Dalhousie University

Department of Mathematics and Statistics

STAT 2450: Introduction to Data Mining with R

Fall 2023

Dalhousie University acknowledges that we are in Mi'kma'ki, the ancestral and unceded territory of the Mi'kmaq People and pays respect to the Indigenous knowledges held by the Mi'kmaq People, and to the wisdom of their Elders past and present. The Mi'kmaq People signed Peace and Friendship Treaties with the Crown, and section 35 of the Constitution Act, 1982 recognizes and affirms Aboriginal and Treaty rights. We are all Treaty people.

Dalhousie University also acknowledges the histories, contributions, and legacies of African Nova Scotians, who have been here for over 400 years.

Instructor	Orla Murphy
Office	Chase 223 and virtual using Collaborate Ultra
e-mail	orla.murphy@dal.ca Please only send emails from your Dalhousie email account with subject line STAT 2450.
Time of Lectures	Tuesday and Thursday: 1:05pm – 2:25 pm
Lecture Location	In-person in CHEM 223
Course Delivery	Lectures are in-person and will not be recorded.
Course Website	Brightspace
Office Hours	Tuesdays 2:30–3:30pm and Thursdays 2:30–3:30pm
Course Prerequisites	MATH 1000.03 or MATH 1215.03 and either (STAT 1060.03 or MATH 1060.03) or (STAT 2060.03 or MATH 2060.03)
Required Text	Introduction to Statistical Learning with Applications in R by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani Available at the bookstore or available for free online at: https://hastie.su.domains/pub.htm
Software	R which can be downloaded from http://www.r-project.org/ Rstudio with the R Markdown add-on is required to organize assignments. It is available at https://www.rstudio.com/

Course Description

From the Academic Calendar:
This course provided as introduction to data mining and R programming, suited for science students. Data mining methods include a vast set of tools developed in different areas for identifying the patterns in data. Students will learn programming methods for manipulating and exploring data through learning the basic ideas of some clustering, regression and classification methods. No prior programming knowledge is assumed.

Learning Outcomes

- Students will learn basic R programming skills including using R as a calculator, handling data, user defined functions, installing and using R packages.
- Students will develop a foundation in statistical learning methods.

Outline:

- Introduction to R and Rstudio
- Data types, data frames, importing and exporting data
- Controlling the program flow in R
- User-defined functions in R
- Built-in functions for classical probability distributions
- Exploratory analysis in R, plots and graphs
- Simulation of coverage of t-confidence interval
- Introduction to linear and polynomial regression
- Introduction to the bootstrap method
- Introduction to tree-based regression
- Model validation techniques
- Introduction to Random Forests

Note that topics may be altered as term progresses.

Evaluations:

Assessment	Weight (% of final grade)
Assignments	40%
Quizzes	25%
Final exam	30%*
Participation	5%

*You need to obtain a grade of at least 50% on your final exam in order to pass the course, regardless of your grades on the remaining course components.

Conversion of numerical grades to Final Letter Grades follows the Dalhousie Common Grade Scale

A+	(90-100)	B+	(77-79)	C+	(65-69)	D	(50-54)
A	(85-89)	B	(73-76)	C	(60-64)	F	(<50)
A-	(80-84)	B-	(70-72)	C-	(55-59)		

Assignments:

- **Late assignments will not be accepted unless a prior extension has been approved.** Extensions may be allowed under extreme circumstances. The student must request an extension *prior* to the deadline. Note that all extensions are at the instructor's discretion and that repeat requests for extensions may not be granted.
- Although students may discuss individual assignments in groups, the submitted work must be your own. Any identical submitted assignments will receive a grade of 0 and, according to university regulations, will be reported to the Academic Integrity Office.
- Assignments will be due roughly every 1-2 weeks. The number of assignments will depend on course progress but is tentatively set to be 7 assignments. Always make sure to check the deadline on the posted assignment.

Quizzes:

- Students will be given a quiz after each main topic (tentative number of quizzes is 7). Quiz dates will be announced in lecture and may be announced up to the lecture before the quiz. Quizzes may be given in person or online on Brightspace (instructor preference).
- Quizzes must be written independently.

Final exam:
The final exam will take place during the December examination period (scheduled by Registrar) and will cover the entire course. Students should check the official exam schedule for the time and location. **You need to obtain at least a 50% on your final exam in order to pass the course, regardless of your grades on the remaining course components.**

Participation:
The participation grade will be based on submission of in-class exercises during class time. These exercises will usually be performed in pairs or small groups. For this reason, students are strongly encouraged to bring a device with access to `R` and Brightspace whenever possible. In the event that students cannot bring an appropriate device, there will be the option of submitting the exercise on paper.

Grading:

- Please note that there will be no grade modifications beyond 2 weeks after the returned assessment. **Grading inquiries accepted only by email.**
- Assignments will be submitted online using Crowdmark, an online grading platform. Instructions for student submission of assignments can be found on the course page, in the Assignment section. There will be a practice submission available before your first assessment. Crowdmark will also be used to grade quizzes. You will be able to access Crowdmark through your course page to submit an assessment or view your grades. Support for using Crowdmark can be found at this `link`.

Student absence

Students experiencing a short-term absence that is no longer than three consecutive days and results in a missed assessment *must*:

- Notify the instructor by email prior to the academic deadline or scheduled time, and
- Download, complete, and submit the Student Declaration form within 3 days of the assessment deadline. Submission must be made on Brightspace.

See: Student Declaration Information for form. No make-up tests, quizzes, or assignments will be offered. **Only one such form submission per student will be accepted in the course.**

In cases where a student is unable to attend a lecture in-person, the student should contact the instructor by email (orla.murphy@dal.ca) in advance of the lecture and arrangements may be made to record the lecture.

Recognition of Mi'kmaq Territory
Dalhousie University would like to acknowledge that the University is on Traditional Mi'kmaq Territory. The Elders in Residence program provides students with access to First Nations elders for guidance, counsel, and support. Visit or e-mail the Indigenous Student Centre at 1321 Edward St or elders@dal.ca. Additional information regarding the Indigenous Student Centre can be found at: https://www.dal.ca/campus_life/communities/indigenous.html

Internationalization
At Dalhousie, 'thinking and acting globally' enhances the quality and impact of education, supporting learning that is "interdisciplinary, cross-cultural, global in reach, and orientated toward solving problems that extend across national borders." Additional internationalization information can be found at: https://www.dal.ca/about-dal/internationalization.html

Academic Integrity
At Dalhousie University, we are guided in all of our work by the values of academic integrity: honesty, trust, fairness, responsibility and respect (The Center for Academic Integrity, Duke University, 1999). As a student, you are required to demonstrate these values in all of the work you do. The University provides policies and procedures that every member of the university community is required to follow to ensure academic integrity.
Information: Academic Integrity

Accessibility
The Student Accessibility Centre is Dalhousie's centre of expertise for matters related to student accessibility and accommodation. If there are aspects of the design, instruction, and/or experiences within this course (online or in-person) that result in barriers to your inclusion, please contact the Student Accessibility Centre:
Academic Support

Student Code of Conduct
Everyone at Dalhousie is expected to treat others with dignity and respect. The Code of Student Conduct allows Dalhousie to take disciplinary action if students don't follow this community expectation. When appropriate, violations of the code can be resolved in a reasonable and informal manner—perhaps through a restorative justice process. If an informal resolution can't be reached, or would be inappropriate, procedures exist for formal dispute resolution.
The full Code of Student Conduct can be found at: Student Code of Conduct

Conduct in the Classroom – Culture of Respect
Substantial and constructive dialogue on challenging issues is an important part of academic inquiry and exchange. It requires willingness to listen and tolerance of opposing points of view. Consideration of individual differences and alternative viewpoints is required of all class members, towards each other, towards instructors, and towards guest speakers. While expressions of differing perspectives are welcome and encouraged, the words and language used should remain within acceptable bounds of civility and respect.

Diversity and Inclusion – Culture of Respect
Every person at Dalhousie has a right to be respected and safe. We believe inclusiveness is fundamental to education. We stand for equality. Dalhousie is strengthened in our diversity. We are a respectful and inclusive community. We are committed to being a place where everyone feels welcome and supported, which is why our Strategic Direction prioritizes fostering a culture of diversity and inclusiveness (Strategic Priority 5.2). Additional diversity and inclusion information can be found at:
Diversity and Inclusion

Fair Dealing Policy
The Dalhousie University Fair Dealing Policy provides guidance for the limited use of copyright protected material without the risk of infringement and without having to seek the permission of copyright owners. It is intended to provide a balance between the rights of creators and the rights of users at Dalhousie. Additional information regarding the Fair Dealing Policy can be found at:https://www.dal.ca/dept/university_secretariat/policies/academic/fair-dealing-policy-.html

Originality Checking Software
The course instructor may use Dalhousie's approved originality checking software and Google to check the originality of any work submitted for credit, in accordance with the Student Submission of Assignments and Use of Originality Checking Software Policy. Students are free, without penalty of grade, to choose an alternative method of attesting to the authenticity of their work and must inform the instructor no later than the last day to add/drop classes of their intent to choose an alternate method. Additional information regarding Originality Checking Software can be found at: Originality Checking Software

Student Use of Course Materials
Course materials are designed for use as part of this course at Dalhousie University and are the property of the instructor unless otherwise stated. Third party copyrighted materials (such as books, journal articles, music, videos, etc.) have either been licensed for use in this course or fall under an exception or limitation in Canadian Copyright law. Copying this course material for distribution (e.g. uploading to a commercial third-party website) may lead to a violation of Copyright law.

DALHOUSIE UNIVERSITY

Student Resources and Support

Important Dates in the Academic Year (including add/drop dates): Important dates

University grading practices: University Grading Practices

Advising
General advising
Science Program Advisors
Indigenous Student Center
Black Students Advising Centre
International Centre

Academic supports
Library
Writing Centre
Studying for Success
Copyright Office
Fair Dealing Guidelines

Other supports and services
Student Health & Wellness Centre
Student Advocacy
Ombudsperson

Safety
Biosafety
Chemical Safety
Radiation Safety

Scent-Free Program

Dalhousie COVID-19 information and updates

Important Notice:
The instructor and university reserve the right to modify elements of the course during the term. If a modification becomes necessary, reasonable notice and communication with the students will be given with explanation and the opportunity to comment on changes. It is the responsibility of the student to check their Dalhousie email and the course website regularly during the term and to note any changes.