

Informatics Student Course Feedback 2017/18

<http://www.inf.ed.ac.uk/teaching/surveys/2017-18>

This report contains feedback from students about a course taught in the School of Informatics during the 2017/18 academic year, in response to the following questions:

- What would you say to students interested in taking this course?
- What did you find most valuable about the course?
- What improvements, if any, would you make to the course?
- Please add any other comments you have about workshops and tutors

Each course organiser receives this report as well as statistics on multiple-choice responses. All these reports, together with student feedback about individual members of teaching staff, are collected and sent to the Director of Learning and Teaching.

Please note that these are personal responses from individual students: some courses only have a few responses and a small sample can be unrepresentative.

Stereotyping and bias, especially unconscious bias, is a serious concern in any survey gathering personal responses. All students received the rubric below before completing the surveys, and you can read a brief introduction to issues of unconscious bias on the university web pages at <http://edin.ac/2iypZBv>

This information is provided for students and staff at the University of Edinburgh: **you may not redistribute or reuse it without permission**. If you would like the information in another format or want to use it in your own publication then please contact the Informatics Teaching Organisation at <http://www.inf.ed.ac.uk/teaching/contact>

Rubric given to all students taking the end-of-course feedback survey

We value your opinions on the courses you take here at the University, as they allow us to shape future delivery and development. We welcome constructive comments about your courses, whether positive or negative, and ask you to give details about any issues in order to help the course organiser to understand and address them.

We encourage you to be aware of the potential for bias in the completion of these questionnaires, so we have developed resources which may be helpful to you:

- Equality, Diversity and Unconscious Bias (<http://edin.ac/2iypZBv>)

You also have a responsibility to provide feedback in a manner which does not breach the University's Dignity and Respect Policy:

- University of Edinburgh Dignity and Respect Policy (<http://edin.ac/1Cq0VZY>)

The results of the questionnaires will never be analysed in a way that seeks to identify individual students from their responses. However, should you wish to remain anonymous, please do not identify yourself in your answers to the survey questionnaire implicitly or explicitly.

Comments Report

What advice would you give to a student taking this course in future?

- Be prepared to learn the material by yourself
- Be prepared to work a lot on the coursework, just to have a single piazza post destroy your solution.
- Blame computing support if any machines required for the coursework is not working properly. Coursework help on Piazza sounded so bureaucratic.
- Don't count on hadoop not being broken last minute, start coursework early
- I would think twice before choosing this course. In my opinion, EXC is a poorly organised course but I am hopeful that things can be improved. The slides were a high-level view of the course. Unfortunately, at lectures not much more relevant content was presented by the lecturers.
- The coursework tasks were ambiguously defined. The marking criteria was mostly about efficiency, while efficiency is not taught. As a result, students implemented their own version of efficiency which was not accepted by the markers.
- If you're the type who does codes things until they work be careful with the coursework; they want optimal solutions.
- It helps if you have experience in programming, mainly Python. It's not required, but will make the course smoother.
- Let me start by saying I did not enjoy this course but I know plenty of people who did. But.... Unless you already have good knowledge of distributed computing and the concepts behind it as well as a very keen interest in the subject, this is not a course you want to take. It might seem interesting but the difficulty of the course quickly leaves it hard to follow. Especially when core concepts are not explained because they are deemed to be simple and therefore not necessary to go in depth about. Furthermore the courseworks are worth 12.5% of your final mark each but unless you are a MapReduce god you should definitely plan at least 25 hours of work per coursework.
- Make sure you are well versed with programming
- Prepare yourself for interesting, but challenging work. Getting a high mark is much harder than in many other courses so don't take this to increase your average.
- Quiz yourself regularly. There was a lot of content in this class that I intuitively felt like was irrelevant fluff but kept coming up.
- Start courseworks early and discuss how you solve things with friends (plenty of times I thought I had made something efficient before finding out about aspects I hadn't considered)
- Start the coursework early
- Start the coursework early and ask questions!
- Take this course, if you're interested in learning how to deal with absurdly large amounts of data and efficient ways for the computer to process it.
- The focus of the course is efficiency not just correctness. It will teach you how to create better solutions and better code, but you might need to change the way you think about problem solving.
- This course is fun but not suitable for students with little programming experience.
- Try to understand very well every concept that is explained. On the coursework try to review your answers a couple of times trying to find flaws and thinking how you would improve it.
- Worthwhile, useful and a bit challenging

Comments Report

What did you find most valuable about the course?

- A very interesting topic
- Coursework helped a lot in understanding the concepts.
- Courseworks gave practical insights, demos were invaluable when performed in-lecture.
- Definitely learn a lot about computing at scale. However coursework felt more like a troubleshooting Hadoop task than carefully and critically thinking about the best approach.
- I enjoyed the content of this course very much.
- I found the opportunity to work on an actual computing cluster to be the most valuable.
- It gives us a holistic view of extreme computing. We also can practise it by projects.
- Learning Hadoop
- Learning a lot about big Data and how companies deal with data in their scale.
- Learning about hadoop and other scalable approaches to coding
- Learning the value of parallel computing
- Learning to use Hadoop and the ideas behind distributed computation.
- Learning what businesses in the industry actually do.
- Lecture material and video records were very good
- Practical experience from the assignments.
- Real-World Examples and Guest Lectures
- Really well explained.
- That I learned to utilize new tools such as GNUParallel and Hadoop, was introduced to new technologies and new ways of thinking when it comes to data processing
- The assignments were intellectually challenging, and made me a more confident programmer - particularly with respect to using the terminal and thinking about distributed programs.
- The challenging assignments.
- The course does cover many important concepts for distributed computing
- The coursework took much more time than expected for most of the tasks (roughly 12.5 according to the school policy). Even though lots of it is spent running the jobs on cluster and debugging.
- I would appreciate if the coursework contained slightly less tasks with remaining tasks left for workshops, or making the coursework weight more.
- The courseworks were a very practical application of the content of the course.
- The instructor complements each other really well. Kenneth provided us with brilliant insights. Also huge props to Volker cause he is really talented in teaching and communicating his thought to the students. Both instructors are available to help students and the Piazza forum is well organized.
- The course is really good to improve/ refresh unix terminal commands.
- The labs were extremely useful for doing the assignments and complementing the classes and the computer cluster

Comments Report

What did you find most valuable about the course? (continued)

- The lecturer was great. Super smart and knew his subject. The course was definitely hard, not like a lot of boring work, but intellectually challenging. Not much of a difference between not doing your homework and not doing it efficiently, which is a bit of a bummer. Helps you realise that most important things are not done at the scale we are used to and there are a lot of things to consider then.
- This Hadoop skills and mapreduce skills are useful to me.

Comments Report

What improvements, if any, would you make to the course?

- Add a lab on pig/spark since they are commonly used by the industry. Also, release lecture presentations prior to the lecture to allow students to follow read them beforehand.
- Assignment marking was tough, you can lose a couple of marks for minor details, and the same amount of marks for sub-optimal code. It would be better if the assignment were marked out of 100 instead of 25, then you could still lose marks for minor details (e.g. not outputting blank lines) but they would be smaller reductions compared to sub-optimal code, which is where most effort was put.
- Coursework feedback was too brief and vague. Marking scheme for coursework is not known, eg how many marks are awarded for correctness and efficiency, etc. Sample solution to coursework should be published. Coursework help on Piazza sounded so bureaucratic.
- Coursework organization was atrocious, and throws you in at the deep end with seemingly unfair marking. If they would provide more detailed feedback and sample solutions, I would at least know where I went wrong.
- I didn't understand the general structure of the course, I don't know a lot about the subject, so I'm not saying there wasn't any, but I wasn't able to understand it and it sometimes felt like every class was a random topic, so perhaps it would be useful to present a broad outline of the whole course on the first lecture.
- I especially had trouble following Kenneth's lectures, I think he assumed we already knew the main idea of the topic instead of starting from the beginning like Volker, although the MapReduce hammer was cool.
- Also, it should be explained what is expected from our projects when they say they our code has to be efficient. As a student from a non-computational background I had a lot of problems understanding what they were expecting from this (and in the end it turned out to be a lot less than what I thought)
- I prefer the assignments to be defined more precisely. I understand that real world problems are often vague, but in the real world you work in different conditions and get different feedback as well.
- I would have liked having introduction lectures on the content (eg. map reduce) before the lectures going into all the practical applications of that method. It would also have been valuable to have had a few lectures on efficiency as I felt the only way I learned about this was from reading the feedback for my assignments which felt a bit late.
- I would like to have done more assignments/labs related to some of the techniques discussed in class but not actually assessed e.g. chord, 2 phase commit, etc.
- Improve the cluster so it doesn't go down while we are trying to complete the coursework.
- Stop live coding in the lectures, we can't read or follow what you are doing.
- Improve the communication of the objectives of the coursework. So many question were asked to clarify things on piazza than it was necessary. Communicate clearly the objective of the task with additional sample output and give us any restrictions that we might have beforehand (languages, frameworks, libraries, etc.)
- Lecture slides don't feel helpful in retrospect.
- More about Hadoop or Spark could be focus on in the course.
- More code examples and well-guided labs that will help you with the courseworks.
- More emphasis on efficiency in lectures.
- More informative feedback. The feedback lets you know you are wrong but not how to improve. For me personally, this means I still don't know how to idiomatically use GNU parallel to combine files with the same space in parallel (w/ locks, w/o locks, in shared memory? etc)
- More labs would be wonderful. Especially about the topics covered at the end of the course.
- More structured delivery of lectures in some cases

Comments Report

What improvements, if any, would you make to the course? (continued)

- Move exam to December
- Nothing much.
- One of the lecturers uses jargon without always explaining the meaning. I know many course mates who will be starting PhDs (and so are generally capable) who cannot follow the lectures and have to learn the material by themselves. I think this is a case of poor teaching rather than inherent difficulty of the material. My best lecturers do not just read off from some hastily constructed slides - they think hard about how to present the material as clearly as possible.
- Please write proper and understandable lecture notes. The slides you are using are confusing.
- Provide more feedback for the results of the coursework as well as discuss best/different solutions in the lectures.
- Sometimes people asked perfectly reasonable questions on Piazza and got snarky responses from the instructors. I felt that created a hostile environment, particularly for non-native speakers. The whole tone of the course was quite 'macho'/ 'alpha male'. A little bit more empathy would help a great deal.
- The course feedback is insufficient. Furthermore, the grading criteria of the coursework is very unclear.
- The course is often presented in a confusing way and challenging concepts are frequently dismissed as trivial. Additionally the labs in this course help with the courseworks a little bit but not enough. So longer (or more labs).
- The coursework is marked a bit too harsh in my opinion. There are a lot of students, markers also make mistakes, so careful about that. A bit too challenging for 10 credits, especially compared to other even 20-creds courses.
- To me, lecturers and instructors had a condescending tone when a student asked for help. In addition, I wish that Yes/No answers on Piazza contained an explanation to help the student who is asking. Students started copying the tone of the lecturers when answering other students' answers, and Piazza became a condescending, Yes/No society.
- I wish Kenneth was more honest and elaborate when answering student's questions (about coursework in particular). That would made him more approachable. Volker sometimes could not answer questions about the lecture slides, and sometimes did not present until the end of the lecture.
- In my opinion, EXC is a poorly organised course but I am hopeful that things can be improved. The slides were a high-level view of the course. Unfortunately, at lectures not much more relevant content was presented by the lecturers. Either the slides or lecturers' presentation should have provided more exam- and coursework-relevant information.
- The coursework tasks were ambiguously defined. Some tasks had not been mentioned in the lectures (e.g. basic GNU parallel was demonstrated in a lab but the coursework asked that we go beyond that). The marking criteria, which was mostly about efficiency, should have been defined or efficiency should have been taught. Otherwise, students implement their own version of efficiency which was not accepted by the markers. As a result, a majority of the students became disappointed and angry after the coursework feedback was returned. This can be observed in Piazza.
- Tutorials, I don't find it fair if we are expected to take an exam in a practical course, without there being any tutorials or thoroughly written notes, lecture slides didn't go deep enough into the topic. Some topics such as consistency models were extremely confusing, and I feel like the explanation was poor for it. Visualisation or better examples for some topics would have been appreciated.

Comments Report

Please add any other comments you have about workshops and tutors

- Also improve the labs as they were rather slow pace and quite primitive and did not help much when it comes to solving the coursework. For hadoop allow us to write with some guidance the configuration to run on the cluster as well as a basic program to utilize most of the features hadoop streaming is offering (changing partition key, setting reverse numerical order, setting number of reducers, etc.). At the GNUParallel lab very little portion of the tutorial
- Good labs!
- I didn't really see the point of the labs. It was hard to know if we were doing the right thing in the exercises or to get useful feedback during the labs. Asking afterwards on Piazza was more effective.
- I would have liked to do more of them, and on different topics. I think a great way of accomplishing the above would be to have optional tutorials where we do e.g. the 1 hour intro to MongoDB.
- Labs are not quite useful, they should be changed to coursework enquiry sessions.
- Labs weren't hugely helpful except for a basic introduction to a couple of the coursework tools.
- Piazza was really useful.
- The Hadoop lab was a good introduction, however, I would have benefited if the section of keys was rewritten to be more straightforward, current and relevant. The GNU lab can be more descriptive in order to help us complete the coursework. The DB lab was a nice addition.
- The lab demonstrator was incredibly good at explaining concepts and very approachable.
- The labs didn't add much to my experience taking the course.
- The labs were just a copy and paste exercise. Would be better to have questions and release solutions for them
- The labs were really informative and well organized. One could do them without the help of the tutors. In general, all the material for the course are well organized.
- Labs of Chord, BigTable, Storm and all other introduced technologies would have been very interesting.