

# Data Science 6601

## Practical Machine Learning

### Winter 2022



**Instructor:** Matthew Hamilton  
**Lectures:** TR 10:30-11:45am on Brightspace Online Rooms. In person: EN-1002.  
**Office Hours:** Available by appointment through email (or Discord) for video call or in-person at EN-2029  
**E-mail:** [mhamilton@mun.ca](mailto:mhamilton@mun.ca) (\*\*please use this email to contact me, not D2L\*\*)  
**Phone:** (709) 864-4560

Note that I **DO NOT** check email in Brightspace (D2L). Please send any emails to my [mhamilton@mun.ca](mailto:mhamilton@mun.ca) address, rather than my D2L address. Also, please include **[DATA SCIENCE]** in the subject line. Official email correspondence within the university must be via a valid [@mun.ca](mailto:@mun.ca) email account. *Every effort will be made to respond to emails within 24h, with the exceptions of evenings, weekends and holidays.*

**Course Content:** <https://online.mun.ca/>

**Prerequisites:** None.

#### Course Objectives:

This course provides a practical introduction to machine learning. The course will mostly focus on applying supervised learning methods to a variety of datasets using R and Python. Model selection and assessment will be discussed and applied.

Topics covered:

#### 1. Introduction to Machine Learning (ML)

- (a) Definition
- (b) Types of ML
- (c) Supervised ML
  - i. Classification and regression
  - ii. Generalization, underfitting, overfitting, bias and variance

#### 2. Introduction to scikit-learn

#### 3. Introduction to caret package

#### 4. Exploring, understanding, and representing data

- (a) Data exploration (recap)
- (b) Data transformations
- (c) Data encoding
- (d) Feature selection

#### 5. Model performance assessment

- (a) Performance metrics
- (b) Cross-validation

## 6. Supervised ML algorithms

- (a) Nearest Neighbours
- (b) Linear Models
- (c) Naive Bayes Classifiers
- (d) Ensembles of Decision Trees
- (e) Support Vector Machines
- (f) Feed forward neural networks
- (g) Other models as appropriate

### Textbook and Resources:

An Introduction to Statistical Learning: with Application in R. G. James, D. Witten, T. Hastie and R. Tibshirani. Springer, 2013. (<https://www-bcf.usc.edu/~gareth/ISL/>).

Introduction to Machine Learning with Python: A Guide for Data Scientists. Sarah Guido, Andreas Muller. O'Reilly Media, 2016.

Machine Learning with R: Expert techniques for predictive modeling, 3rd Edition. Brett Lantz. Packt, 2019.

Additional materials assigned throughout the course.

### Evaluation:

	Value	Approximate due date
<b>Assignments</b>	<b>50%</b>	-
Assignment 1	10%	<b>September 29</b>
Assignment 2	10%	<b>October 6</b>
Assignment 3	10%	<b>October 20</b>
Assignment 4	10%	<b>November 3</b>
Assignment 5	10%	<b>November 17</b>
<b>Project</b>	<b>25%</b>	-
Proposal	2%	<b>October 13</b>
Content / presentation	15%	<b>November 17-24</b>
Final Report	8%	<b>December 1</b>
<b>Final Exam</b>	<b>25%</b>	<b>December</b>

### COVID-19/Health Related Course Information:

Welcome to the Fall 2022 term! This course is designed to be held in-person. Our class lectures have been carefully designed to emphasize safety while providing a rich learning experience for all students. Should health directives or the overall situation connected to COVID-19 change over the course of the term, a back-up plan for remote delivery is in place to ensure that the course will continue and to minimize disruption to the student experience.

To protect yourself and those around you, it is important to stay home if you feel unwell or if you are under quarantine, because you have potentially been exposed to the virus. Please keep me informed so we can work together to allow you to keep up with the course materials, should you need to miss classes. You will not be penalized if you need to stay home for quarantine. Memorial University has recognized the importance of academic leniency as we work to keep our campus safe for all.

There is nothing more important than your mental and physical health. Doctors' notes are not required for medical absences in this course. You are encouraged to seek appropriate medical attention from the Student Wellness and Counselling Centre. I am committed to working with students with pre-existing medical and mental health needs, as

well as new needs that may arise within the semester. I encourage you to reach out to the Blundon Centre as early as possible to discuss any adjustments you think may be necessary in this course. Let's explore the options to help you succeed, no matter what is going on.

In the event of a COVID-19 related disruption, **instruction will switch to remote delivery**. Some course topics may be cancelled depending on when the disruption occurs. Remote lectures for our class will include a mixture of synchronous and asynchronous delivery. If this transition occurs, we will update the course syllabus and post to Brightspace to announce the revised lecture schedule.

### **Important notes to read, understand and follow for this course:**

#### **1. General Notes**

- Students are expected to adhere to those principles which constitute proper academic conduct. A student has the responsibility to know which actions, as described under Academic Offences in the University Regulations, could be construed as dishonest or improper. Students found guilty of an academic offence may be subject to a number of penalties commensurate with the offence including reprimand, reduction of grade, probation, suspension or expulsion from the University. For more information regarding this policy, students should refer to the University Regulations for Academic Misconduct (Section 6.12) in the University Calendar.
- Further course information (e.g., assignments, lab exercises, quizzes/exams, announcements, grades, etc) can be found in Brightspace (D2L).
- No late submissions will be accepted. If, for special circumstances (such as medical or bereavement), you miss a deadline for a grade item, notify your instructor as soon as possible and not later than 48 hours, providing an explanation for the missed work and any related documentation (if documentation is required). Failure to do this can result in a mark of 0% for that work. Please refer to the **current University policy** regarding medical notes and the information to be in them. For more information, please see the University Calendar - University Regulations - General Academic Regulations (Undergraduate) - **6.7.5 Exemptions from Parts of the Evaluation** and **6.15 Appeal of Decisions** or consult the Registrar's Office. If your reasons for the missed work are acceptable, then your instructor will provide details of any alternate evaluation scheme.
- The lectures and all material delivered or provided in this course, including any visual or audio recording thereof, are subject to copyright. It is prohibited to record or copy by any means, in any format, openly or surreptitiously, in whole or in part, in the absence of express written permission from the course instructor (M. Hamilton) any of the lectures, materials provided or published in any form during or from the course.
- The use of visual and/or audio recording in the classroom is prohibited, excluding any provisions made for students with special needs.
- The use of personal electronic devices or programs during class (e.g., phones, laptops, tablets) is expected, but please make sure to turn off sound volume (e.g. ringers) to minimize any disruptions.
- In line with the Newfoundland and Labrador Human Rights Act, we are committed to student equity and the provision of a safe learning environment regardless of race, colour, nationality, ethnic origin, social origin, religious creed, religion, age, disability, disfigurement, sex (including pregnancy), sexual orientation, gender identity, gender expression, marital status, family status, source of income or political opinion.
- We respectfully acknowledge the territory in which we gather as the ancestral homelands of the Beothuk, and the island of Newfoundland as the ancestral homelands of the Mi'kmaq and Beothuk. We would also like to recognize the Inuit of Nunatsiavut and NunatuKavut and the Innu of Nitassinan, and their ancestors, as the original people of Labrador. We strive for respectful relationships with all the peoples of this province as we search for collective healing and true reconciliation and honour this beautiful land together.
- Attribution of sources: Students are expected to demonstrate the ability to complete the assigned work independently. While you are encouraged to use all resources at your disposal, you are representing all of your submission as your independent work unless you indicate otherwise. Therefore it is important include attribution of all your sources as part of any work submission. Failure to include attribution of your source material may constitute plagiarism and will be subject to penalties as academic misconduct. Collaboration with a classmate may result in submissions that are considered to be copied rather than independent. You may avoid penalties for academic misconduct by including appropriate attribution or explanation of such collaboration as part of your submission. Marks may still reflect a failure to submit independent work.

## 2. Assignments

- Assignments are due at **11:59 p.m.** on the specified date, electronically through the Brightspace (D2L) Assignment dropbox tool. (See table above.) No late assignments will be accepted. Be aware that the files you submit for evaluation should be uploaded on or before the due date and much before the cut off time, **11:59 p.m. Newfoundland Time**. Even if you are late by a few seconds you will not be allowed to submit your work; hence you should try to upload the files at least 15 minutes before the cut off time since your system clock is not synchronized with the CITL's system clock and the cutoff time is based on CITL's system clock. Please note that if your files have been correctly uploaded, you will get a confirmation receipt from the **Dropbox** tool. If you do not receive this receipt, please contact the CITL Support team (<https://www.citl.mun.ca/support/>). **\*\*\*It is your responsibility to make sure that the CORRECT files are actually uploaded, so please DO check for the confirmation that your files have been uploaded.\*\*\***
- **Assignments are to be done individually. If cheating/plagiarism is detected for an assignment, the assignment will be given 0%. The department head will also be notified of any academic misconduct.**

## 3. Exams

- **Final Exam**
  - The final exam will be scheduled by the **Registrar's Office** for some time during the regular exam period. For the exact final exam date and time, check on **Memorial Self-Service**.
  - **Final exams are to be done individually. If academic misconduct is detected for a final exam, the case will be reported to the department head, who will forward it to the Senate Committee on Undergraduate Studies for investigation and appropriate action.**
- **No supplementary examination will be given for this course.**

## 4. Student support

- Memorial University of Newfoundland is committed to supporting inclusive education based on the principles of equity, accessibility and collaboration. Accommodations are provided within the scope of the University Policies for the Accommodations for Students with Disabilities ([www.mun.ca/policy/site/policy.php?id=239](http://www.mun.ca/policy/site/policy.php?id=239)). Students who may need an academic accommodation are asked to initiate the request with the Glenn Roy Blundon Centre at the earliest opportunity ([www.mun.ca/blundon](http://www.mun.ca/blundon)).
- Help with any Brightspace (D2L) technical issues is available by contacting the **CITL Support Centre** via phone (1-866-435-1396, or locally at 709-864-8700), online chat (when available) or via their ticketing system. This contact information is available at: <https://www.citl.mun.ca/support/>
- Note that, while the due times for submissions are at **11:59pm Newfoundland Time**, help may not be available after 4:30pm on the due date. The last few hours from 4:30pm – 11:59pm should be used to finalize your work and submit it. Any questions you may have about the assignments should be asked well in advance of the due date/time to allow time for help to be given.
- Please note that Memorial University offers a broad range of supports. For a more comprehensive list of student supports and resources, please check out the following links:
  - <https://www.mun.ca/munup/support/>
  - <https://www.mun.ca/main/students.php>
  - <https://www.munsu35.ca/resource-centres/>