

Course Report WASP Graduate School

Date: 2023-03-24

Authors¹: Raazesh Sainudiin and Marina Axelson-Fisk

Scalable Data Science and Distributed Machine Learning, 6hp

Semester: Fall 2022

Number of registered students: 41

Answering frequency (course evaluation²): 25/41 (61%)

Examination results

Number of students examined: 41

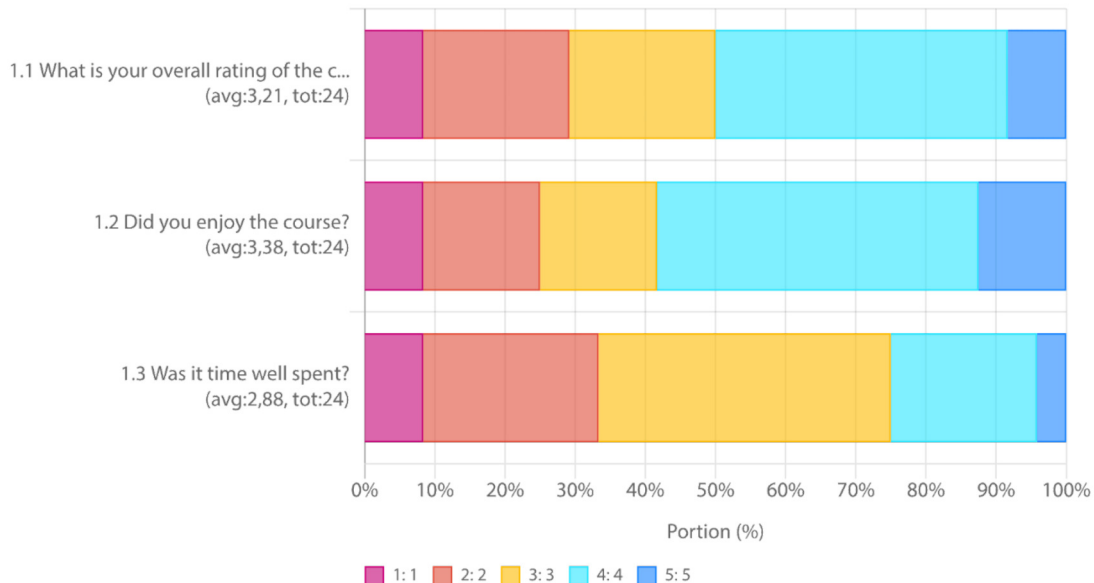
Fail: 0 (0%)

Pass: (41) (100%)

Brief summary of student viewpoints and suggestions

Results of WASP base-line quantitative questions

- What is your overall rating of the course? (1-5) Mean: **3.21** (distribution below)
- Did you enjoy the course? (1-5) Mean: **3.38** (distribution below)
- Was it time well spent? (1-5) Mean: **2.88** (distribution below)
- *Other selected quantitative results required to make an evidence-based proposal to improve the course are shared as images along with reflections.*



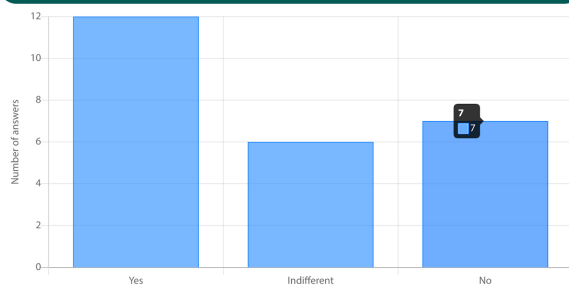
¹ The report should be written by the examiner together with the teachers and possibly others, such as teaching assistants

² retrieved on 2023-03-24 from <https://www.lyyti.fi/e/?p=q&h=2875ad3eb13d143b419f737c77c49550>

Question 2 – Did you find the academic, technical and entrepreneurial experiences shared by the invited industrial speakers insightful? – was answered negatively by 7 out

of 25 students (28%), positively by 12 students (48%) and neutrally by 6 students. Given, that the invited speakers – a total of six AI entrepreneurs from Uppsala, Stockholm, Vilnius and Palo Alto were invited in 2022 – were not compensated for their time or travel costs, but offered WASP-sponsored lunch and dinner, and only 28% of the surveyed students perceived their presentations to be un insightful, it may be worthwhile having at least 3 invited entrepreneurial speakers representing AI/Data industry in

2. Did you find the academic, technical and entrepreneurial experiences shared by the invited industrial speakers insightful?

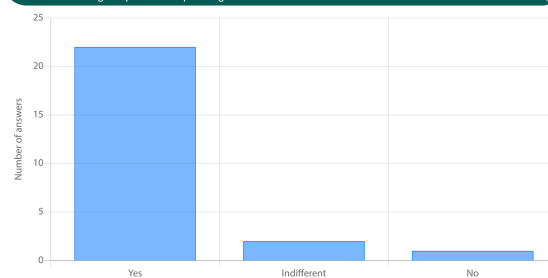


the Nordics, North America and Europe/Eurasia in 2024 (those outside Uppsala-Stockholm area will be invited online via zoom as done for the speaker from Palo Alto in 2022).

Question 3 – Did you appreciate the course design for the theoretical part of the course (jointly given with Reza Zadeh’s Stanford materials) that allowed you to go as deep as you wanted to via flipped classroom lectures with relatively detailed notes and minimal reading comprehension quiz assignments? –

was answered negatively by only one student. For some historical context based on experiences from 2020, this format for theoretical background was chosen to cater to the large variance in interests and backgrounds of the students. It seems to have been successful. Therefore the 2024 version will be a more thorough variant of the 2022 version covering introductory materials on analysis of algorithms in more detail than done in 2022 for those who need all prerequisites in self-contained lecture notes and videos (only 50% of students have had undergraduate course in analysis of algorithms based on a live show-of-hands survey done by the lecturer).

3. Did you appreciate the course design for the theoretical part of the course (jointly given with Reza Zadeh’s Stanford materials) that allowed you to go as deep as you wanted to via flipped classroom lectures with relatively detailed notes and minimal reading comprehension quiz assignments?

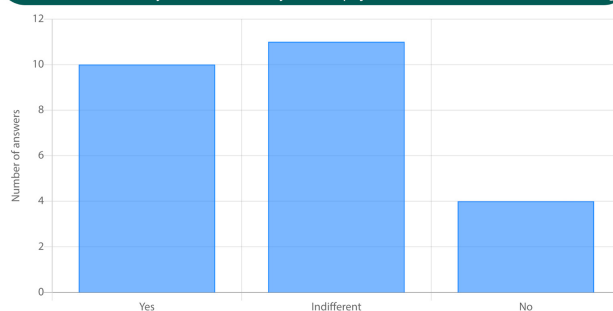


Question 4 – Would you have preferred to have access to videos of Raaz going

through notebook codes and theoretical materials that you had access to from the other modules of ScaDaMaLe that were not covered as part of the interaction time in the first two modules but may have been relevant to your course project? –

Given that 10 students have a preference to have access to videos of the lecturer going through the content in the codes in notebooks, it may be sensible to provide this to the students if time permits. Also, this will allow those who

4. Would you have preferred to have access to videos of Raaz going through notebook codes and theoretical materials that you had access to from the other modules of ScaDaMaLe that were not covered as part of the interaction time in the first two modules but may have been relevant to your course project?

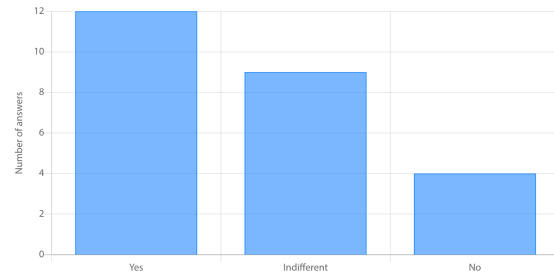


want to absorb the material to absorb in a flipped setting at their most comfortable video playback speed (0.25x,0.5x,1.0x,1.5x,2.0x). Such videos can be made optional as it is not necessary to understand the codes, given the codes are self-documented in detail as

evidenced by 15 students (60%) in the survey not really preferring such videos of all the codes and materials designed for self-study, depending on the learning path dictated by the student's group project.

Question 5 – If this course were to be taught again by Raaz in 2 years, do you think it is reasonable to expect the students to go to Uppsala for the first two modules but go to a conference centre at the geographic mid-point of Sweden that is accessible by Swedish Railways for the third module? – is answered negatively only by 4 students (16%), and answered positively and neutrally by 12 (48%) and 9 (36%) students, respectively.

5. If this course were to be taught again by Raaz in 2 years, do you think it is reasonable to expect the students to go to Uppsala for the first two modules but go to a conference centre at the geographic mid-point of Sweden that is accessible by Swedish Railways for the third module?

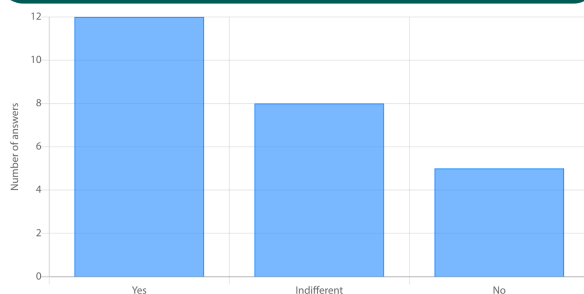


Therefore, it is not unreasonable to have one of the modules, especially the first module (with a gorgeous Norrland's fall scenery), meet in Torpshammar, close to the Geographical midpoint of Sweden, where the lecturer will be confined during the frost-free growing season (62° N June–September 2024 & 2026) in a research field station, already being built by repurposing an abandoned traditional Swedish farm from 1901, for experiments in general living systems theory jointly with other scientists and engineers, in a long-term international research project over the next 2-7 years. The first module of 2024 can meet at the conference centre in Torpshammar³ (accessible by Swedish Rail) and a bulk cost-effective deal can be secured by planning over a year ahead in time. The remaining two modules can be organised in Uppsala University as done in 2022.

Other reflections on travel per hp: The lecturer is also happy to do all three modules in Uppsala or reduce three visits to just two visits with the first of the two visits over 2 over-night stays instead of just one. These considerations which are based on requests of the students in the survey for reducing the amount of travel per hp may be possible if the Graduate School deems it necessary. The lecturer's personal view is that it is entirely reasonable to expect the students to travel with one overnight stay per 2hp module and not make any exceptions just because all three modules are given by the same lecturer.

Question 6 – Do you think that the new cohort of PhD students might benefit from a 30-45 minutes long introduction to 'starting your own company in Sweden?' during the Community-Building Summer School or during the mandatory course on Societal Legal and Ethical Aspects of AI?

6. Do you think that the new cohort of PhD students might benefit from a 30-45 minutes long introduction to 'starting your own company in Sweden?' during the Community-Building Summer School or during the mandatory course on Societal Legal and Ethical Aspects of AI?



'starting your own company in Sweden?' during the Community-Building Summer School or during the mandatory course on Societal Legal and Ethical Aspects of AI? – is answered negatively by 5 students (20%). Given 12 (48%) and 8 (32%) students answered positively and neutrally, the suggestion seems to be favoured. The question was posed based on the observation that many students simply lacked the basic

economic commonsense understanding of how academia and industry are entwined in ensuring the competitiveness of the local economy. Many students seem to have rather deeply ingrained false notions that academia is somehow self-existent and independent of

³ <https://www.torpshammar.com/> and <https://www.torpshammar.com/konferens/>

industry, and are unaware of historical facts around the birth of Vapnik Chervonenkis theory of statistical machine learning originating in US Bell Labs, or current facts about latest academically relevant research coming from various private industrial labs for instance. The lecturer believes strongly that just 30-45 minutes of interactions during the community-building summer school is enough to set the economically realistic framing of the inherent interaction between entrepreneurship, industry, and academia in most countries including Sweden today.

Answers to free text-questions to be (shortly) summarised under “Strengths” and “Weaknesses”

- What was the best aspect of the course?
- What would you suggest improving?
- What advice would you like to give to future participants?
- Other comments. Is there anything else you would like to add?

“Strengths” according to students⁴

- “The collaboration with Stanford and Reza was great”; “The stanford video lectures and the project.”; “I liked the theoretical parts”
- “excellent in terms of not only providing up-to-date theoretical and practical lessons, but also providing insight to entrepreneurial experiences in various fields and to the opportunities in terms of applying research to real world problems”
- “Socialization and the chinese restaurant plus learning scalable coding and having hands on the codes. In short, practicality mixed with theory.”; “The social part”.
- “Comprehensive lecture notes and easy-to-follow notebooks providing lots of practical examples of running code.”;
- “The topic of the course is very relevant for careers in data science, I also very much appreciated the “open-endedness” of the course.”
- I like the combination of theory and practice. The material is comprehensive....
- “The hands-on project”...
- Personally, I greatly appreciate the organization of the course materials and the insight to the application areas. I believe it motivated the students to look for application areas of their research work in the real world....

“Weaknesses” according to students¹

- “I think the first module is unnecessary to participate in in person and it can be performed online. The Others are ok to be offline. The idea of going to a conference centre at the geographic mid-point of Sweden for the third module sounds great!”
- “Too many guest presentations that seemed not aligned with the learning outcomes of the course...”
- “reduce the course material a bit, but dive deeper in scala/python implementations and the ideas behind their form of implementation...”
- “Lectures where we just scroll through notebooks are not for me.”
- “Have more conventional lectures during the in-person gatherings in Uppsala. It was difficult to follow notebook walkthroughs live.”
- “It is unclear why we needed to go to Uppsala just to watch notebooks being scrolled through”
- “I think its well structured. But might need to add additional materials for future.”

⁴ Based on both quantitative results and key viewpoints from students’ free-text answers

- “Raaz adds a lot of side information when he talks (I guess psychologically he is a perfectionist). :) I am fine with that, but some cannot follow especially if English is not their mother language.”
- “My main gripe is regarding logistics around the course. It is not necessarily time-efficient to go to Uppsala every time. I would rather have more modules, but more split up between online and offline sessions.”
- The "political" (e.g. the Marxist analysis of Amazon Echo) stuff should be either cut or be less vague and hard to understand. If you're going to talk about your societal grand theory of everything, don't just ramble about it, have some structure. Or stick to the actual specifics of the course. The Swedish word "flummig" comes to mind - either say something meaningful and definite, or keep your peace.
- “I thought Raaz's enthusiasm was good and made some of the more tedious material accessible, but I think he should give more time for students to contribute during the lectures and interactive sessions.”

Comments from teachers on the implementation and outcome of the course⁵

- Overall, the lecturer feels that the 2022 version of the course went significantly better than the 2020 version as the three modules were tightly entwined and the collaboration with Professor Reza Zadeh at Stanford for the theoretical part of the course was successful.

Proposed changes/comments/asures

- There is room for improvement. The main difference in 2024 would involve more flipped classroom materials for the applied/coding aspects of the course as requested by many students. So students don't have to watch me scroll through the coding cells in the notebooks.
- Minor Points: To prepare students for discussions especially around the formal *descriptive game-theoretic setting* for the *data science process* in the broader geopolitical economic context, reading comprehension quizzes have to be assigned so students are *forced* to actually read the required discussion materials. Only one student actually read the reading on “Anatomy of AI”, the article by Microsoft Researchers on Amazon Echo (based on a raise-your-hands-survey conducted before the discussion).
- The lecturer has already shared some of his thoughts when reflecting on the student answers to the five specific additional survey questions above for improving the 2024 version of the course.
- A lot more structured interactions will be planned in the 2024 versions focusing more on discussions and problem-solving based on preparatory work already done before the physical meeting.
- Perhaps only three or four instead of six or seven industrial speakers will be invited in the 2024 version.

⁵ Including changes effected during the course