School of Computer Science – Module Assessment Sheet for 2021-2022

Module	Big Data / COMP4103 (BIG) / Semester 2
Module Convenor(s)	Isaac Triguero

Assessment Name	Group work	**Weight**	50%
Description and Deliverable(s)	Research group project – The aim is to develop a big data solution in Python for either a specific data mining/machine learning algorithm or to solve a particular big data problem. The group work includes some warm-up exercises with various deadlines to help students get the necessary background for their projects. This assignment is broken down into different pieces: - 4 warm-up exercises with weekly deadlines on MapReduce, Spark RDDs, Spark DataFrames and Spark MLlib - Expression of Interest to take a particular project or suggest one (PDF file) - 6-page paper in IEEE format (PDF file) - Github repository with the code (Python code) - Peer review scores (CSV file) - Presentation (Slides, live demo) Full details in the group project brief.		
Release Date	- Ex01 - MapReduce exercises – 10th February 2022 - Research Project Brief - 21st February - Ex02 - RDD exercises – 24th February 2022 - Ex03 - DataFrame exercises – 3rd March 2022 - Ex04 - MLlib exercise – 10th March 2022		
Submission Date	- Ex01 - MapReduce exercises – 18th February 2022 - Ex02 - RDD exercises – 4th March 2022 - Ex03 - DataFrame exercises – 11th March 2022 - Expression of Interest – 18st March 2022 - Ex04 - MLlib exercise – 25th March 2022 - **[Optional]** Preliminary Project Submission – 8th April 2022 - Final Submissions - 4th May 2021 - Presentations/Conference – 10th May 2022		
Late Policy **(University of Nottingham default will apply, if blank)**	- Solutions to the weekly assignments will be provided the week after the deadline. Work submitted after the deadline will be subject to a penalty of 5 marks (the standard 5% absolute) for each late working day out of the total 100 marks, and maximum deadline of 5 working dates. - All groups need to demo their project on the 'conference' day. Exceptions to this will need to be supported by the extenuating circumstances procedure.		
Feedback Mechanism and Date	Written individual feedback in Moodle within two weeks from submission.		
Assessment Criteria	The different components of the group project are weighted as follows: - 4 Warm-up exercises (20 marks, 5 marks each) - Expression of interest (15 marks) - First Project submission (0 marks – not assessed – optional to get feedback) - Final Project submission (50 marks) - Presentation (15 marks) **Detailed Assessment Criteria:**		

	- **Warm-up exercises** (20 marks): Each group should submit their solutions to the 4 warm-up exercises. *Full details provided in the specification documents.* - **Expression of Interest** (15 marks)**:** Each group should submit a document containing a table with the 3 chosen projects and initial thoughts (up to 300 words for each project). They can suggest their own project. o **Motivation and initial analysis** ▪ *Understanding of the chosen topics.* ▪ *Has the team provided enough motivation as to why they want to carry out this research?* ▪ *Are the initial suggestions appropriate for a Big Data project?* - **Final Submission** (50 marks): *Each group should submit a paper-based description of their big data solution together with the code produced.* o **Motivation and context** ▪ Do the team understand the need for a big data solution for their project? ▪ Do they provide a compelling argument as to why is needed? o **Design/Methodology** ▪ Explanation of the methodology. ▪ Is the design efficient? (e.g. does it reduce as much as possible data movement across workers?) ▪ Innovation of the proposed methodology. Does it provide a standard divide and conquer strategy, or has the team thought of an elaborated strategy to alleviate the bottleneck of the methods to tackle big data. o **Experiments and results** ▪ Are the experiments well designed to test the proposed solutions? ▪ Do the results support the original motivation? ▪ Is the analysis coherent? o **Writing** ▪ Clear description, reproducibility ▪ Quality of visual elements, illustrations, tables. ▪ Quality of References o **Code – software quality** ▪ Efficiency, suitability of Spark operations to solve the problem ▪ Documentation o **Response to Reviews** ▪ Justification of changes (if any). Has the team considered the reviews provided by their peers? o **Contribution/Reflection and Peer reviews**: One page to motivate the project and their reflection. In addition, each member of the team is asked to anonymously assess the contribution of each other member of the team. This will feed into the final mark of the project. - **Group Presentation** (15 marks): *Each group is asked to deliver a 12-min presentation summarising their contribution + 5 mins for Q&A.* o Quality and clarity of the presentation o Response to questions from the panel and public o Understanding of their solution o Individual participation in the presentation. All members are expected to participate equally. Full details provided in the group project brief.

Assessment Name	Examination	**Weight**	50%
Description	Paper-based exam (alternatively on Moodle for remote students)		
Release Date	June 2022, exact date TBA.		
Submission Date	2 hrs after release.		
Late Policy **(University of Nottingham default will apply, if blank)**	For short-timed (e.g. 2 hrs) exams in Moodle and Rogō, an additional 30 minutes will be permitted for all students to cover IT issues and upload time. This will be factored into the overall exam duration in Moodle and Rogō. Beyond the above, late submission of exams in not accepted.		

Reassessment Method	**Weight**
Exam	100%

Hours Break-down for this Module

Activity	Per Week	Total Hours
Lecture – delivery key material	2 × 10	20
Research Project (50%) – Lab time (10 hours) is part of the group work as the students are expected to use this time to work on their assignments. Remaining 40 hours are allocated to the research group work and submissions. The labs/group project will use the concepts introduced in the lectures, so it's preparing students for the exam!		50
Examination (50%) – revise and further study		30
Total (20 credits)		**100**