

Course Information

DSC441 (previously IS467)
Fundamentals of Data Science
Autumn Quarter 2019-2020
Thursday 5:45PM-9:00PM
Loop Campus, CDM Center 218
Course Management System: <http://d2l.depaul.edu>

Instructor Information

Instructor: Thiruvarangan (Thiru) Ramaraj
Office: CDM Center, Room 703
Office Hours: Will be posted on Blue Star
Phone: 312-362-8223
Email: tramaraj@depaul.edu

Course Description

The greatest challenge facing data warehousing professionals is extracting valuable information from the masses of data in the warehouse. One of the most significant and powerful technologies to address this concern is data mining.

Data mining uses statistical analysis and modeling techniques to uncover patterns and relationships hidden in large databases—patterns that ordinary methods might miss. Data mining is only one step in the knowledge discovery process. Other steps include identifying the problem to be solved, collecting and preparing the right data, interpreting and deploying models, and monitoring the results. The real key to success, however, is to have a thorough understanding of your data and your business.

This course will illustrate the knowledge discovery process and how the technology works with sample applications of data mining. The course will cover the following topics:

- What data mining is and is not (Chapter 1)
- Relationship between data mining, data warehouse, and query tools (Chapter 4)
- Applications and trends in data mining (Chapter 13)
- Data understanding and preparation for the data mining process (Chapters 2 & 3)
- Model building, algorithms and technology:
 - o Supervised learning:
 - § Classification and Prediction (Chapters 8& 9)
 - o Unsupervised learning
 - § Clustering (Chapters 10 & 11)
- Data Mining for complex data objects (Chapter 13) –if time permits

Course Objectives

After completing the course, the students will be able to:

- Students will be able to understand and apply basic concepts, terminology, models and methods in the field of data mining
- Students will be able to apply different data mining algorithms to large datasets
- Students will be able to evaluate and compare different data mining algorithms
- Students will be able to identify, design, and recommend different data mining processes for specific problems

Required Textbook

Data Mining: Concepts and Techniques, by Han and Kamber and Pei
Morgan Kaufman Publishers, Third Edition

Textbook webpage:

<http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf>

Prerequisites: IT403: Statistics and Data Analysis

Grading

The homework assignments will be worth 60% of the course grade and the final exam will be worth 40%. The first four assignments will be worth 10% each. The fifth assignment will be worth 20% and it will be designed to demonstrate your ability to apply the knowledge and the data mining techniques learned during this course to a dataset relevant to the course topics. It is the most extensive assignment among the five homework assignments.

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

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

Assignment	Weight in final grade
Homework Assignments	60%
Final Exam	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-	Very Good
85-89	B+	
80-84	B	
75-79	B-	
70-74	C+	Satisfactory
65-69	C	
60-64	C-	Poor
55-59	D+	
50-54	D	
0 – 50	F	

Homework Assignments and Final Exam Policies

Homework assignments

There will be 5 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.

Final exam

There will be a final exam that will be worth 40% of the course grade; the final exam is a closed book and notes exam, but students are allowed to bring a calculator (no phones are allowed) and a one page (single-sided) of formulas.

For online students

Online students must schedule their final exam at the D2L website during the time frame specified by your instructor. Students living within the Chicago land area are considered local and will be expected to take their exams at a DePaul University campus. Time slots vary by campus and day. Online students living outside the Chicago land area (remote) will have their exams administered by a qualified proctor. You will need to find an acceptable proctor in your area before you register for your exam. Detailed information about online exams is at <https://www.cdm.depaul.edu/onlinelearning/Pages/Exams.aspx>

Software

SPSS will be taught in class and all the assignments can be implemented using SPSS, but students may use any data mining tool of their choice when completing class assignments. There will be two hands-on SPSS tutorials on September 19th (week 2) and October 17th (week 6) and an SPSS practice final session on November 7th (week 9, if needed). The tutorials will be recorded and available to both in-class and online students. (see useful information document for more info on the software).

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.