CSCI 3151 Course Syllabus

Foundations of Machine Learning

January 2024

Dr. Sageev Oore

Instructor Information

Instructor: Sageev Oore **Email:** sageev@dal.ca **Class meeting time:** MW 8:35-9:55 in GB 127 **Tutorial meeting times:** F 8:35-9:55 in GB 127 **TA:** TBD **Course Rep:** TBD	**Office:** Microsoft Teams, private channel (Sageev Oore) **Office hours:** TBD **ONLINE COMPONENT:** Some classes and tutorial sessions may be conducted online. In that case you will receive notice at least one day in advance, and the session will then be held on the corresponding channel on Teams.

Important Dates *[link]*

Course Description & Prerequisites *[link to Academic Calendar]*

Learning Outcomes

- Acknowledge the Academic Integrity Policy.
- Apply optimization algorithms to minimize multivariate cost functions used in machine learning models.
    - Given a differentiable multivariate cost function for a regression problem, be able to compute partial derivatives as needed and implement gradient descent optimization
    - Be able to incorporate basic regularization terms into the optimization
- Demonstrate basic understanding of probability (e.g., Bayes, MLE) and information theory (e.g., entropy, mutual information).
    - Compute conditional, marginal distributions, entropy and mutual information, given sufficient information
- Use Python and NumPy to code linear algebra operations, and make effective use of vectorization.
- Demonstrate an understanding of the characteristics of supervised, unsupervised, and generative models.
- Understand the principles of traditional supervised machine learning methods (e.g., linear regression, SVM, decision trees, naive bayes) sufficiently well to be able to implement these methods in Python and Numpy, analyze their appropriateness in a variety of contexts, and be able to clearly explain the intuitions behind them.
- Implement traditional supervised machine learning methods (e.g., linear regression, SVM, decision trees, naive Bayes) using the Scikit-learn library.
- Understand overfitting, including how and when to look for it, and how to avoid it
- Use feature selection techniques to improve generalization
- Design evaluation methods for supervised machine learning methods

- Implement traditional unsupervised machine learning methods (e.g. k-means, HAC) in Python and/or using the Scikit-learn library.
- Design evaluation methods for unsupervised machine learning methods
- Understand how to set hyperparameters of machine learning algorithms using cross validation
- Design and implement neural networks using a specialized library (e.g. Keras)
- Demonstrate an understanding of the basics of text and image representation.
- Apply neural networks for advanced text and image processing.

Prerequisite Learning Outcomes **(of a student enrolling in this course)**

- Use basic data structures (linked lists, queues, stacks, trees) [CSCI 1101].
- Develop complete applications comprising multiple classes that integrate various object-oriented programming concepts [CSCI 1101].
- Describe the difference between data, information, and knowledge [CSCI 2141].
- Describe the purpose, function, and importance of a data dictionary, system catalog, and indexing [CSCI 2141].
- Select and use appropriate abstract data types, data structures, and algorithms to solve moderately complex problems [CSCI 2110].
- Familiar with Null Hypothesis, Binomial and Normal Distribution and its PDF, t-test, Random variables, and conditional probabilities [STAT 2060].
- Familiar with linear algebra: including vector and matrix operations, vector spaces, and inner product [MATH 1030].
- Familiar with derivatives of functions of one and multiple variables, the concept of gradient and directional derivative, unconstrained and constrained optimization [MATH 1000]
- See also the topics graph shown in class for more context.

Course Rationale

This Course introduces machine learning concepts and tools, and prepares students to take subsequent courses in machine learning, data science, and natural language processing. The course includes mathematics fundamentals for machine learning and uses Python libraries including numpy and SciKitLearn.

Format & Communication

- Classes will be presented in a mix of the following formats (depending on the topic):
- **(1) regular lecture:**
    - material will be presented in class, in-person
    - this may include in-class activities/exercises/quizzes, which may be worth some small part of your term mark (as described in the marking scheme)
    - **Collective Note-taking**: There will be a shared google doc(s) which will be maintained by the instructor and TAs but which will be created by the students, where the class will have the opportunity to make collective notes about the main topics discussed in class. The TAs and instructor will initiate this document to provide the class with a model of what it can contain, and then will transfer the responsibility to the class to enter the content. The TAs and instructor will continue to monitor this document, allowing for discussions and Q&A. This will also provide a review document for any tests/exams.
- **(2) flipped classroom model:**
    - **before the lecture:** Online videos and/or readings will be assigned before the class (usually by the end of the previous class). Students are expected to have completed any such assigned materials at their convenience before the class (asynchronously).

    - **lecture after being introduced to the topic**: These will typically begin with a short quiz (online or in-person TBA) to test the concepts in the assigned readings or videos. The instructor will then discuss the concepts in more depth and/or solve problems on the class content, for or in collaboration with the class (**either synchronously or asynchronously**). If a class meeting is held virtually or asynchronously, then it will be recorded and made available so that it is not mandatory to attend that class live. However, students not attending synchronously are required to watch the lecture and complete the pre-class quiz (if any) before the next class.
    - **after the lecture**: Students may be assigned further problems to solve, theoretical concepts to reflect on, or just work on their assignments. At this stage, students ask questions on the forum or individually.
    - **Collective note-taking:** The same document described above will also apply to material presented in the flipped classroom model. This will all be discussed in class and demonstrated in the first two weeks of class.
- **Tutorials** will be run synchronously. Tutorials will review background material required for the course, and discuss practice problems through Python notebooks that will help students do the experimental questions in the assignments.
- **Mix of Lectures and Tutorials**: *Depending on the material, we may occasionally hold a tutorial during lecture time, or a lecture during the tutorial time.* If a lecture/tutorial is planned to be virtual, then we will announce that during the previous lecture/tutorial.
- **Reflection**: At the end of term, students will have the opportunity to write a short reflection about their experience in the course, for an optional 0.5% credit. More details will be given at that time.
- **Assignments** will be submitted electronically in the form of a single Python notebook, which will include both the theoretical and the experimental/implementation questions. Submission details will be provided in class and online.
- **In-class exercises** will be handed in on paper, and/or possibly a working program will be shown to the TA during tutorial time.
- The class will make use of the following modes of interaction:
    - **Teams Discussions** for questions and answers pertaining to the assignments and the lectures.
    - If any lectures or tutorials are held entirely online, this will be done through the MS Team as well. In those cases, student participation will be via the chat window, and may contribute to the overall course grade. How to join video. - Introduction to MS Teams - How to download MS Teams.
    - Students are encouraged to self-organize by creating study groups as private Microsoft Teams (no instructor participation), or connect on other media of their choice (e.g. Facebook, Whatsapp or Discord).
- **Course announcements** will be posted to the General channel of the class MS Team and/or made in class. It is the student's responsibility to check their Dal e-mail and MS Team of the class on a daily basis. It is also the student's responsibility to be aware of the material covered, and announcements made, in class. To access your Dal e-mail see: https://www.dal.ca/dept/its/o365/services/email.html

Evaluation Criteria

- **Assignments** (20% over all assignments)
    - 3-5 assignments (each worth between 5% and 10%, as described in class),
    - Assignments and exercises are to be done individually unless otherwise specified.
    - Assignments will be checked for plagiarism.

- **Exercises** (5% over the term)
    - In-class and other small exercises throughout the term.

- **2 Midterms** (35%)
  - As long as the course stays in person, then there will be **two** in-class tests/midterms, consisting of multiple-choice and written answer questions.
  - As long as the course stays in person, then each midterm will be worth exactly 17.5% of the final grade.
  - In-class midterms are closed-book. We may discuss in class the possibility of a single common sheet that will be provided to everyone during the midterm.
  - **o** For full credit, your solutions to written answer questions must be clearly and fully explained in your own words, to demonstrate an understanding of the concepts involved (e.g. correct keywords alone may not receive any marks). Note that an answer containing a correct statement followed by an incorrect statement may receive a lower score than one containing only the correct statement. Be clear. Examples of strong and weak answers will be shown in class and/or in online discussion.
  - We will discuss in class the details of how midterm and exam questions will be graded.
  - The duration of each midterm will be specified in advance.

- **Final Exam** (40%)
  - There will be a written final exam, held during the scheduled exam period.

- **Bonus Marks**
  - There may be opportunity throughout the term to accumulate bonus marks through active participation.
  - Bonus marks will be discussed in class as a means of potentially rounding up final grades at the end of term.

- **Oral Exam (possible)**
  - For any submitted work, the instructor may always choose to request a followup oral examination (given by the instructor or by one of the TA's) on Teams of a student to confirm the student's command of the material and ownership of answers. The oral exam will be recorded.
  - An oral exam request is not an accusation; it is an opportunity for the instructor to have a short in-person conversation with you about the course material.

Notes

- A minimum grade of C is required in this course if it is core to your FCS degree, or if it will be used as a prerequisite for a subsequent CSCI course.
- As of 2019, students who receive a grade lower than C in the same required CS course twice will be dismissed.
- The grade conversion scale in Section 17.1 of the Academic Regulations, Undergraduate Calendar will be used.
- **A student must pass** (i.e. achieve 50%) **both** (1) the assignment component and (2) the combined midterms and final exam, to pass the course.
- **It is up to the discretion of the instructor to use remote proctoring in online testing.** Students may be required to download proctoring software onto their devices. Students who cannot meet system requirements for remote proctoring should contact the instructor for an alternate assessment. (Typical system requirements are: (i) Mac OS or Windows, (ii) a web-cam, and (iii) an internet connection.)
- There will be no converting of final grades to pass/ill this semester. All letter grades will be final.

Student Declaration of Absence

The Student Declaration of Absence policy for missed or delayed course requirements shall apply.
https://www.dal.ca/campus_life/safety-respect/student-rights-and-responsibilities/academic-policies/student-absence.html

Exam Requirements

- The final exam will be in-person, with the following conditions::
    - Photo ID is required
    - Closed book. We will discuss in class the possibility of a single common sheet that may be provided to everyone during the midterm. Furthermore, this option may be dependent on overall class participation.
    - Unless decided otherwise, a non-programmable calculator will be allowed.
- The midterms will be held in-class, under the same conditions as the final exam
- If the final exam or midterm need to be online due to public health conditions,
    - remote proctoring software will be used
    - questions will be randomly ordered
    - each question must be answered and submitted before proceeding to the next question
    - No going back to previous questions is allowed

Academic Standards

Failure to properly attribute sources in your work will be treated as an academic standards issue and points may be deducted for not following citation requirements. For example, forgetting to quote text taken from other sources, failure to include in-text citations, or failure to include the required information in the citations or references. Please see the resources on proper citation provided by the Dalhousie Writing Center (https://dal.ca.libguides.com/c.php?g=257176&p=5001261).
Please note that if it appears that the error was made with intent to claim other people's work as your own such as a lack of both citations and references, an allegation of plagiarism will be submitted to the Faculty Academic Integrity Officer, which could result in consequences such as a course failure.

Tentative Schedule

The schedule below is tentative and subject to change. If we need to add material (e.g. more detailed review of linear algebra) then there will be corresponding questions about this material on the midterms and final exam.

Readings, videos, homework, etc will be assigned as the term progresses.

One of the main reading references will be the lecture notes by Roger Grosse, as discussed in class and links will be provided (e.g. through Teams).

Date	Topic				Week
8 Jan 2023	**Introduction** What is this course about?				W1

	Dependency graph **Classification Intro**				
10 Jan 2023	**Nearest Neighbours**				W1
12 Jan 2023	**T: Probability Review I**				W1
15 Jan 2023	**Information Gain, Decision Trees**				W2
17 Jan 2023	**Ensembles, Bias-Variance**				W2
19 Jan 2023	**T: DTs, Lin Alg**				W2
22 Jan 2023	**Linear Regression I**				W3 drop date 1
24 Jan 2023	**Linear Regression II**				W3
26 Jan 2023	**Linear Classification I**				W3
29 Jan 2023	**L: Midterm Review**				W4
31 Jan 2023	**Midterm 1**				W4
2 Feb 2023	**DaI Closed**				W4
5 Feb 2023	**Linear Classification II**				W5 drop date 2
7 Feb 2023	**SVMs (vid)**				W5
9 Feb 2023	**T: Catch up/Reading**				W5
12 Feb 2023	**Clustering I** **K-means**				W6
14 Feb 2023	**Clustering II** **GMMs**				W6
16 Feb 2023	**Clustering III** **EM**				W6
19 Feb 2023	**Reading Week**				W7

21 Feb 2023	**Reading Week**				W7
23 Feb 2023	**Reading Week**				W7
26 Feb 2023	**Lin Alg II** Eigenvalues, SVD				W8
28 Feb 2023	**MV Gaussians**				W8
1 Mar 2023	**PCA**				W8
4 Mar 2023	**Naive Bayes**				W9
6 Mar 2023	**Generative Models I**				W9
8 Mar 2023	**Generative Models II**				W9
11 Mar 2023	**T: Midterm Review**				W10
13 Mar 2023	**Midterm 2**				W10
15 Mar 2023	**No tutorial**				W10
18 Mar 2023	**Neural Networks I**				W11
20 Mar 2023	**Neural Networks II**				W11
22 Mar 2023	**T: Neural Networks**				W11
25 Mar 2023	**T: as needed**				W12
27 Mar 2023	**Extra Topics**				W12
29 Mar 2023	**Dal Closed**				W12
1 Apr 2023	**Extra Topics**				W13
3 Apr 2023	**Question-based Review**				W13
5 Apr 2023	**T: Question-based Review**				W13
8 Apr 2023	**Friday Makeup Classes**				W14

9 Apr 2023	**Friday Makeup Classes**				W14
11-23 Apr 2023	**Exam period**				W15

General FAQ [link] about this course

Electronic communications etiquette

(adapted from Prof. S. Adamo's course policy)

Students often don't appreciate how many emails their professors and TAs receive in a single day, most of which require a quick response. E-mails constitute professional correspondence and the informal messages students may use to communicate with their friends are not appropriate in this context. This point is more likely to get lost in online courses, where face-to-face contact does not exist, and students have additional stressors due to the online delivery mode.

Since I am inundated by email and messages, it is not the best way to reach me for most matters (unless I explicitly ask you to email me). If you do email me, you can increase your chances of receiving a reply by making sure that you:

- Read the syllabus and website to determine whether the information you require is already available
    - We will not respond to emails that ask about information contained in the course syllabus or on the course website - not even to tell you that the email will not be answered
    - We will not respond to emails that ask whether 'anything important' was or will be covered during an absence - not even to tell you that the email will not be answered
    - Email should not be an alternative to taking the initiative and making the time to look for the information you require
- Enter the course number in the subject line: CSCI3151 (no space)
- Provide a professional salutation ("Hey!" is not appropriate)
- Write in full, grammatically correct sentences
    - Your professors do not want to receive text-speak
- Proof-read your email message to be sure that it is clear and provides all pertinent information
    - Avoid the time-wasting back-and-forth that happens when a respondent must request more information before being able to act
- Do not write to request an exception to the syllabus
    - All students are subject to all the rules as outlined in this course syllabus and no exceptions will be made to any class policies under any circumstances
- Provide a professional closing
- Include your FULL name and Dal student ID
- Before you hit the send button, proof-read your message once again and delete aggressive, hostile, sarcastic or abusive language. Such messages will receive no response. Useful advice on your email's tone is available here https://blog.hubspot.com/sales/email-etiquette-tips-rules (item #11).

Responsible Computing Policy

Usage of all computing resources in the Faculty of Computer Science must be within the Dalhousie Acceptable Use Policies (https://www.dal.ca/dept/university_secretariat/policies/information-management-and-technology/acceptable-use-policy-.html) and the Faculty of Computer Science Responsible Computing Policy. For more information please see https://www.dal.ca/content/dam/dalhousie/pdf/faculty/computerscience/policies-procedures/fcs_policy_local.pdf

Use of Plagiarism Detection Software

All submitted code may be passed through a plagiarism detection software, such as the plagiarism detector embedded in Codio, the Moss (https://theory.stanford.edu/~aiken/moss/) Software Similarity Detection System, or similar systems. If a student does not wish to have their assignments passed through plagiarism detection software, they should contact the instructor for an alternative. Please note, that code not passed through plagiarism detection software will necessarily receive closer scrutiny. https://cdn.dal.ca/content/dam/dalhousie/pdf/dept/university_secretariat/policy-repository/OriginalitySoftwarePolicy.pdf

Student Health and Wellness

Taking care of your health is important. As a Dalhousie student, you have access to a wide range of resources to support your health and wellbeing. Students looking to access physical or mental health & wellness services at Dalhousie can go to the Student Health & Wellness Centre in the LeMarchant Building. The team includes: registered nurses, doctors, counsellors and a social worker. Visit **dal.ca/studenthealth** to learn more and book an appointment today.

Students also have access to a variety of online mental health resources, including telephone/texting counselling and workshops/training programs. Learn more and access these resources at d**al.ca/mentalhealth**.

Culture of Respect[1]

Every person has a right to respect and safety. We believe inclusiveness is fundamental to education and learning. Misogyny and other disrespectful behaviour in our classrooms, on our campus, on social media, and in our community is unacceptable. As a community, we must stand for equality and hold ourselves to a higher standard.

What we all need to do:

1. **Be Ready to Act:** This starts with promising yourself to speak up to help prevent it from happening again. Whatever it takes, summon your courage to address the issue. Try to approach the issue with open-ended questions like “Why did you say that?” or “How did you develop that belief?”
2. **Identify the Behaviour:** Use reflective listening and avoid labeling, name-calling, or assigning blame to the person. Focus the conversation on the behaviour, not on the person. For example, “The comment you just made sounded racist, is that what you intended?” is a better approach than “You’re a racist if you make comments like that.”

---

[1] Source: Speak Up! © 2005 Southern Poverty Law Center. First Printing. This publication was produced by Teaching Tolerance, a project of the Southern Poverty Law Center. Full ”Speak Up” document found at: http://www.dal.ca/dept/dalrespect.html. Revised by Susan Holmes from a document provided April 2015 by Lyndsay Anderson, Manager, Student Dispute Resolution, Dalhousie University, 902.494.4140, lyndsay.anderson@dal.ca www.dal.ca/think.

3. A**ppeal to Principles:** This can work well if the person is known to you, like a friend, sibling, or co-worker. For example, “I have always thought of you as a fair-minded person, so it shocks me when I hear you say something like that.”
4. **Set Limits:** You cannot control another person’s actions, but you can control what happens in your space. Do not be afraid to ask someone “Please do not tell racist jokes in my presence anymore” or state “This classroom is not a place where I allow homophobia to occur.” After you have set that expectation, make sure you consistently maintain it.
5. **Find or be an Ally:** Seek out like-minded people that support your views, and help support others in their challenges. Leading by example can be a powerful way to inspire others to do the same.
6. **Be Vigilant:** Change can happen slowly, but do not let this deter you. Stay prepared, keep speaking up, and do not let yourself be silenced.

University Statements

This course is governed by the academic rules and regulations set forth in the University Calendar and the Senate. https://academiccalendar.dal.ca/Catalog/ViewCatalog.aspx?pageid=viewcatalog&catalogid=117&loaduseredits=False

Territorial Acknowledgement

Dalhousie University is located in Mi’kma’ki, the ancestral and unceded territory of the Mi’kmaq. We are all Treaty people.
Dalhousie acknowledges the histories, contributions, and legacies of the African Nova Scotia people and communities who have been here for over 400 years.

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.” 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. 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. (read more: http://www.dal.ca/dept/university_secretariat/academic-integrity.html)

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: https://www.dal.ca/campus_life/academic-support/accessibility.html for all courses offered by Dalhousie with the exception of Truro.

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). (read more: http://www.dal.ca/cultureofrespect.html)

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. (read more: https://cdn.dal.ca/content/dam/dalhousie/pdf/dept/university_secretariat/policy-repository/Code%20of%20Student%20Conduct%20rev%20Sept%202021.pdf)

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. (read more: 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. (read more: https://cdn.dal.ca/content/dam/dalhousie/pdf/dept/university_secretariat/policy-repository/OriginalitySoftwarePolicy.pdf)

Learning and Support Resources

Please see https://www.dal.ca/campus_life/academic-support.html

Student Use of Course Materials

These course materials are designed for use as part of the CSCI courses 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 material to a commercial third-party website) may lead to a violation of Copyright law.