Response to Student Feedback
received for the Aalto course CS-EJ3211
"Machine Learning with Python"

corresponding teacher: Alex Jung (first.last (at) aalto.fi)

August 27, 2019

We express our sincere gratitude for the insightful and constructive comments and suggestions provided by the students of the first edition (under preliminary course code "CS-EV Vaihtuvasisältöinen opinto") of this course during spring 2019. We have tried to address all comments to the extent possible.

Major modifications we implemented include the following:

- We have positioned this course now as a basic entry point to machine learning. The focus is on the modelling of real-world problems ("business cases") as machine learning problems (ML) and their solution using some off-the-shelf methods provided by Python libraries.
- The presentation of the material has been adapted such that students can follow easily if they are familiar with the concept of matrices and vectors and have been programming in some high-level (object oriented) programming language.
- We now discuss in some detail the use of Python libraries such as scikit-learn for solving simple machine learning problems.
- We have extended the amount of visible grader tests ("sanity checks") in order to better guide students through assignments.
- We have replaced most of the examples with new "business cases" which are modelled as machine learning problems and solved using few lines of Python code.

We now respond in a point-by-point fashion to a selection of relevant student feedback.

S1 *Sometimes I couldn't understand what I had done wrong in an exercise.*

T1 We have added more sanity checks in the notebooks to help in debugging student solutions. Moreover, we point students more clearly towards using the discussion forum for specific questions on the coding assignments.

S2 *The provided course book is in a way good to provide mathematical foundation behind, but has very little benefits in more practical sense ('translate' mathematical algorithms into computer code). Some inconsistent mathematical notations in the assignment ipynb, which can be confusing and doesn't help. Also, inconsistent numpy array dimensions even in the same task. It's not in anyway useful to students if the tasks has more complications because of some little insignificant practicalities (different array dimensions, etc).*

T2 We have made efforts to make the notebooks more self-contained and also provide more guidance for how to implement ML methods in Python. We have also carefully revised the notebooks to avoid any inconsistencies in notation of conventions regarding variable types used for matrices and vectors.

S3 *Unclear how final grade will be assigned.*

T3 We now indicate the grading scheme more clear (see Section "Grading") right from the start of the course.

S4 *Registration and course enrolment was tricky and bugged (both FiTech and Aalto).*

T4 The registration process has been revised by both FITech and Aalto. FITech is now using Studyinfo.fi (Opintopolku) for enrolments. Studyinfo is hopefully more user friendly compared to the system used before. The registrations collected by FITech are automatically transferred to Aalto systems making the registration process smoother and faster for the student.

S5 *The intermediate tests for the code snippets should be more complete to validate the steps and catch the errors early enough. Now, it was possible to pass everything and notice a possible bug in the last test, when it's much more laborious to find the cause.*

T5 We have added more intermediate tests (sanity checks) in the notebooks.

S6 *More tests to know whether the solution is correct. The feedback should be right away, not after weeks. And possibility to correct the code if it was wrong.*

T6 The students are informed about the achieved points at the time they submit a notebook. We are hesitant to release correct solutions during the course for organisational reasons.

S7 *I don't want to use my name on public forums. Slack is bad.*

T7 We now point out (in the MyCourses Section "Slack Discussion Forum") how to use Slack without revealing the full name and instead using a nickname.

S8 *It's not clear to me how these ML techniques are specifically related to python, and the actual content was very very similar to ML Basic Principles. With that said, I was hoping for some more focus on practical application, but everything stayed very abstract, and issues like incompatible matrix/vector shapes distracted from the overall learning.*

T8 We have adjusted the content of this course so that the focus is now more on formulating applications as Machine learning problems. Moreover, instead of discussing particular ML algorithms in details we demonstrate how to apply several different methods using Python libraries. We have also revised the notebooks carefully to avoid any inconsistencies in matrix shapes.

S9 *The demands of the assignments and sub-tasks were very uneven in my opinion. Some whole assignments were very quick and easy to complete, but on the other hand, some assignments took much more time and effort, and even sub-tasks took days to solve. Asking for help in Slack was at times frustrating as most replies referred to the existing theory, which I had already read multiple times. Also, some of the Demos were not working properly apparently, which was a bit confusing.*

T9 We have tried to make the workload for each of the (sub) tasks more even. The demos have been revised carefully in order to better guide students through the problems.

S10 *For some of the tasks, it was a bit hard to know if the solution is correct at all.*

T10 We have added sanity checks and also point to the Slack channel to get better feedback on the correctness of solutions for coding assignments.

S11 *Sometimes it was difficult to follow the lecture's book and how the explained methods can be practically implemented.*

T11 The role of the course book is now de-emphasized and it mostly serves as further reading for students who wish to deepen their understanding beyond the scope of this course. The notebooks have been geared toward implementation of ML methods using Python libraries.

S12 *I would have benefited of verbal feedback on the mistakes I did in coding. Sure I do understand that it's really hard to provide such on this kind of large course.*

T12 We are considering options for offering (virtual) personal meetings with TAs or the teacher of the course.

S13 *While I do understand that the math is the precise way to describe the world, the way how it is used in the "book" and in the exercise, makes it very hard to read and has an awful lot of useless repetition in it (as a result of to define things over and over again). We're here for the ML content, not for the Math content really; it doesn't make the course less authoritative if it sometimes uses more "layman terms in two sentences" instead of "half a page precise math formula". ToDo: I really hope you could review the material with someone outside of academic world. :-)*

T13 The notebooks and coding assignments are now more focused on hands-on skills for applying ML methods via Python libraries instead of deriving these methods from first principles. However, we provide many pointers to additional material allowing students to deepen their understanding beyond the scope of the course.

S14 *I heard about this course in an 'AI Monday' session, and there it was presented as 'not needing much mathematics or programming skills'. This image did not match reality. I do have fairly good maths and coding skills, only from quite a while ago, i.e. not fresh in my mind. Given the above, this course was not for me. It was for more (recently) math and tech savvy folks. Which is a bit of a pity, because we would need open, hands-on courses for a broader audience. Finally, even if the subject matter is somewhat math and coding intensive, there are less and more pedagogical ways to teach it. I switched to Deeplearning.ai's courses, and those offered me clearly more positive learning experiences. Good teaching breaks down a more complex topic into approachable pieces, and encourages from there, instead of trying to create an impression of mastery by keeping things difficult and laborious to approach. This is partly cultural, I'm afraid, to our detriment. Of course, promoting the course should raise more realistic expectations in the first place.*

T14 We have tried to revise the notebooks such that they convey the "big picture" and also focus on the correct application of ML methods that are readily available in Python libraries. In particular, we do not require students to apply linear algebra or calculate derivatives of functions by hand.

S15 *It is better to have all the materials and assignments early. I was waiting for the release of assignments. I am losing my tempo after a long one week gap. P.S. : Being self taught course, I expect so.*

T15 We will release the notebooks weekly such that students can start on the new round within one week.

S16 *There is no exercise lectures*

T16 We have now added more pointers to video lectures of Prof. Andrew Ng which should serve as a good substitute for having "classroom-style" lectures. Moreover, we are also investigating alternative means for providing contact teaching, e.g. via online meetings.

S17 *I felt that from round 3 or 4 onwards the exercises started to require quite little time compared to the beginning. Perhaps this was because I was learning but I also felt that the latter rounds should have contained more student tasks.*

T17 We have tried to make the workload more even towards the later exercises.

S18 *It would be really beneficial to learn more about how to apply the higher level ML toolkits, like Scikit, and to use those to solve harder problems. Perhaps this can be a follow-up course if you just have the resources. I would also like to see more emphasis on the computational complexity and how to measure or predict those depending on the amount of data, selected features and algorithms. This is really important to master when one starts to apply these skills in real life were you have to pay for the compute when you start to scale it. Again, this might be part of the "advanced" ML course.*

T18 The new course edition will exactly follow this suggestion by making extensive use of readily available ML methods via Python libraries such as scikit-learn. Moreover, we now included tasks related to computational complexity and scalability of ML methods.

S19 *Ensure that the exercises contain correct information and instructions. Too often the instructions contained mistakes, like the shapes of the function parameters. This was confusing in the beginning and caused me to waste time. Towards the end I had learned enough to spot and avoid these mistakes. In that sense perhaps the mistakes allowed me to learn more ;).*

T19 We have revised the notebooks carefully to avoid any (unintentional) inconsistencies and typos.