MET CS 699 Data Mining
2026 Spring 1 Online Course Syllabus

Instructor

Jae Young Lee, Ph.D., jaeylee@bu.edu

Course Duration

Start: January 13, 2026
End: March 2, 2026

Course credits

4 credits

Course Description

Data mining is one of the most important components in the knowledge discovery process. The course surveys various data mining applications, methodologies, techniques, and models. Topics include classification, numeric prediction, association rules, sequential pattern mining, clustering, and time series analysis. You will implement algorithms in the R language.

Prerequisites: CS 669 and CS 546 or instructor's consent.

Course Overview

- Overview of the Machine Learning Process
- Data Exploration and Visualization
- Data Summaries and Dimension Reduction
- Ethics in Machine Learning
- Prediction and Classification
- Association Rule Mining
- Clustering
- Time Series Analysis
- Case Studies

Course Learning Objectives

By successfully completing this course, you will be able to do the following:

1. Describe what Data Mining is and where it is applicable.
2. Apply a range of Data Mining techniques.
3. Apply practical examples using the R language.

Course Outline

- **Readings** – Each module has both textbook readings and online lectures. Your professor may suggest additional readings during the running of the course.
- **Discussion** – There are group-level threaded discussions for each module. These discussions are moderated by your facilitator. Postings for each discussion should be completed by the assigned due dates. There are also course-level general discussions boards, which are not graded, for you to use to discuss any issues with your classmates. Please see the Discussion Module on the home page for more details.
- **Assignment** – There are assignments that are due throughout the course. They are accessed from the Assignments menu.
- **Assessments** – Weekly quizzes and a final exam are also listed in the course calendar and accessed from the Assessments menu.

Module 1 – Introduction and Data Basics

- Basics of Data Mining
- Pre-processing and Data Cleaning
- Ethics in Data Science
- Data visualization
- Feature reduction

Module 2 – Prediction and Classification – 1

- Evaluating Model Performance
- Multiple Linear Regression
- Logistic Regression
- *k*-Nearest Neighbors

Module 3 – Prediction and Classification – 2

- Naïve Bayes
- Decision Trees
- Regression Trees

Module 4 – Prediction and Classification – 3

- Parameter Tuning
- Support Vector Machine
- Neural Network
- Discriminant Analysis
- Ensemble Methods

Module 5 – Association Analysis

- A/B Testing
- Uplifting

- Association Analysis
- Association Rule Evaluation
- Collaborative Filtering

Module 6 – Clustering and Time Series Analysis

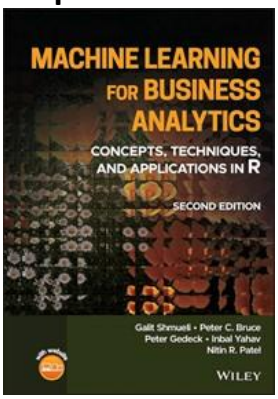
- Clustering
- Introduction to Time Series

Module 7 – Prepare for and take the final exam

You will prepare for and take the proctored final exam.
The course will remain open two weeks after the final exam, so that you can continue discussions and ask any questions about database technology, your grades, or the course. This is also a time when we enter into a dialog with you where we endeavor to learn how we can modify the course so that it better meets students' needs.

Course Materials

Required Book



Shmueli, G., Bruce, P.C., Gedeck, P., Yahav, I., & Patel. N.R. (2023). ***Machine learning for business analytics: concepts, techniques, and applications in R*** **(2nd edition)**

Publisher: Wiley

ISBN-13: 978-1119835172

How to Purchase and Access eBook through First Day™ from Your Course

To enhance your learning experience and simplify access to the right materials for your class, your course materials have been integrated directly into your course.

Benefits of This Program

- Exclusive preferred pricing
- Guaranteed the right materials
- Single Sign-On
- Ready to go on day one
- Course materials charge will be placed on your student account
- Option to Opt-Out on the first day of class.

Accessing Your Materials

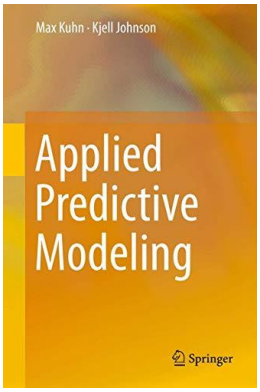
- Log in to your course site on the course launch day after 10 AM.
- On the homepage, click "**Course Materials (B&N First Day)**".

Notes:
Boston University will bill you at the discounted price as a First Day course material charge on your student account later in the semester.

Please be advised it is NOT recommended that you Opt-Out, as these materials are required in the course to complete the course. You can choose to Opt-Out on the first day of class, but you will be responsible for purchasing your course materials at the Opt-Out price.

For more information and FAQs, check Customer Care.

Recommended Book (not required)



Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling*.

Publisher: Springer

ISBN-13: 978-1461468486

This text is for reference only and is **not required** to be purchased for this course.

An e-book is available at Vitalsource.com and through Amazon.

Supplemental Material

Download the following zip file containing the supplementary course materials to your local hard disk from the course site, under "Syllabus and Resources": Supplements. This zip file provides the data sets used in the module readings in case you'd like to try those examples as part of your weekly reading.

Required Software

- R
- RStudio

Study Guide and Timeline of Deliverables

This course starts on a Tuesday. The modules in this course run from Tuesday to Monday.

Module 1 Study Guide and Deliverables

(January 13 – January 19, 2026)

Module Theme and Topics:

Overview of the Machine Learning Process, Data Exploration and Visualization, Data Summaries and Dimension Reduction, and Ethics in Machine Learning

Required Readings:

- Module 1 online content
- *Machine Learning for Business Analytics* (2nd edition):
  - Chapter 1 – Introduction
  - Chapter 2 – Overview of the Machine Learning Process (skip Sections 2.5, 2.7 and 2.8)
  - Chapter 3 – Data Visualization (skip Section 3.5)
  - Chapter 4 – Dimension Reduction (Sections 4.8, 4.9, and 4.10 are optional)
  - Chapter 22 – Responsible Data Science (Optional)

Discussions:

- Discussion 1 postings end **Tuesday, January 20, at 6:00 AM ET**

Assignments:

- Assignment 1 due **Tuesday, January 20, at 6:00 AM ET**

Assessments:

- Quiz 1 due on **Tuesday, January 20 at 6:00 AM ET**

Live Classrooms:

- **Tuesday, January 13th from 8 PM to 10 PM ET** and **Friday, January 16th from 8 PM – 9 PM ET** (Live sessions will be recorded) **(2 sessions per week)**
- Facilitator: **7 PM to 8 PM ET Wednesday, January 14th**

Module 2 Study Guide and Deliverables

(January 20 – January 26, 2026)

Module Theme and Topics:

Prediction and Classification – 1

- Evaluating Model Performance
- Multiple Linear Regression
- Logistic Regression
- *k*-Nearest Neighbors

Required Readings:

- Module 2 online content
- *Machine Learning for Business Analytics* (2nd edition):
    - Chapter 2 – Overview of the Machine Learning Process: Section 2.5 (Predictive Power and Overfitting)
    - Chapter 5 – Evaluating Predictive Performance (skip Section 5.4)
    - Chapter 6 – Multiple Linear Regression
    - Chapter 7 – k-Nearest Neighbors
    - Chapter 10 – Logistic Regression
    - Chapter 16 – Cluster Analysis: Section 16.2 (Measuring Distance Between Two Records)

Discussions:

- Discussion 2 postings end **Tuesday, January 27, at 6:00 AM ET**

Assignments:

- Assignment 2 due **Tuesday, January 27, at 6:00 AM ET**

Assessments:

- Quiz 2 due on **Tuesday, January 27 at 6:00 AM ET**

Live Classrooms:

- **Tuesday, January 20th from 8 PM to 10 PM ET** and **Friday, January 23rd from 8 PM – 9 PM ET** (Live sessions will be recorded) **(2 sessions per week)**
- Facilitator: **7 PM to 8 PM ET Wednesday, January 21st**

Module 3 Study Guide and Deliverables

(January 27– February 2, 2026)

Module Theme and Topics:

Prediction and Classification – 2

- Naïve Bayes
- Decision Trees
- Regression Trees

Readings:

- Module 3 online content
- *Machine Learning for Business Analytics* (2nd edition):
    - Chapter 8 – The Naïve Bayes Classifier
    - Chapter 9 – Classification and Regression Trees (skip Section 9.9)

Discussions:

- Discussion 3 postings end **Tuesday, February 3, at 6:00 AM ET**

Assignments:

- Assignment 3 due **Tuesday, February 3, at 6:00 AM ET**

Assessments:

- Quiz 3 due on **Tuesday, February 3 at 6:00 AM ET**

Live Classrooms:

- **Tuesday, January 27th from 8 PM to 10 PM ET** and **Friday, January 30th from 8 PM – 9 PM ET** (Live sessions will be recorded) **(2 sessions per week)**
- Facilitator: **7 PM to 8 PM ET Wednesday, January 28th**

Module 4 Study Guide and Deliverables

(February 3 – February 9, 2026)

Module Theme and Topics:

Prediction and Classification – 3

- Parameter Tuning
- Support Vector Machine
- Neural Network
- Discriminant Analysis
- Ensemble Methods

Readings:

- Module 4 online content
- *Machine Learning for Business Analytics* (2nd edition):
    - Chapter 11 – Neural Nets (skip Sections 11.6 and 11.7)
    - Chapter 12 – Discriminant Analysis (skip Sections 12.5 – 12.8)
    - Chapter 13 – Ensemble Methods: Section 13.1 (Ensembles)
    - Chapter 9 – Classification and Regression Trees: Section 9.9 (Improved Prediction: Random Forest and Boosted Trees)

Discussions:

- Discussion 4 postings end **Tuesday, February 10, at 6:00 AM ET**

Assignments:

- Assignment 4 due **Tuesday, February 10, at 6:00 AM ET**

Assessments:

- Quiz 4 due on **Tuesday, February 10 at 6:00 AM ET**

Live Classrooms:

- **Tuesday, February 3rd from 8 PM to 10 PM ET** and **Friday, February 6th from 8 PM – 9 PM ET** (Live sessions will be recorded) **(2 sessions per week)**
- Facilitator: **7 PM to 8 PM ET Wednesday, February 4th**

Module 5 Study Guide and Deliverables

(February 10 – February 16, 2026)

Module Theme and Topics:

Association Analysis

- A/B Testing
- Uplifting
- Association Analysis
- Association Rule Evaluation
- Collaborative Filtering

Readings:

- Module 5 online content
- *Machine Learning for Business Analytics (2nd edition)*:
    - Chapter 14 – Interventions (skip Section 14.3)
    - Chapter 15 – Association Rules and Collaborative Filtering

Discussions:

- Discussion 5 postings end **Tuesday, February 17, at 6:00 AM ET**

Assignments:

- Assignment 5 due **Tuesday, February 17, at 6:00 AM ET**

Assessments:

- Quiz 5 due on **Tuesday, February 17, at 6:00 AM ET**

Live Classrooms:

- **Tuesday, February 10th from 8 PM to 10 PM ET** and **Friday, February 13th from 8 PM – 9 PM ET** (Live sessions will be recorded) **(2 sessions per week)**
- Facilitator: **7 PM to 8 PM ET Wednesday, February 11th**

Module 6 Study Guide and Deliverables

(February 17 – February 23, 2025)

Module Theme and Topics:

Clustering and Time Series Analysis

- Clustering
- Introduction to Time Series

Readings:

- Module 6 online content
- *Machine Learning for Business Analytics (2nd edition)*:

- Chapter 16 – Cluster Analysis (Skip Section 16.2)
- Chapter 17 – Handling Time Series
- Chapter 18 – Regression-Based Forecasting
- Chapter 19 – Smoothing and Deep Learning Methods for Forecasting (Skip Section 19.5)

Discussions:

- Discussion 6 postings end **Tuesday, February 24, at 6:00 AM ET**

Assignments:

- Assignment 6 due **Tuesday, February 24, at 6:00 AM ET**

Assessments:

- No quiz

Live Classrooms:

- **Tuesday, February 17th from 8 PM to 10 PM ET** and **Friday, February 20th from 8 PM – 9 PM ET** (Live sessions will be recorded) **(2 sessions per week)**
- Facilitator: **7 PM to 8 PM ET Wednesday, February 18th**

Course Evaluation:

Please complete the course evaluation once you receive an email or Blackboard notification indicating the evaluation is open. Your feedback is important to MET, as it helps us make improvements to the program and the course for future students.

Final Exam Details

The Final Exam is an online proctored exam available from **Wednesday, February 25, 2026, at 6:00 AM ET to Saturday, February 28, 2026, at 11:59 PM ET**.

The Computer Science department requires that all final exams be administered using an online proctoring service that you will access via your course in Blackboard. In order to take the exam, you are required to have a working webcam and computer that meets the exam proctoring service's system requirements. A detailed list of those requirements can be found on the How to Schedule page. Detailed instructions regarding your proctored exam will be forthcoming from the Assessment Administrator. You will be responsible for scheduling your own appointment within the defined exam window.

Please note that student activity during the final exam is monitored and recorded by the proctoring vendor. Accessing any unauthorized material during the final exam is a major violation of the course policy and can result in serious academic disciplinary actions.

Final Exam Duration: 3 hours

Grading Information

Please check the Study Guide in the syllabus for Live Classroom dates and specific due dates for assignments and assessments.

Grading Structure and Distribution

The grade for the course is determined by the following:

Graded Items	Percentage of Grade
Assignments	30%
Discussions	10%
Quizzes	20%
Proctored Final Exam	40%
Total	100%

94 ≤ G	A
90 ≤ G < 94	A−
87 ≤ G < 90	B+
83 ≤ G < 87	B
80 ≤ G < 83	B−
77 ≤ G < 80	C+
73 ≤ G < 77	C
70 ≤ G < 73	C−
60 ≤ G < 70	D
G < 60	F

Graded Items

- **Discussions:** Graded discussion periods are held Day 1 (Tuesday) until 6:00 AM ET on Day 1 of the following week. You're certainly welcome to continue a discussion past the grading period, but that additional posted material will not affect your discussion grade. The discussion grading rubric below is the guide we use to evaluate your discussion contributions.

  Students will be participating in discussions that will be graded on a 100-point scale. For more information on discussion grading, check out the Discussion Rubric.

Criteria	65–69	70–79	80–89	90–94	95–100
Participation	Very limited participation	Participation generally	Reasonably useful relevant	Frequently relevant and	Continually relevant and

		lacks frequency or relevance	participation during the discussion period	consistent participation throughout the discussion period	consistent participation throughout the discussion period
Community	Mostly indifferent to discussion	Little effort to keep discussions going or provide help	Reasonable effort to respond thoughtfully, provide help, and/or keep discussions going	Often responds thoughtfully in a way that frequently keeps discussions going and provides help	Continually responds thoughtfully in a way that consistently keeps discussions going and provides help
Content	No useful, on-topic, or interesting information, ideas, or analysis	Hardly any useful, on-topic, or interesting information, ideas, or analysis	Reasonably useful, on-topic, and interesting information, ideas, and/or analysis	Frequently useful, on-topic, and interesting information, ideas, and analysis	Exceptionally useful, on-topic, and interesting information, ideas, and analysis
Reflection and Synthesis	No significant effort to clarify, summarize, or synthesize topics raised in discussions			Contributes to group's effort to clarify, summarize, or synthesize topics raised in discussions	Leads group's effort to clarify, summarize, or synthesize topics raised in discussions

- **Assignments:** There will be six assignments, one per every module. The purpose of assignments is to help students apply what they have learned in the module to application problems. Check the Study Guide for the specific due date.

- **Quizzes:** There will be five quizzes, one every week except in Week 6 and Week 7. The primary goal of quizzes is to help students keep current with the course material.

- **Proctored Final Exam:** There will be a proctored Final Exam in this course using a proctor service. Detailed instructions regarding your proctored exam will be forthcoming from the Assessment Administrator. You will be responsible for scheduling your own appointment.

Ungraded Items

- **Ungraded Discussion Forums:** There are ungraded discussion forums throughout the course. You are encouraged to share your knowledge and learn from your peers.

Discussions forums are provided for your benefit. Some discussion forums involve the teaching team members; others are among students.

- **Live Classroom Sessions:** Live Classroom sessions will be offered during this course in Modules 1 through 6.
    - Your participation, while not mandatory, will be valuable to you and the class. To participate in the Live Classroom discussion, you will need to go to the "Live Classroom/Offices" area.
    - Live Classroom sessions will be recorded and archived for further viewing. You can go to the "Live Classroom Recordings" area to view the recordings.

Late Work Expectations

For the assignments and discussions due on Tuesdays at 6:00 AM ET, there will be a 5% per day penalty (5% if submitted by Wednesday at 6:00 AM ET, 10% if submitted by Thursday at 6:00 AM ET, and 15% if submitted by Friday at 6:00 AM ET). Work submitted after Friday at 6:00 AM ET will receive no credit.

Quizzes should be completed well in advance of the assignment and discussion deadline as they are designed to prepare you to do the assignment. As such, quizzes are not accepted late.

If you believe you have a true emergency and require a deviation from this policy, contact the professor to make alternative arrangements. This could involve you receiving a different assignment from the rest of the class.

Policy for the Use of Generative AI

You may use Generative AI algorithms on assignments and quizzes in this class unless a problem directly prohibits it. On discussions, Generative AI may be used to help you research what to write and to improve work you have written. Generative AI algorithms are not permitted on the final exam.

You are responsible for knowing whether the information provided by the Generative AI is correct. If you are allowing a Generative AI to incorporate work from others, you are responsible for producing valid citations and references. Generative AI algorithms may not be aware of all the details presented in our lectures, online readings, and in your textbook.

You should also follow all student guidelines in the Generative AI Assistance (GAIA) policy.

Course Policies and Academic Conduct

Academic Integrity: Plagiarism is the passing off of another's words or ideas as your own, and it is a serious academic offense. Plagiarism and cheating also defeat the purpose of getting an education. Plagiarism and cheating cases will be handled in accordance with the disciplinary

procedures described in the College of Arts and Sciences Academic Conduct Code. You are expected to know and abide by the code, which can be read online: Academic Conduct Code. Penalties range from failing an assignment or course (first offense) to suspension or expulsion from BU. If in doubt, cite your source. If you have any questions about academic integrity, please ask your instructor.

Incidents of academic misconduct will be reported to the Academic Conduct Committee (ACC). The ACC may suspend/expel students found guilty of misconduct.

Important Message on Final Exams

Dear Boston University Computer Science Online Student,

As part of our ongoing efforts to maintain the high academic standard of all Boston University programs, including our online MSCIS degree program, the Computer Science Department at Boston University's Metropolitan College requires that each of the online courses includes a proctored final examination.

By requiring proctored finals, we are ensuring the excellence and fairness of our program. The final exam is administered online.

Specific information regarding final-exam scheduling will be provided approximately two weeks into the course. This early notification is being given so that you will have enough time to plan for where you will take the final exam.

I know that you recognize the value of your Boston University degree and that you will support the efforts of the University to maintain the highest standards in our online degree program.

Thank you very much for your support with this important issue.

Regards,

Professor Lou Chitkushev, Ph.D.
Associate Dean for Academic Affairs
Boston University Metropolitan College

Disability and Access Services

In accordance with university policy, every effort will be made to accommodate students with respect to speech, hearing, vision, or other disabilities. Any student who may need an accommodation for a documented disability should contact Disability and Access Services at 617-353-3658 or at access@bu.edu for review and approval of accommodation requests.

Once a student receives their accommodation letter, they must send it to their instructor and/or facilitator each semester. They must also send a copy to their Faculty & Student Support Administrator, who may need to update the course settings to ensure accommodation is in place. Accommodation cannot be implemented if the students do not send their letters.