Math 4685/6685 Statistical Learning I

Course Description: Math 4685/6685 is the first course in a two-course sequence designed to introduce students to the statistical foundations of tools used in the modeling and understanding of big data in the context of developments in computer science. Emphasis is on application to real data sets. Topics include linear regression, shrinkage methods, classification algorithms, principle component analysis, support vector algorithms and clustering methods.
Prerequisite: Math 4614 or Math 4635 and Math 3242 or permission of the instructor.

Learning Outcomes: Upon completion of this course, students will be able to select, apply, and evaluate appropriate techniques from statistical and computational methods to typical data sets using linear modeling and classification.

Textbook: *An Introduction to Statistical Learning with Applications in R* by Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, Springer, 2013. Available at http://www.springer.com/us/book/9781461471370.

Grading:	**Math 4685**	**Math 6685**
Homework	25%	20%
In-class work	25%	20%
Exams	40%	40%
Project	10%	20%

Students taking Math 6685 will complete two data science projects, while Math 4685 students will complete only one.

Grading scale:

90-100%	A
80-89%	B
70-79%	C
60-69%	D
Below 60%	F

Tentative Schedule:

Week 1:	What is Statistical Learning?
Week 2:	Introduction to R programming or another computational platform
Week 3:	Assessing model accuracy
Week 4:	Statistical foundations of simple linear regression
Week 5:	Statistical foundations of multiple linear regression
Week 6:	Statistical considerations in the regression model
Week 7:	Statistical foundations of logistic regression
Week 8:	Foundations of discriminant analysis
Week 9:	Principles of variable selection
Week 10:	Shrinkage methods

Week 11: Dimension reduction methods
Week 12: Statistical foundations of analysis in high dimensions
Week 13: non-parametric classifiers
Week 14: Statistical foundations of hierarchical clustering
Week 15: Statistical foundation of non-hierarchical clustering

Academic Misconduct:
The University of Memphis defines cheating and plagiarism as:

Plagiarism - The adoption or reproduction of ideas, words, statements, images, or works of another person as one's own without proper attribution.

Cheating - Using or attempting to use unauthorized materials, information, or aids in any academic exercise or test/examination. The term academic exercise includes all forms of work submitted for credit or hours.

If plagiarism or cheating occurs, the student will receive a failing grade on the assignment and (at the instructor's discretion) a failing grade in the course. The course instructor may also elect to forward the incident to the Office of Student Conduct for further disciplinary action. More information is available at http://www.memphis.edu/studentconduct/pdfs/csrr.pdf.

Special accommodation:
If you need special accommodation, please let the instructor know immediately.