

Economics 706/420

Fall 2020

R. Davidson

Course Outline

This course was new two years ago. 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 Ph.D. courses. This means that I cannot tell with any certainty just how much work we will be able to cover.

The first text for the course is *Deep Learning*, by Ian Goodfellow, Yoshua Bengio, and Aaron Courville.

[This link](#) takes you to the website for the book. Its contents are completely available online. The book is also available [in hardcover](#). It is published by the MIT Press.

The book contains three main sections. The first, *Applied Math and Machine Learning Basics*, provides necessary background material. It is the final chapter of this section that first discusses machine learning explicitly, and this will be our starting point. It may be necessary, at least for some students, to consult some of the earlier chapters on mathematical prerequisites.

The second section is what we will spend most time on. It is entitled *Deep Learning: Modern Practices*. These “Modern Practices” will constitute the principal points of interest for this course. The third section, *Deep Learning Research*, is concerned with where machine learning is going. It makes for fascinating reading, but we will certainly not have time to look at this section. Students intending to go further with machine learning may find it interesting reading, although even it will no doubt be out-dated pretty soon.

Once we have got a good idea of the basics, we will go on to the study of the modern practices. This study necessarily involves work with the computer. Machine Learning cannot be understood in any useful way unless at least some of the different ways of training an algorithm have been implemented by each student. Some, but I hope not all, students may still be suffering from the analogue of math anxiety regarding computer programming, and I will try to be helpful in getting people over this nowadays unnecessary anxiety.

The second 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, and also as a valuable complement to the main textbook for theoretical considerations. The book was then called *Hands-on Machine Learning with SciKit-Learn and TensorFlow*, and it is available from the [O’Reilly website](#), to which I believe that McGill people can get free access. However, the book has been updated, with the new title *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd Edition*. Keras is a software layer that makes programming an algorithm even easier than with TensorFlow, the leading platform until very recently.

Most of my own recent programming projects have been in the C++ programming language, but, unless one has quite a lot of prior experience with C++ , it has a steep learning curve. In the last few years, I have been working with Python, an interpreted language

that has, for the most part, a straightforward syntax, and can be learnt swiftly by anyone with even just a little experience of programming.

The relative simplicity of programming in Python is probably the main reason for which Python has quite the best set of libraries for machine learning, and not just for deep learning. Although deep learning will be the main focus of the course, I plan to look at some other machine-learning techniques, for which the Python libraries are equally useful. The study of some of these other techniques reveals how much all modern machine-learning approaches have in common, despite the fact that some are much better adapted than others for specific applications.

Provisionally, here is a summary of the topics I covered over the last two years, and hope to be able to cover this year as well.

- Artificial neural networks (ANN)
- Deep feedforward networks and back-propagation
- Regularisation and optimisation of deep networks
- Convolutional networks
- Recurrent networks
- Gated units
- Random forests
- Auto-encoders

Resources

There is a super-abundance of resources available online for studying machine learning, and for implementing it. Machine learning is often coupled with the buzzword *Big Data*, and this is simply because machines usually learn better if they have a lot of data available to train their algorithms. Many big datasets are available online, the best known, and probably the most comprehensive, being

<https://www.kaggle.com/datasets>

Here are some of the available resources which I found useful. First, some books that may or may not be available free on the Internet.

- *Deep Learning*, Goodfellow, Bengio, and Courville, MIT Press (2016) – our main text.
- *Hands-on Machine Learning with SciKit-Learn, Keras, and TensorFlow*, Aurélien Géron, O'Reilly 2017.

This is an excellent reference for people wishing to learn how to use the Python libraries named in the title of the book. It really is hands-on, with detailed instructions for the examples given.

- *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”.

- *Artificial Intelligence with Python*, Prateek Joshi, Packt 2017.

From this book, the interesting part is Chapter 16: use tensorflow to do linear regression with a perceptron.

- *Building Machine Learning Systems with Python (2nd ed)* Willi Richert and Luis Pedro Coelho, Packt 2013.

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.

- *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

I wish to consult you all before committing myself to some specific 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. Then the way in which this was done could 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.

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, not especially at McGill, but in other North American universities. 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/>)