Co se O line

This course was new in 2018. It was then, and still is, inevitably somewhat experimental. It is also addressed to a rather heterogeneous audience, as it bears two numbers, one for Honours courses, the other for PhD courses. This means that I cannot tell with any certainty just how much work we will be able to cover.

The main text for the course is a book by Aurélien Géron, which I have found to be very useful for learning how to program machine-learning algorithms. The book was originally called *Hands-on Machine Learning with SciKit-Learn and TensorFlow*, and it is available from the O'Reilly website, to which McGill people can get free access. However, the

way unless at least some of the di erent ways of training an algorithm have been implemented by each student. Some, but I hope not all, students may still be su ering from the analogue of math anxiety regarding computer programming, and I will try to be helpful in getting people over this nowadays unnecessary anxiety.

This book is valuable for having some classic explicit examples: Chapter 1: First (tiny) application; Chapter 2: Iris dataset; Chapter 7: House prices.

- *Introduction to Machine Learning with Python*, Andreas C. Müller and Sarah Guido, O'Reilly 2017.

  The particular chapter of interest here is Chapter 7: Text data.

- *Machine Learning with R*, Brett Lantz, Packt 2013.

  I don't myself care much for R, as I find its syntax rebarbative. However, it has many aficionados, and so I draw attention to this book.

- *GANs in Action*, Jakub Langr and Vladimir Bok, Manning

  This book covers some pretty advanced topics, but using only the most basic mathematics. Myself, I think that some topics would be clearer with just a little more sophisticated math. The idea, as the title suggests, is to exposit the Generative Adversarial Networks, introduced mainly by Ian Goodfellow in his PhD thesis at the Université de Montréal. In order to do so, it has to provide a rather extensive overview of many of the topics we cover at an earlier stage, and so it can serve as a valuable additional reference.

- *Perceptrons* (Expanded edition 1988), Marvin Minsky and Seymour Papert, MIT 1969.

  This book more or less put an end to earlier attempts to develop artificial intelligence. It contains even so the seed of the idea that subsequently developed into artificial neural networks, here called \perceptrons".

- *The Elements of Statistical Learning*, 2nd ed, Hastie, Tibshirani, and Friedman, Springer 2008.

  This counts as the definitive reference for statistical machine learning, written before deep learning became popular. It is quite mathematically advanced, and is definitely directed to an audience of statisticians. It is available, quite free of charge, at https://web.stanford.edu/ hastie/Papers/. Look for ESLII.

- *An Introduction to Statistical Learning with applications in R*, James, Witten, Hastie, and Tibshirani, Springer 2013, reprinted 2017.

  Statisticians love R, and so this is a companion volume to the one above, pitched at a lower mathematical level. It too can be found, free of charge, online, at https://www-bcf.usc.edu/ gareth/ISL/. There is a download link for the PDF version of the book.

- Here are URLs for some online tutorials for Python libraries of immense value for what we want to do with machine learning.

  https://www.tutorialspoint.com/numpy

  A wider range of tutorials is at

  https://www.tutorialspoint.com/

Method of Evaluation

As each year since I began to teach this course, I will consult you all before committing myself definitively to the method of evaluation. What I want is for each student to undertake the training of an algorithm, preferably but not necessarily using deep learning, on some (reasonably) big dataset. Although I say each student, it is also possible for students to work in a team of two. If some of you are still a little computer-shy, you might find it valuable to team up with another student who is more experienced with computers. The results of the work should then be written up, along with the relevant source code, and submitted, along with the results, as the document on which a final mark could be based. One way in which such a document could be created is to use a Jupyter notebook; go to

https://jupyter.org

for information on what a Jupyter notebook is, and how to work with one.

Unless there is strong opposition, I plan also to set a couple of assignments that will carry some weight in the final grade, but will not be allowed to lower the grade based solely on the big project. The idea behind the assignments is just to give you practice in machine learning and confidence in carrying it out.

Another possibility would be a formal exam, but that would be contrary to the spirit of the course. While some weight could perhaps be given to students' ability to answer theoretical questions, or to carry out some mathematical operations successfully, I wouldn't want this weight to be very great.

Academic Honesty

You'll have seen the following in all of your course outlines, because the McGill Senate requires that it should appear in all of them. I used to think of it as a pure formality, but a disturbing number of cases of plagiarism have been detected in recent years. So, please take seriously all the admonitions in the following text.

1) Right to submit in English or French written work that is to be graded [approved by Senate on 21 January 2009]: In accord with McGill University's Charter of Students' Rights, students in this course have the right to submit in English or in French any written work that is to be graded. This right applies to all written work that is to be graded, from one-word answers to dissertations.

2) According to Senate regulations, instructors are not permitted to make special arrangements for final exams. Please consult the calendar, section 4.7.2.1, General University Information and Regulations, at http://www.mcgill.ca.

3) Academic Integrity statement [approved by Senate on 29 January 2003]: McGill University values academic integrity. Therefore all students must understand the meaning and consequences of cheating, plagiarism and other academic offences under the Code of Student Conduct and Disciplinary Procedures.

   (see http://www.mcgill.ca/students/srr/honest/ for more information).

   Et en français:

L'université McGill attache une haute importance à l'honnêteté académique. Il incombe par conséquent à tous les étudiants de comprendre ce que l'on entend par tricherie, plagiat et autres infractions académiques, ainsi que les conséquences que peuvent avoir de telles actions, selon le Code de conduite de l'étudiant et des procédures disciplinaires (pour de plus amples renseignements, veuillez consulter le site http://www.mcgill.ca/students/srr/honest/)