. To solve real-life problems with the sophisticated data mining techniques.
3. To build a solid theoretical background in data mining and explore the recent topics for future research and study.

Required Textbook
Data Mining: Concepts and Techniques by Jaiwei Han, Micheline Kambber & Jian Pei 4th edition, Morgan Kaufmann, 2011. ISBN: **0123814790.**

Recommended Textbooks:
Machine Learning by Tom M. Mitchell, McGraw-Hill Science/Engineering/Math, 1997. ISBN: **0070428077**
Pattern Recognition and Machine Learning by Christopher M. Bishop, Springer, 2007. ISBN: **0387310738**
Data Mining in Agriculture by Antonio Mucherino, Petraq Papajorgji, Panos M. Pardalos, Springer, 2009. ISBN: **0387886141**

Instructor   Dr. Panos M. Pardalos
        http://www.ise.ufl.edu/pardalos
        Office: 401 Weil Hall   **Phone:** 352-392-1464
        Email: pardalos@ufl.edu
Office Hours By appointment

Course Outline
A tentative list of topics for the class is given next. This list might be shortened or lengthened depending on the pace of the class.

1. Introduction to Data Representation and Mining
2. Statistical Method
3. Support Vector Machines and Proximal Support Vector Machines
4. Clustering by k-means
5. Biclustering
6. Bayesian Networks
7. Dimensionality Reduction
8. Validation Methods
9. Application in Matlab
10. Robust Data Mining
11. Massive Data Sets and Future Challenges

Grading Method & Policy

Three in class exams(closed book)	60%
Homework	20%
Project	20%

You will receive numerical grades for all your homework, project and exams. The final letter grade will be determined primarily by the curve. The break between "B" and "B-" will be approximately set at the average of total scores of students receiving letter grades. Letter grades will be monotonic in total course scores. Your grade will be solely based on your performance in this course and not on outside factors like your wish to graduate this semester or possibility of losing a scholarship.

A "C-" will not be a qualifying grade for critical tracking courses. In order to graduate, students must have an overall GPA and an upper-division GPA of 2.0 or better ("C" or better). Note: a "C-" average is equivalent to a GPA of 1.67, and therefore, it does not satisfy this graduation requirement. For more information on grades and grading policies, please visit: https://catalog.ufl.edu/ugrad/current/regulations/info/grades.aspx.

Late Assignments
Assignments are late if not turned in at the first of the period due. Prior approval, or acceptable medical documentation, is necessary for late assignments to receive any credit.

Attendance and Make-up Exams
Requirements for class attendance and make-up exams, assignments, and other work in this course are consistent with university policies that can be found at: https://catalog.ufl.edu/ugrad/current/regulations/info/attendance.aspx.

Honor Code
All students admitted to the University of Florida have signed a statement of academic honesty committing themselves to be honest in all academic work and understanding that failure to comply with this commitment will result in disciplinary action. This statement is a reminder to uphold your obligation as a University of Florida student and to be honest in all work submitted and exams taken in this course. If you conduct any dishonest act (e.g. cheating on an exam, bringing in extra material not allowed during the exams, or copying someone else's homework), you will get an F on that assignment/exam. Further action is possible depending on the severity and repetition of cheating.

Accommodation for Student with Disabilities:
Student with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, www.dso.ufl.edu/drc/) by providing appropriate documentation. Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure as early as possible in the semester.

Online Course Evaluation:
Students are expected to provide feedback on the quality of instruction in this course by completing online evaluations at https://evaluations.ufl.edu. Evaluations are typically open during the last two or three weeks of the semester, but students will be given specific times when they are open. Summary results of these assessments are available to students at https://evaluations.ufl.edu/results/