COMP6235 Referral coursework

Module:	Foundations of Data Science			Lecturers:	ES, MB
Assignment:	Repeat			Weight:	100%
Deadline:	14/12/18	Feedback:	18/01/2019		

Instructions

The assessment for an external repeat of the module is based on 100% coursework. You will have to complete the individual courseworks set during the term, e.g.:

CW1 (worth 30% of the marks), see: http://edshare.soton.ac.uk/19467/ and http://edshare.soton.ac.uk/19466/ for the corresponding data set.

CW2 (worth 30% of the marks), see description of coursework when released in week 8 on the wiki page (https://secure.ecs.soton.ac.uk/noteswiki/w/COMP6235/1819)

The deadlines for CW1 and CW2 apply as specified in the course wiki:

https://secure.ecs.soton.ac.uk/noteswiki/w/COMP6235/1819

Additionally, instead of CW3 which is sets as a group assignment, you will have to complete the following individual coursework, detailed below. This individual coursework is worth 40% of the marks in the module.

Instructions for individual coursework 3

In this coursework, you will develop a data science application, demonstrating all aspects of the data science pipeline covered in the module. This application may be on any topic in any domain of your choosing, but should include more than one dataset, and use data science techniques to provide insight towards a real-world problem. You will be expected to follow good software engineering practices, evaluate the effectiveness or reliability of your model or analysis, and justify design and methodological choices.

You are not allowed to pick a topic that is similar to the one you have had in your COMP6235 group project before.

Submission

You will submit two items:

- a PDF report about the work undertaken in the style of a conference paper, including a link to a GitHub repository containing the datasets, the software developed, and related documentation. The report should use the standard 2017 ACM conference proceedings style (using the sigconf style option for the template). The paper should be at most 6 pages in length, including all references and, if applicable, appendices.
- a PDF or Powerpoint set of slides describing the work done, which will presented to the module leader in the week December 16th-December 21st. The presentation (which can be carried out remotely, e.g. via skype) will take 45 minutes, of which: 15 minutes will be dedicated to your presentation; 15 minutes to Q&A; and 15 minutes feedback. The module leader will contact you in due time to propose a timeslot for the meeting.

Both report and slides should be submitted by December 14th 2018 4 pm.

Marking Scheme

The conference paper and presentation will be marked as a single piece of work using the following criteria. The weights refer to the share of points available for each category. The total number of points adds up to 100.

- **Choices and justifications** – The student should choose a suitable challenge, and justify the reasoning behind the choice, and the technologies and techniques used to solve it - 15%
- **Analysis of results** - The overall effectiveness, as well as strength and weaknesses of your application should be discussed - 30%
- **Technical implementation** - The application should be complete, should run without errors, and behave as defined - 35%
- **Report** - The report should be clear and professional in tone. High quality references should be used to justify statements – 20%

Standard ECS late submission penalties apply.

Feedback

You will receive feedback to your work in the meeting with the module leader. A summary of the feedback will be provided to you in writing as well by January 18th.