

INTRODUCTION TO MACHINE LEARNING

SYLLABUS: CSC 311 FALL 2019

1. Instructors.

- Richard Zemel
Email: csc311prof@cs.toronto.edu
Office: Pratt 290C
Office Hours: - Wednesday 1pm-2pm
- Murat A. Erdogdu
Email: csc311prof@cs.toronto.edu
Office: Pratt 286B
Office Hours: Friday 11am-1pm

2. Lectures. This course has three identical sections:

- L0101: Monday 11:00-13:00 at RW 117
Monday 15:00-16:00 at RW 117
- L0201: Wednesday 11:00-13:00 at BA 1170
Wednesday 15:00-16:00 at BA 1170
- L0301: Thursday 16:00-18:00 at MB 128
Thursday 18:00-19:00 at MB 128

3. Teaching Assistants.

- Anil, Cem
- Cao, Tianshi
- Chan, Harris
- Flam-Shepherd, Daniel
- Hosseinzadeh, Rasa
- Killian, Taylor
- Liao, Shun
- Madras, David
- Mehralian, Ehsan
- Nestor, Bret
- Ng, Nathan
- Razdaibiedina, Anastasiia
- Zhou, Audrina

4. Course webpages. The course webpage contains all course information, additional readings, assignments, announcements, office hours, etc. Please check it regularly!

- erdogdu.github.io/csc311
- q.utoronto.ca

5. Course Evaluation.

- 4 assignments: 40%
- Midterm exam: 20%
- Final exam: 40%

6. Course Outline. This course covers several commonly used machine learning algorithms and related methodological concepts. Topics may include:

1. Introduction to ML & Nearest Neighbours
 - (a) Supervised vs unsupervised learning
 - (b) Training, test, validation
2. Decision Trees & Ensembles
3. Linear Regression & Classification
4. General Issues
 - (a) Overfitting and generalization
 - (b) Cross validation
 - (c) Bias variance tradeoff
5. SVMs & Boosting
6. Neural Networks
7. Principal Components Analysis & K-Means
8. Mixture Models and Expectation-Maximization
9. Matrix Factorization
10. Bayesian Linear Regression
11. Reinforcement Learning
12. Recent Topics

7. Prerequisites. CSC207H1, Mat235Y1/Mat237Y1/Mat257Y1/(minimum of 77% in Mat135H1 and Mat136H1)/(minimum of 73% in Mat137Y1)/(minimum of 67% in Mat157Y1), Mat221H1/Mat223H1/ Mat240H1, STA247H1/STA255H1/STA257H1

Exclusion: CSC411H1, STA314H1, Ece421H1. Note: Students not enrolled in the Computer Science Major or Specialist program at the UTSG, UTM, or UTSC are limited to a maximum of three 300-/400-level CSC/ECE half-courses.

8. Textbooks. There is no required course textbook. The following materials can be helpful.

- Christopher M. Bishop (2006). Pattern Recognition and Machine Learning
- Ian Goodfellow, Yoshua Bengio and Aaron Courville (2016), Deep Learning
- Kevin Murphy (2012). Machine Learning: A Probabilistic Perspective
- Trevor Hastie, Robert Tibshirani, Jerome Friedman (2009). The Elements of Statistical Learning
- Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani (2017). Introduction to Statistical Learning
- David MacKay (2003). Information Theory, Inference, and Learning Algorithms

9. Assignments. There will be 4 assignments in this course. The assignments will be released on the course webpage.

9.1. *Collaboration policy.* Collaboration on the assignments is not allowed. Each student is responsible for his/her own work. Discussion of assignments should be limited to clarification of the handout itself, and should not involve any sharing of pseudocode or code or simulation results. Violation of this policy is grounds for a semester grade of F, in accordance with university regulations.

10. Exams. There will be a midterm exam (tentatively) on Oct 21, 6pm. Details will be announced in class and on the course webpage. You can bring an optional A4 cheat sheet - double-sided. Final exam is in December, date TBD. You can bring two optional A4 cheat sheets - double-sided.

11. Late policy. If you are traveling, you may email your submission to one of the course staff one week in advance of the deadline. Ten percent of the value will be deducted for each late day (up to 3 days, then submission is blocked). No credit will be given for assignments submitted after 3 days. Extensions will be granted only in special situations, and you will need a Student Medical Certificate or a written request approved by the course coordinator at least one week before the due date.

12. Grading concerns. Any requests to have graded work re-evaluated must be made within one week of the date the grade is released. Re-evaluation may result in a decrease in the grade.

13. Computing. In the assignments and project, you may need to write your own programs, debug them, and use them to conduct various experiments, plot curves, etc. You may use any programming language, but `Python`, and `R` might be preferable. On the midterm and final exams, you will not be asked to understand a particular languages syntax and will not need to provide code in any particular language. You may be asked to provide pseudo-code.

14. Missed Tests.

- If a test is missed for a valid reason, you must submit documentation to the course instructor.
- If a test is missed for a valid medical reason, you must submit the University of Toronto Verification of Student Illness or Injury form to your instructor within one week of the test.
- The form will only be accepted as valid if the form is filled out according to the instructions on the form.
- Important: The form must indicate that the degree of incapacitation on academic functioning is moderate, serious, or severe in order to be considered a valid medical reason for missing the term test. If the form indicates that the degree of incapacitation on academic functioning is negligible or mild then this will not be considered a valid medical reason.
- If the midterm test is missed for a valid reason then the final test will be worth 60% of your final grade. Other reasons for missing a test will require prior approval by your instructor. If prior approval is not received for non-medical reasons then you will receive a term test grade of zero.