

## Data Analysis and Data Mining in Systems engineering

EIN 4905/EIN 6905 Section 1A83

**Class Periods:** Tuesday, periods 1-2, 7:25 am – 9:20 am and Thursday, period 2, 8:30 am – 9:20 am

**Location:** FLG 0220

**Academic Term:** Spring 2017

### **Instructor:**

Distinguished Professor Panagote (Panos) M. Pardalos, <http://www.ise.ufl.edu/pardalos/>

Email Address: pardalos@ufl.edu

Office Phone Number: 352-392-1464

Office Hours: By appointment, WEIL 401

### **Teaching Assistants:**

TBA

### **Course Description**

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.

### **Course Pre-Requisites / Co-Requisites**

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

### **Relation to Program Outcomes (ABET):**

| Outcome   | Coverage* |
|---|-----------|
| a. Apply knowledge                                  | Medium    |
| b1. Conduct experiments                             |           |
| b2. Statistical design of experiments               |           |
| c. Design   |           |
| d. Function on teams                                |           |
| e. Solve problems                                   | High      |
| f. Professional and ethical responsibility          |           |
| g. Communicate                                      |           |
| h1. Economic impact                                 |           |
| h2. Global, societal, and environmental impact      |           |
| i. Lifelong learning                                |           |
| j. Contemporary issues                              |           |
| k. Techniques, skills, and tools for degree program | High      |

\*Coverage is given as high, medium, or low. An empty box indicates that this outcome is not part of the course.

### **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

**Materials and Supply Fees**

No fees.

**Required Textbooks and Software**

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

**Recommended Materials**

- 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**

**Online Resources**

Most of the material for this class will be available on E-learning. It is therefore crucial that you know how to efficiently use E-learning.

**Course Schedule**

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

**Attendance Policy, Class Expectations, and Make-Up Policy**

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>.

**Grading Method and Policy**

|                                     |     |
|-------------------------------------|-----|
| Three in class exams (closed books) | 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.

### ***Students Requiring Accommodations***

Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, <https://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.

### ***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/evals>. 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/>.

### ***University Honesty Policy***

UF students are bound by The Honor Pledge which states, "We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment." The Honor Code (<https://www.dso.ufl.edu/sccr/process/student-conduct-honor-code/>) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor or TAs in this class.

### ***Software Use***

All faculty, staff, and students of the University are required and expected to obey the laws and legal agreements governing software use. Failure to do so can lead to monetary damages and/or criminal penalties for the individual violator. Because such violations are also against University policies and rules, disciplinary action will be taken as appropriate. We, the members of the University of Florida community, pledge to uphold ourselves and our peers to the highest standards of honesty and integrity.

### ***Student Privacy***

There are federal laws protecting your privacy with regards to grades earned in courses and on individual assignments. For more information, please see:

<http://registrar.ufl.edu/catalog0910/policies/regulationferpa.html>

## ***Campus Resources:***

### ***Health and Wellness***

#### **U Matter, We Care:**

If you or a friend is in distress, please contact [umatter@ufl.edu](mailto:umatter@ufl.edu) or 352 392-1575 so that a team member can reach out to the student.

**Counseling and Wellness Center:** <http://www.counseling.ufl.edu/cwc>, and 392-1575; and the University Police Department: 392-1111 or 9-1-1 for emergencies.

#### **Sexual Assault Recovery Services (SARS)**

Student Health Care Center, 392-1161.

**University Police Department** at 392-1111 (or 9-1-1 for emergencies), or <http://www.police.ufl.edu/>.

### ***Academic Resources***

**E-learning technical support**, 352-392-4357 (select option 2) or e-mail to Learning-support@ufl.edu. <https://lss.at.ufl.edu/help.shtml>.

**Career Resource Center**, Reitz Union, 392-1601. Career assistance and counseling. <https://www.crc.ufl.edu/>.

**Library Support**, <http://cms.uflib.ufl.edu/ask>. Various ways to receive assistance with respect to using the libraries or finding resources.

**Teaching Center**, Broward Hall, 392-2010 or 392-6420. General study skills and tutoring. <https://teachingcenter.ufl.edu/>.

**Writing Studio, 302 Tigert Hall**, 846-1138. Help brainstorming, formatting, and writing papers. <https://writing.ufl.edu/writing-studio/>.

**Student Complaints Campus:** [https://www.dso.ufl.edu/documents/UF\\_Complaints\\_policy.pdf](https://www.dso.ufl.edu/documents/UF_Complaints_policy.pdf).

**On-Line Students Complaints:** <http://www.distance.ufl.edu/student-complaint-process>.

### ***Teaching Improvement***

We are interested in being the best instructors possible. In particular, we would like to know of the problems you face during the semester as soon as they occur. It is a waste for us to learn at the end of the semester that we were not speaking sufficiently loud to be heard, that our handwriting was not readable, that nobody understood the pictures that were drawn on the board or that the software used for the class was very difficult to use. We want you to feel free to make suggestions to improve the content of the class, its exposition and our instructing skills. You can address these suggestions directly to us (in a polite manner) or anonymously by leaving comments in the instructor mailbox. We will consider carefully all these suggestions and if necessary, we will address them in class.

***Feedback***

If you foresee any problem with adhering to the guidelines set in this syllabus, please discuss them with the instructor as soon as possible. The sooner problems are discussed, the more likely it is that they can be solved.