Course Information

Two sections:	Tuesday 5:45-9:00 PM	Wednesday 5:45-9:00 PM
Locations (Loop):	CDM 00226	LEWIS 01107

Instructor Information

Instructor: Eli T. Brown
Office: CDM Center, Room 711

Office Hours:	Tue & Wed, 9:00-9:45pm
	Wed 12:30-1:30pm
	Wed 4:30-5:00pm
	Other times by appointment
Phone:	312-362-8381
Email:	ebrown80@cdm.depaul.edu

Course Description

The course will teach advanced statistical techniques to discover insight from data. The course topics include visualization techniques to summarize and display high dimensional data, dimensional reduction techniques such as principal component analysis and factor analysis, clustering techniques for discovering patterns, and classification techniques for decision making. The methods will be implemented using standard computer packages.

Course Goals

At the end of this course, the student should have a basic understanding of the following topics and be able to identify which approach is appropriate for a given data set and data analysis task to be performed:

- Multivariate linear regression (least-square estimation & normal equations, model building & variable selection)
- Canonical Correlation (to assess the relationship between two sets of variables)
- Principal component analysis & Factor analysis (Eigenvalues and Eigenvectors, scree plots, dimension reduction, factor rotation)
- Discriminant analysis (Fisher's discriminant function)
- Cluster analysis (similarity measures, hierarchical clustering & non-hierarchical clustering).
- Multidimensional Scaling and visualization techniques (if time permits)

Recommended Books

• Hair, Black, Babin, & Anderson, "Multivariate Data Analysis", Published by Prentice Hall, ISBN-13: 9780138132637, 2010 (7th edition).
• Johnson & Wichern, "Applied Multivariate Statistical Analysis", Published by Prentice Hall, ISBN-13: 9780131877153, 2008 (6th edition).

Prerequisites

CSC324/CSC423: Data Analysis and Regression

Grading

The homework/programming assignments will be worth 40% of the course grade. There will be two midterms, one around the middle of the quarter and one on the last day of class (called a midterm because it will focus on the material of the latter half of the course). Each midterm will be worth 20% of the grade. There will be a take-home final exam that's more of a small project, due on the University-assigned day for the final, also worth 20% of the grade.

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%
Midterm I	20%
Midterm II	20%
Take-home Final Exam	20%

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+	Very Good
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/Programming Assignments, Papers' Reviews, and Exam Policies

Homework/programming assignments
There will be 4 homework/programming assignments, which are due at the beginning of class one or two weeks after they are assigned. Late assignments will be accepted up to one lecture later than the assigned due date with a 25% penalty – this penalty will be assessed in full to assignments turned in from the end of class on the day that the assignment is due up until the beginning of next lecture. No assignments will be accepted beyond the beginning of class one lecture beyond the due date. The assignments must be submitted online on the Course on Line site at https://col.cti.depaul.edu

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

Midterm:
There will be two midterms given in class. One in the sixth week and one during the final class period. They will be worth 20% of the course grade each; the midterms are 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.

Final Exam
The final exam is a take-home exam due on the University-assigned exam day. It is meant to demonstrate your ability to apply the knowledge and the techniques learned during this course to a dataset relevant to the course topics. It will be written in the format of a technical report outlining the problem, summarizing the data, describing the methodology, presenting and visualizing the results, and making recommendations.

Online students must schedule their final exam at the COL website (https://col.cdm.depaul.edu) during the time frame specified by your instructor (Midterm I: February 9 – 13, Midterm II: March 8 – 13). 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. They can also take the exam with the other in-class students at the official exam time. 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
http://blogs.cti.depaul.edu/colwiki/Wiki%20Pages/How%20Do%20I%20Take%20My%20Exams.aspx

Software
There are a number of great statistics packages available and you may use any that you choose for assignments. SPSS will be used in class, although I can mix it up if students prefer. There will be also two lab sessions scheduled during the second and fifth week of school, though the schedule is subject to change. Students may use any statistical analysis tool of their choosing when completing class assignments.

Attendance
It is expected that you will attend every class; it is the single most important action you can take in mastering the course objectives. 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 online discussions, 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://col.cdm.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.

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. You will be informed of all such changes.

Class Cancellation
      Unless DePaul University closes because of weather, we will have class.
      Note: I will be away with research obligations next week but we will have fill-in professors.

School policies:

Online Instructor Evaluation
Course and instructor evaluations are critical for maintaining and improving course quality. To make evaluations as meaningful as possible, we need 100% student participation. Therefore, participation in the School's web-based academic administration initiative during the eighth and ninth week of this course is a requirement of this course. Failure to participate in this process will result in a grade of incomplete for the course. This incomplete will be automatically removed within seven weeks after the end of the course and replaced by the grade you would have received if you had fulfilled this requirement.

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.

Academic Integrity Policy

I expect that you have read and understood DePaul's policy on Academic Integrity: http://academicintegrity.depaul.edu/ It is part of this syllabus; follow it.

Plagiarism

The university and school policy on plagiarism can be summarized as follows: Students in this course, as well as all other courses in which independent research or writing play a vital part in the course requirements, 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 a report, examination paper, computer file, lab report, or other 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.

Incomplete

An incomplete grade is given only for an exceptional reason such as a death in the family, a serious illness, etc. Any such reason must be documented. Any incomplete request must be made at least two weeks before the final, and approved by the Dean of the School of Computer Science, Telecommunications and Information Systems. Any consequences resulting from a poor grade for the course will not be considered as valid reasons for such a request. Students must formally request an incomplete by filling out a Request for Incomplete Grade form, available at the CDM main office, and submitting it to me.

Resources for Students with Disabilities

Students who feel they may need an accommodation based on the impact of a disability should contact the instructor privately to discuss their specific needs. All discussions will remain confidential. To ensure that you receive the most appropriate accommodation based on your needs, contact the instructor as early as possible in the quarter (preferably within the first week of class), and make sure that you have contacted either:

PLuS Program (for LD, AD/HD) at 773-325-4239 in SAC 220
The Office for Students with Disabilities (for all other disabilities) at 773-325-7290 Student Center 307