Rutgers University
The State University of New Jersey
Department of Economics - CCAS
MS in Data Science Program
Spring 2026

Class Information

Course Title:	Applied Data Mining & Machine Learning (ADM & ML) Economics 50:220:422:92/MS in Data Science 56:219:531:92
Instructor:	Dr. I-Ming Chiu
Office:	ARMITAGE 435, Phone: (856) 225-6012
E-mail address:	ichiu@camden.rutgers.edu
Class Meeting:	ATG-101 (Armitage Hall), 6:00-8:50 pm, Tuesday [Note: We will meet in person in ATG-101, except for three virtual sessions held via Zoom on 2/3, 2/10, and 2/17.]
Office Hours:	1:00-3:00 pm, Tuesday or by appointment
Course Description:	In today's world, massive amounts of data are generated every moment—from emails and social media posts to financial transactions. The key challenge is how to turn this data into meaningful patterns, insights, and predictions. This is the focus of **Data Mining & Machine Learning (DM & ML)**. DM & ML uses mathematical and statistical algorithms to uncover structure in large datasets and convert raw information into actionable knowledge. This course introduces major supervised and unsupervised learning methods, including linear and logistic regression, K-Nearest Neighbors, Support Vector Machines, Naïve Bayes, Association Rules, and K-Means Clustering. Each topic is paired with case studies to build practical, hands-on skills. The goal is to equip students with modern analytical tools that are in high demand across the job market. For an overview of data-science-related career prospects, see: https://www.indeed.com/career/data-scientist/salaries
Main Readings:	Bin Yu & Rebecca L. Barter, Veridical Data Science: The Practice of Responsible Data Analysis and Decision Making, MIT Press 2024 (book abbreviation: VDS) Fred Nwanganga & Mike Chapple, Practical Machine Learning in R, 1st edition, Wiley Publishing 2020 (book abbreviation: PMLR). Ethem Alpaydin, Introduction to Machine Learning, 4th edition, MIT Press 2020 (book abbreviation: IML). Brett Lantz, Machine Learning with R, 4th edition, Packt Publishing 2019 (ISBN: 978-1-80107-132-1; book abbreviation: MLR).

Rutgers University
The State University of New Jersey
Department of Economics - CCAS
MS in Data Science Program
Spring 2026

Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, An Introduction to Statistical Learning/with Applications in R (Python), 2nd edition, Springer 2021 (book abbreviation: ISLR).

Other References: Hadley Wickham & Garrett Grolemund, R for Data Science: Import, Tidy, Transform, Visualize, and Model Data, O'Reilly Media 2017.

Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, and Vipin Kumar, Introduction to Data Mining 2nd edition, Pearson 2018.

Computing: All the computations will be done using an open source statistical software R. It can be downloaded at http://www.r-project.org. You're encouraged to download and use RStudio at the following site, which is a very user-friendly IDE (integrated development environment) for R. https://www.rstudio.com/products/rstudio/

R Installation: https://www.youtube.com/watch?v=8NvvydRwxEI (for Mac)
https://www.youtube.com/watch?v=TsnGd6p9oTk (for PC)

R Download: https://cran.r-project.org/bin/windows/base/
https://cran.r-project.org/bin/macosx/

RStudio Download: https://posit.co/download/rstudio-desktop/

Class Material: Data, handouts, assignments, and additional readings will be posted on Canvas website.

Online Learning: https://www.datacamp.com/ (Learn Data Science online)

Useful Websites: https://bookdown.org/f_lennert/introduction-to-r/ (Learn R)

https://www.analyticsvidhya.com/ (Machine Learning)

Academic Calendar: https://camden.rutgers.edu/registrar/catalogs-calendars/2025-2026#spring

Grading: **Contributions to Final Grade**

- Attendance	5%
- DataCamp courses	10%
- Take-home problems	20%
- Midterm Exams (2)	40%
- Final Exam or Project	25%
- Participation (extra credit)	5%

Grading Policy: Term grades will be based on the final distribution of the above grading weights.

Exam Preparation: The exam questions will be drawn from three sources: (i) homework assignments, (ii) course lectures, and (iii) reading material.

Class Attendance: Class attendance is essential for learning achievement. When missing a class, it would cost you more time to learn on your own. I strongly recommend the following steps for your successful learning: (1) attend every class and take notes; (2) review everything you learn from the class immediately, never put it off; (3) ask questions and participate in class discussions.

Academic Conduct: Make up exams will be given **only upon prior notice**. I request prior knowledge of any expected absence from an exam. If this is not feasible, you can document a valid reason for missing the exam. Unexcused absence on any exam will result in a grade of zero. Dishonesty in seeking an excused absence or in the examination process will result in a grade of zero on the exam involved and in university discipline. More detailed information can be found at the following site: https://academicintegrity.rutgers.edu.

Disability Services: Rutgers University welcomes students with disabilities into all of the University's educational programs. In order to receive consideration for reasonable accommodations, a student with a disability must contact the appropriate disability services office at the campus where you are officially enrolled, participate in an intake interview, and provide documentation:

https://ods.rutgers.edu/students/documentation-guidelines. If the documentation supports your request for reasonable accommodations, your campus's disability services office will provide you with a Letter of Accommodations. Please share this letter with your instructors and discuss the accommodations with them as early in your courses as possible. To begin this process, please complete the Registration form at https://webapps.rutgers.edu/student-ods/forms/registration.

Here is the link to the Office of Disability Service: https://success.camden.rutgers.edu/disability-services

Rutgers University
The State University of New Jersey
Department of Economics - CCAS
MS in Data Science Program
Spring 2026

Course Outline

Topic 0	What is Data Mining & Machine Learning
Topic 1	Introduction to Computing and Modeling Using R
Topic 2	The Familiar Linear Regression Model
Topic 3	Naïve Bayes
Topic 4	K-Nearest Neighbors (KNN)
Midterm Exam 1	Date: TBA
Topic 5	Decision Trees & Ensemble methods using Random Forest
Topic 6	Logistic Regression Model
Topic 7	Linear Model Selection and Regularization
Topic 8	Support Vector Machines (SVM)
Midterm Exam 2	Date: TBA
Topic 9	Dimension Reduction (focus on PCA) & Clustering PCA: Principal Components Analysis
Topic 10	Deep Learning: Introduction to Neural Networks (RNN, CNN, and Transformer)
Topic 11	Imbalanced Data
Topic 12	Association Rules (a.k.a. Market Basket Analysis)
Final Exam/Project	6:00-9:00 pm, Tuesday, May 12, 2026.