NJIT New Jersey's Science & Technology University

THE COLLEGE OF SCIENCE AND LIBERAL ARTS

THE DEPARTMENT OF MATHEMATICAL SCIENCES

MATH 678-102: Introduction to Statistical Methods in Data Science
Spring 2020 Course Syllabus

NJIT Academic Integrity Code: All Students should be aware that the Department of Mathematical Sciences takes the University Code on Academic Integrity at NJIT very seriously and enforces it strictly. This means that there must not be any forms of plagiarism, i.e., copying of homework, class projects, or lab assignments, or any form of cheating in quizzes and exams. Under the University Code on Academic Integrity, students are obligated to report any such activities to the Instructor.

COURSE INFORMATION

Course Description: This course introduces to students concepts in statistical methods used in data science, including data collection, data visualization and data analysis. Emphasis is on model building and statistical concepts related to data analysis methods. The course provides the basic foundational tools on which to pursue statistics, data analysis and data science in greater depth. Topics include sampling and experimental design, understanding the aims of a study, principles of data analysis, linear and logistic regression, resampling methods, and statistical learning methods. Students will use the R statistical software.

Number of Credits: 3

Prerequisites: Math 661 or Math 663 or permission by instructor.

Course-Section and Instructors

Course-Section	Instructor
Math 678-102	Professor W. Guo

Office Hours for All Math Instructors: Spring 2020 Office Hours and Emails

Required Textbook:

Title	*An Introduction to Statistical Learning: with Applications in R*
Author	Gareth James, et al.
Edition	1st (2013 ed.)
Publisher	Springer
ISBN #	978-1461471370
Reference	*The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, by Hastie, Tibshirani, and Friedman; Publisher: Springer, 2nd edition (2009); ISBN: 978-0387848570.

University-wide Withdrawal Date: The last day to withdraw with a **W** is **Monday, April 6, 2020**. It will be strictly enforced.

POLICIES

DMS Course Policies: All DMS students must familiarize themselves with, and adhere to, the Department of Mathematical Sciences Course Policies, in addition to official university-wide policies. DMS takes these policies very seriously and enforces them strictly.

Grading Policy: The final grade in this course will be determined as follows:

Homework	25%
Project	15%
Midterm Exam	25%
Final Exam	35%

Your final letter grade will be based on the following tentative curve.

A	90 - 100	**C+**	75 - 79
B+	85 - 89	**C**	70 - 74
B	80 – 84	**F**	0 - 69

Attendance Policy: Attendance at all classes will be recorded and is **mandatory**. Please make sure you read and fully understand the Math Department's Attendance Policy. This policy will be strictly enforced.

Exams: There will be one midterm exam held in class during the semester and one comprehensive final exam. Exams are held on the following days:

Midterm Exam	March 7, 2019
Final Exam Period	May 10 - 16, 2019

The final exam will test your knowledge of all the course material taught in the entire course. Make sure you read and fully understand the Math Department's Examination Policy. This policy will be strictly enforced.

Makeup Exam Policy: There will be **NO MAKE-UP QUIZZES OR EXAMS** during the semester. In the event an exam is not taken under rare circumstances where the student has a legitimate reason for missing the exam, the student should contact the Dean of Students office and present written verifiable proof of the reason for missing the exam, e.g., a doctor's note, police report, court notice, etc. clearly stating the date AND time of the mitigating problem. The student must also notify the Math Department Office/Instructor that the exam will be missed.

Cellular Phones: All cellular phones and other electronic devices must be switched off during all class times.

ADDITIONAL RESOURCES

Math Tutoring Center: Located in the Central King Building, Lower Level, Rm. G11 (See: Spring 2020 Hours)

Further Assistance: For further questions, students should contact their instructor. All instructors have regular office hours during the week. These office hours are listed on the Math Department's webpage for Instructor Office Hours and Emails.

All students must familiarize themselves with and adhere to the Department of Mathematical Sciences Course Policies, in addition to official university-wide policies. The Department of Mathematical Sciences takes these policies very seriously and enforces them strictly.

Accommodation of Disabilities: Disability Support Services (DSS) offers long term and temporary accommodations for undergraduate, graduate and visiting students at NJIT.

If you are in need of accommodations due to a disability please contact Chantonette Lyles, Associate Director of Disability Support Services at 973-596-5417 or via email at lyles@njit.edu. The office is located in Kupfrian Hall, Room 201. A Letter of Accommodation Eligibility from the Disability Support Services office authorizing your accommodations will be required.

For further information regarding self identification, the submission of medical documentation and additional support services provided please visit the Disability Support Services (DSS) website at:

- https://www.njit.edu/studentsuccess/accessibility/

Important Dates (See: Spring 2020 Academic Calendar, Registrar)

Date	Day	Event
January 21, 2020	T	First Day of Classes
January 31, 2020	F	Last Day to Add/Drop Classes
March 15 - 22, 2020	Su-Su	Spring Recess: No Classes/ University Open
April 6, 2020	M	Last Day to Withdraw
April 10, 2020	F	Good Friday - University Closed
May 5, 2020	T	Friday Classes Meet - Last Day of Classes
May 6 & 7, 2020	W & R	Reading Days
May 8 - 14, 2020	F - R	Final Exam Period

Course Outline

Date	Lecture	Sections	Topic	Assignment
WEEK 1 - 1/22	1	Chapter 1	Introduction to Data Science	
WEEK 2 - 1/29	2	Chapter 2	Statistical Learning; KNN	Homework 1
WEEK 3 - 2/5	3	Chapter 3	Linear Regression; R Lab	
WEEK 4 - 2/12	4	Chapter 3	Linear Regression (Cont.)	Homework 2
WEEK 5 - 2/19	5	Chapter 4	Logistic Regression	
WEEK 6 - 2/26	6	Chapter 4	LDA, QDA; R Lab	Homework 3
WEEK 7 - 3/4	7	Chapter 5	Cross-Validation and Bootstrap	
WEEK 8 - 3/11		**MIDTERM EXAM: MONDAY ~ MARCH 11, 2020**		
WEEK 9 - 3/18		SPRING RECESS (NO CLASSES)		

WEEK 10 - 3/25	8	Chapter 6	Linear Model Selection; R Lab	Homework 4
WEEK 11 - 4/1	9	Chapter 6	Shrinkage Methods and Dimension Reduction Methods	Course Project
WEEK 12 - 4/8	10	Chapter 7	Nonlinear Modeling; R Lab	Homework 5
WEEK 13 - 4/15	11	Chapter 8	Tree-Based Methods: Bagging, Random Forests, Boosting	
WEEK 14 - 4/22	12	Chapter 9	Support Vector Machines	Homework 6
WEEK 15 - 4/29	13	Chapter 10	Unsupervised Learning	
WEEK 16 - 5/6			Reading Day 1	Deadline of the project report
WEEK 17 - 5/13			**FINAL EXAM: MONDAY ~ MAY 13, 2020**	

Updated by Professor W. Guo - 1/21/2020
Department of Mathematical Sciences Course Syllabus, Spring 2020