

### ***Course Information***

DSC540 (previously CSC529)  
Advanced Machine Learning  
Fall 2018-2019  
Wednesday 5:45PM-9:00PM  
Loop Campus, LEWIS 01216  
Course Management System: <http://d21.depaul.edu>

### ***Instructor Information***

Instructor: Daniela Stan Raicu  
Office: CDM Center, Room 718  
Office Hours: Wednesday, 4:15-5:00pm, 9:00pm-9:45pm  
Thursday, 4:15pm-5:00pm, 9:00pm-9:45pm  
Other times by appointment  
Advising Hours: Wednesday, 3:15-4:15 pm  
Thursday, 11-12pm, 3:15-4:15pm  
Phone: (312) 362-5512  
Fax: (312) 362-6116  
Email: [draicu@cdm.depaul.edu](mailto:draicu@cdm.depaul.edu)  
Home page: <http://facweb.cs.depaul.edu/~dstan>

### ***Course Description***

The course is for students with prior background in data mining or machine learning techniques. The course will cover advanced modeling techniques, such as ensemble learning, extended linear models, probabilistic graphical models, mixture and latent variable models, and matrix factorization. First the theoretical foundations of these techniques will be presented and augmented with in-class examples and homework problems. Second, the state-of-the-art research related to these techniques will be presented and augmented with paper reviews that highlight the practical applications of these advanced data mining techniques. Applications of the models will be presented in various domains, including social computing, visual computing, and biomedical and health informatics.

### ***Course Learning Goals***

At the end of the course, students should be able to:

- understand the basics behind each data mining method as well as the respective cons and pros
- understand how information in real world applications can be formulated and represented as different genres of data, such as matrices, sequences, data streams, graphs/networks
- select, combine, and apply specific data mining techniques for certain data types and challenges, and understand, explain, and interpret the obtained results, and
- identify recent trends and open directions in the field of data mining.

### ***Recommended Books***

- *Data Mining: Practical Machine Learning Tools and Techniques* by Witten, Frank, and Hall, 3rd Edition, ISBN 978-0-12-374756-0
  - This book has a focus of practical applications and the use of the WEKA toolkit.
- *Probabilistic Graphical Models, Principles and Techniques* by Daphne Koller and Nir Friedman, ISBN 978-0-262-01319-2
  - This book has a focus on theoretical foundations of probabilistic graphical models <http://mitpress.mit.edu/books/probabilistic-graphical-models>
  - eBook available through the DePaul library at [http://depaul.worldcat.org/title/data-mining-practical-machine-learning-tools-and-techniques/oclc/706802868&referer=brief\\_results](http://depaul.worldcat.org/title/data-mining-practical-machine-learning-tools-and-techniques/oclc/706802868&referer=brief_results)
- *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* by Hastie, Tibshirani, Friedman
  - This book has a focus on theoretical foundations of data mining; PDF available at <https://web.stanford.edu/~hastie/Papers/ESLII.pdf>

***Prerequisites***

DSC424 (formally CSC 424) and (DSC 441 –formerly IS467, ECT 584, CSC 578, or CSC478)

***Grading***

The homework/programming assignments will be worth 40% of the course grade, paper reviews will be 20%, and the final project will be worth 40%.

The summary of the weights of each assignment for contributing to the final grade is as follows:

Assignment	Weight in final grade
Homeworks & Programming Assignments	40%
Paper Review	20%
Final Project	40%

The final grade will be assigned according to the following scale:

Percentage Grade	Letter Grade	Manner of fulfillment
95-100	A	Excellent
90-94	A-	
85-89	B+	
80-84	B	
75-79	B-	Very Good
70-74	C+	
65-69	C	
60-64	C-	
55-59	D+	Satisfactory
50-54	D	
0 – 50	F	
		Poor

***Homework/Programming Assignments, Papers' Reviews, and Final Project Policies*****Homework/programming assignments**

There will be 4 homework assignments during the quarter. Work to be submitted for the course is generally due one or two weeks after it was assigned; late submissions are allowed with a 5%, 10%, and 15% penalty for a one day, two days, and three days, respectively. No late work will be accepted after three days since the assignment was due.

The assignments must be submitted online on the D2L site at <https://D2L.depaul.edu>. Only legible, organized homework which shows your work will be graded. Include your name, section number, date, and homework number on the first page of your assignment. It is your responsibility to check that your files are uploaded correctly on D2L; you should always keep a copy of your submission.

Extra credit points will be given for additional problems in assignments and midterm, paper reviews, active participation in the lectures and Discussion Forum.

**Paper Review:**

Throughout the quarter, the students will be also provided with a list of research papers related to the theory discussed in class. Each student will have to review these papers and participate in their discussion. Two paper reviews will have to be submitted and they will represent 20% of the final grade.

The paper readings will be selected from the recent literature in major journals and conference proceedings in the field of data mining. They include but are not limited to the ACM KDD Conference on Knowledge Discovery and Data Mining (KDD), the IEEE International Conference on Data Mining (ICDM), the SIAM International Conference on Data Mining, the IEEE Transactions on Knowledge and Data Engineering (TKDE), and the ACM Transactions on Knowledge Discovery from Data (TKDD).

### Final Project

The purpose of the final project is to demonstrate students' ability to apply the knowledge and the techniques learned during this course. The final project for this class is more extensive analysis task, chosen by students from among the topics we discuss. Final projects will include a presentation to the rest of the class at the end of the quarter, in place of a final exam. As part of your final project, you will also be asked to critique your classmates' projects. These critiques will be collected, collated, and passed on anonymously to the presenter.

Whenever it is possible, it is recommended that the DL students attend the final presentations to participate in the live discussions of the final project. However, appropriate accommodations will be made for the DL students not being able to give the presentations in class; the DL students will still have to submit their critiques on the other projects.

Deliverables for the final project:

Proposal (October 10<sup>th</sup>): One-page proposal describing the problem, data, proposed approach, and at least three references other than text book or class notes.

Presentation (November 7<sup>th</sup> & November 14<sup>th</sup>): Each project is to be presented using PowerPoint, and the PPT file will have to be submitted to be published on course web site.

Report (November 15<sup>th</sup>): The report will be written in the format of a paper (abstract, introduction, literature review, methodology, results, discussion, conclusions and future work). The literature review for the final report consists of reading and summarizing about 5 to 6 published papers on the review topic. While the internet can serve as a good source of information, the DePaul Library also has extensive holdings, most of them available electronically.

### ***Software***

The use of Matlab will be taught in class. There will be also a lab session scheduled during the second week of school to familiarize the students with Matlab and the Statistics Toolbox. Students may use any data mining/data science/machine learning tool of their choice when completing class assignments (see useful information document for more information).

### ***Attendance***

It is expected that you will attend every class and remain for the duration; it is the single most important action you can take in mastering the course objectives. Coming 15 minutes late or leaving 15 minutes constitutes an absence for the student. You are responsible for all material covered, assignments delivered or received, and announcements made in class sessions that you miss. For distance learning students, this means viewing the classes in a timely manner, participate in the discussion forum, and being sure to email or call in any questions that you have.

### **For online students:**

Recordings of each lecture will be available a few hours after the "live" class, and can be found at the course website <https://d2l.depaul.edu>. Online students are expected to watch the lectures every week and to keep up with the course information posted on the course website.

### ***Email***

Email is the primary means of communication between faculty and students enrolled in this course outside of class time. Students should be sure their email listed under "demographic information" at <http://campusconnect.depaul.edu> is correct.

### ***Changes to Syllabus***

This syllabus is subject to change as necessary to better meet the needs of the students. Significant changes are unlikely, and will be thoroughly addressed in class. Minor changes, especially to the weekly agenda, are possible at any time. If a change occurs, it will be thoroughly addressed during class and posted under Announcements in D2L.

### ***Class Cancellation***

Unless DePaul University closes because of weather, we will have class.

### ***Attitude***

A professional and academic attitude is expected throughout this course. Measurable examples of non-academic or unprofessional attitude include but are not limited to: talking to others when the instructor is speaking, mocking another's opinion, cell phones ringing, emailing, texting or using the internet whether on a phone or computer. If any issues arise a student may be asked to leave the classroom. The professor will work with the Dean of Students Office to navigate such student issues.

### ***Civil Discourse***

DePaul University is a community that thrives on open discourse that challenges students, both intellectually and personally, to be Socially Responsible Leaders. It is the expectation that all dialogue in this course is civil and respectful of the dignity of each student. Any instances of disrespect or hostility can jeopardize a student's ability to be successful in the course. The professor will partner with the Dean of Students Office to assist in managing such issues.

### ***Cell Phones/On Call***

If you bring a cell phone to class, it must be off or set to a silent mode. Should you need to answer a call during class, students must leave the room in an undistruptive manner. Out of respect to fellow students and the professor, texting is never allowable in class. If you are required to be on call as part of your job, please advise me at the start of the course.

## ***School policies:***

### ***Online Course Evaluations***

Evaluations are a way for students to provide valuable feedback regarding their instructor and the course. Detailed feedback will enable the instructor to continuously tailor teaching methods and course content to meet the learning goals of the course and the academic needs of the students. They are a requirement of the course and are key to continue to provide you with the highest quality of teaching. The evaluations are anonymous; the instructor and administration do not track who entered what responses. A program is used to check if the student completed the evaluations, but the evaluation is completely separate from the student's identity. Since 100% participation is our goal, students are sent periodic reminders over three weeks. Students do not receive reminders once they complete the evaluation. Please see <https://resources.depaul.edu/teaching-commons/teaching/Pages/online-teaching-evaluations.aspx> for additional information.

### ***Academic Integrity and Plagiarism***

This course will be subject to the academic integrity policy passed by faculty. More information can be found at <https://resources.depaul.edu/teaching-commons/teaching/academic-integrity/Pages/default.aspx>

The university and school policy on plagiarism can be summarized as follows: Students in this course should be aware of the strong sanctions that can be imposed against someone guilty of plagiarism. If proven, a charge of plagiarism could result in an automatic F in the course and possible expulsion. The

strongest of sanctions will be imposed on anyone who submits as his/her own work any assignment which has been prepared by someone else. If you have any questions or doubts about what plagiarism entails or how to properly acknowledge source materials be sure to consult the instructor.

### ***Withdrawal***

Students who withdraw from the course do so by using the Campus Connection system (<http://campusconnect.depaul.edu>). Withdrawals processed via this system are effective the day on which they are made. Simply ceasing to attend, or notifying the instructor, or nonpayment of tuition, does not constitute an official withdrawal from class and will result in academic as well as financial penalty. The dropping dates can be found at: <https://academics.depaul.edu/calendar/Pages/default.aspx>

### ***Retroactive Withdrawal***

This policy exists to assist students for whom extenuating circumstances prevented them from meeting the withdrawal deadline. During their college career students may be allowed one medical/personal administrative withdrawal and one college office administrative withdrawal, each for one or more courses in a single term. Repeated requests will not be considered. Submitting an appeal for retroactive withdrawal does not guarantee approval.

All students are required to manage their class schedules each term in accordance with the deadlines for enrolling and withdrawing as indicated in the University Academic Calendar. Information on enrollment, withdrawal, grading and incompletes can be found at:

<http://www.cdm.depaul.edu/Current%20Students/Pages/PoliciesandProcedures.aspx>

### ***Excused Absence***

In order to petition for an excused absence, students who miss class due to illness or significant personal circumstances should complete the Absence Notification process through the Dean of Students office. The form can be accessed at <http://studentaffairs.depaul.edu/dos/forms.html>. Students must submit supporting documentation alongside the form. The professor reserves the sole right whether to offer an excused absence and/or academic accommodations for an excused absence.

### ***Incomplete Grades***

An incomplete grade is a special, temporary grade that may be assigned by an instructor when unforeseeable circumstances prevent a student from completing course requirements by the end of the term and when otherwise the student had a record of satisfactory progress in the course. All incomplete requests must be approved by the instructor of the course and a CDM Associate Dean. Only exceptions cases will receive such approval. Information about the Incomplete Grades policy can be found at <http://www.cdm.depaul.edu/Current%20Students/Pages/Grading-Policies.aspx>

### ***Students with Disabilities***

Students seeking disability-related accommodations are required to register with DePaul's Center for Students with Disabilities (CSD) enabling them to access accommodations and support services to assist with their success. There are two office locations:

- Loop Campus – Lewis Center #1420 – (312) 362-8002
- Lincoln Park Campus – Student Center #370 – (773) 325-1677

Students who register with the Center for Students with Disabilities are also invited to contact Dr. Gergory Moorhead, Director of the Center, privately to discuss how he may assist in facilitating the accommodations to be used in a course. This is best done early in the term. The conversation will remain confidential to the extent possible.

Please see <https://offices.depaul.edu/student-affairs/about/departments/Pages/csd.aspx> for Services and Contact Information.