CN7031

Big Data Analytics

SEPTEMBER 2025-26

Computer Science and Digital Technologies (CSDT)

MODULE GUIDE 2025/26

University of East London

Accessibility Notice

If you require this document in an Alternative Format please contact Link qae@uel.ac.uk.

There are also a range of free tools in built on devices to help you access this document and other readings. You can find more information on the Ability net My Computer My Way website and the Assistive Technology and Ease of Access Intranet page.

Contact Information

Module Leader

Name: ***Dr Amin Karami***

Email: a.karami@uel.ac.uk

Tel: ***020 8223 3347***

Room Number: ***EB.1.98***

The Module Leader/Other Tutors and Contact Details were correct at point of publication. You will be notified of any changes.

Module Introduction

This is a 30-credit, level seven module within the Computer Science & Digital Technologies Department and is a core requirement of several MSc courses, such as MSc Big Data Technologies, MSc Artificial Intelligence, MSc Cloud Computing, and MSc Computer Science.

This module is intended to deliver Big Data systems and applications including hands-on exercises and tutorials based on Hadoop and Spark ecosystems. The core content of this module is large-scale data analytics using extensively Python and SQL over two widely used in-parallel computing ecosystems and frameworks namely Hadoop and Spark. We will take students to the cutting-edge topics to provide them with most important and significant aspects of large-scale data analytics.

The most important thing to say is that good engagement on this module is absolutely essential for success. The feedback from previous students was very positive and as long as you engage with classes, there is no reason why you should not be very successful on this module. Along with the course texts, that will be made available to you, all the teaching support material you need is on an accompanying Moodle site, and it is important that you consult this site regularly throughout the course. Microsoft Teams will be our platform for on-line communication.

We hope you enjoy this module and look forward to teaching you this term again.

Dr Amin Karami (CN7031 Module leader)

MODULE AIMS AND LEARNING OUTCOMES

- This module aims to provide students with the core theoretical and practical background required for big data analytics and developing big data systems. It will provide you with an insight into areas of big data management and advanced analytics. You will develop in-depth practical skills through using tools and techniques from the forefront of the emerging field of data analytics.

- At the end of this module, students will be able to:

Knowledge

1. Explain the current and emerging concepts, technologies and principles relevant to Big Data analytics.
2. Illustrate understanding of the current advanced techniques, tools and methods applicable on Big Data visualisation and applications.

Thinking skills

3. Examine approaches for preparing Big Data from different and heterogeneous data sets.
4. Critically evaluate appropriate visualization techniques within a business context.

Subject-based practical skills

5. Design, implement, and test a system working with Big Data.
6. Utilise batch and real-time Big Data and their visualisations effectively.

Skills for life and work (general skills)

7. Work effectively in groups to develop a software project.
8. Present Big Data analytics to technical and non-technical audiences

Key Information

The module will be delivered face-to-face for both lectures and tutorials (laboratory based practical works). Lectures and Tutorials will be used to introduce both the theoretical and practical aspects of the module. You will learn large scale data analytics and modelling using Hadoop and Spark frameworks. Hence, these sessions will be used to reinforce the lecture and tutorial materials and put the theory into practice. The majority of your time on this module will be spent in private study. You are expected to use this private study time to review the lecture and tutorial materials, carry out the suggested reading for that week, prepare for practical as instructed and work on your coursework tasks.

As with all modules at UEL, extensive use will be made of a virtual learning environment called Moodle and a communication tool called Microsoft Teams.

You can access all your Moodle and Teams sites via UEL's TrackMyFuture portal:

https://trackmyfuture.uel.ac.uk

You can also directly access your CN7031 Moodle site by going to the following link:

https://moodle.uel.ac.uk/course/view.php?id=82303

The Moodle site will contain a variety of information and resources including teaching and learning materials (links to reading, lecture slides, lab exercises, etc), a calendar of important events and coursework deadlines and important news regarding the operation of this module.

You should check Moodle regularly!

Teaching Schedule

When the module starts:

- **Lecture:** from 1st Oct onwards
- **Tutorial:** from 3rd Oct onwards
- Please find your allocated lecture and tutorial time slots from https://ueltt.uel.ac.uk, your weekly in-person attendance is required.

Tentative Module Contents

- **W 1 (1-2 Oct):** Module Intro on Big Data [Amin]
- **W 2 (8-9 Oct):** Hadoop + HDFS [Amin]
- **W 3 (15-16 Oct):** Data Acquisition using Sqoop [Fahimeh]
- **W 4 (22-23 Oct):** Hadoop Hive for Structured Data [Fahimeh]
- **W 5 (29-30 Oct):** Unstructured Data in Hadoop [Amin]
- **W 6 (5-6 Nov):** Apache Spark + Spark DF [Fahimeh] [CRWK handout]
- **W 7 (12-13 Nov):** Spark Data Source API + Spark RDD I [Fahimeh]
- **W 8 (19-20 Nov):** HuggingFace with PySpark [Amin]
- **W 9 (26-27 Nov):** Spark RDD II [Amin]
- **W 10 (3-4 Dec):** Spark RDD for Unstructured Data [Amin]
- **W 11 (10-11 Dec):** Module review + CRWK Preparation [Amin]
- W 12 (15th – 19th Dec.): CRWK Presentation and Submission

More information

- Link to your **personal timetable** (UEL login required)
  - **Link to the Docklands Campus Map**
  - **Link to the Stratford Campus Map**
  - **Link to University Square Stratford**
  - **Guide to Room Numbers**

Engagement and Attendance Requirements

Evidence shows that good attendance has a positive impact on academic performance. As a university, we are committed to supporting you in reaching your full potential.
Student Engagement, Retention and Success Unit (SERT) is here to support you in reaching your full potential. By monitoring attendance and engagement, we can identify early signs that you may need support and work with you to address any challenges that could be affecting your learning.

If you are experiencing challenges that impact your ability to engage with your course, you can speak to a member of the team in person at either the Docklands or Stratford campus, or remotely by telephone. To book an appointment or ask a question, please use MyPortal or call The Hub on 0208 223 4444 to speak with our Duty Advisor (available weekdays, 10am-1pm, excluding bank holidays).

Assessment Information

A group-based (3-5 members) CRWK (100%) including two tasks (RDD and DF)

- Presentations: week 12 (15th-19th December 2025)
- Turnitin Submission: 14th December 2025, 10pm

All the group members must attend the presentation. If you do not attend, you fail the module. Every member of a group will be assessed individually.

Assessment Task	Submission Deadline
Part 1: Group based CRWK Report (60 marks) Specification, Design, and Implementation	Submit by Sunday 14th December 2025, 10pm (via Moodle)
Part 2: Presentation (40 marks) Presentation and individual assessment	15th-19th December 2025, 9am-5pm (online via Teams)

Assessment criteria:

Provide students with clear information about what they need to do to pass this assessment. Further guidance on writing assessment criteria can be found in the Advance HE CEBE Guide to Writing Assessments Criteria

How to submit your work:

All submissions for this component(s) of this Module must be submitted according to the instructions given in the CRWK file usually at week 5 or 6. If you fail to submit this component as directed (via Moodle site), a mark of 0 will be awarded for the component.

Reassessment Arrangements

In case of failure or non-submission of CRWK, there's no need to worry. You can have a 2nd attempt (re-assessment), but the highest grade you can get is capped to 50%. This usually takes place in March or April.

Problems Completing an Assessment

You may be in a position where you cannot submit your coursework or take an exam due to circumstances that are:

- Unforeseeable - in that you could have no prior knowledge of the event concerned, and
- Unpreventable - in that you could do nothing reasonably in your power to prevent such an event, and
- Expected to have a serious impact on performance.

More information on support available to you can be found via the Mitigating Circumstances webpages (UEL login required) and the Students' Union

Online System Failures

If you experience a problem submitting your work online, you should notify your lecturer/tutor by email immediately.

Deadlines are not extended unless there are significant systems problems. If UEL finds that the issue with the system was significant, you will receive an email notifying you of the issue and that you have been given a 24-hour extension. **If you don't receive any email that specifically states you have been given an extension, then the original deadline has not been changed**.

Best advice: Don't wait until the last minute to submit your assessments electronically.

More information

- A guide to submitting your work through **Turnitin** (UEL login required)
- A guide to viewing and understanding the similarity report in Turnitin (UEL login required)
- **Assessment & Feedback Policy**

Referencing

As a student you will be taught how to write correctly referenced essays. UEL's standard **Harvard referencing** system is from Cite Them Right (UEL login required). Cite them Right is the standard Harvard referencing style at UEL for all Schools, however professional body requirements will take precedence for instance the School of Psychology which uses the APA system.

- MyFeedback – submitting your draft for advice on writing and referencing. (UEL login required)
- Academic skills drop-in sessions (Academic Tutoring). (UEL login required)
- Library guide on citation and referencing (UEL login required)
- Video guide to Getting Started on Zotero (UEL login required)
- Zotero Student Written Guide
- Student Handbook page on Academic Misconduct and Plagiarism (UEL login required)
- MyFeedback – submitting your draft for advice on writing and referencing. (UEL login required)
- Academic skills drop-in sessions (Academic Tutoring). (UEL login required)
- Advice on academic writing and statistics (Academic Tutoring) (UEL login required)

Reading and Resources

Core:

T. White (2015), *Hadoop: The Definitive Guide*. O'Reilly Media, Inc.

Jules S. Damji, Brooke Wenig, Tathagata Das, Denny Lee (2020) Learning Spark, 2nd Edition. ISBN: 9781492050049

M. Zaharia and B. Chambers (2018), *Spark: The Definitive Guide*. O'Reilly Media

Other Resources and Forms:

Bengfort, B. and Kim, J. (2016). *Data Analytics with Hadoop: An Introduction for Data Scientists*. O'Reilly Media.

Bahga, A. and Madisetti, V. (2016). *Big Data Science & Analytics: A Hands-On Approach*. VPT.

Return of Work and Feedback

Feedback is crucial for your learning, and it is an important part of the academic cycle. It tells you what the strengths are of your work, what its weaknesses are and how it can be improved.

Why is Feedback Important?

Feedback is the most effective way to:

- Help you understand how to succeed in your assessments;
- Help you produce better work for the future;

- Signpost you to other resources for assistance.

If you pay attention to feedback, particularly where the same comment is made in several modules, you can use the information to improve.

Where do I get Feedback?

- When a tutor comments on your answers in seminars/lectures/workshops
- General comment on assessment performance in lectures and seminars
- General comment on questions prepared for seminars
- When another student makes comments on your presentation
- When you produce practice questions for a tutor who gives comments
- When you receive written comments on your work submitted either as coursework or exam
- When you look at general feedback on module performance on UEL Direct.
- When you see your Academic Adviser with all your assessment feedback for general advice. You should always do this after each assessment period.

Formal results are ONLY available in UEL Direct and will be published within 8 working days of the Board, where results are formally confirmed. Any other results are provisional / indicative but not approved.

You will receive feedback throughout your this module through the following methods:

- one-to-one or individualised *(i.e. tutorials, conversations with supervisors, or individualised comments on assignments)*
- generic feedback *(i.e. use of rubrics, 'Quickmarks' in Turnitin or standardised forms)*
- peer feedback *(i.e. feedback from other students)*
- informal feedback *(i.e. through in-class discussions or online forums)*
- self-evaluation *(i.e. online checklists or reflective submissions)*
- other please provide details here

(Please delete the above as appropriate)

Feedback and students' marks should be provided within 15 working days of the due date for summative work (i.e. work that counts towards the final course grade) and formative work (i.e. work that is developmental and designed to help you improve).

Whilst feedback will be given on draft/formative work, it shouldn't be assumed that every aspect will be identified.

Student Feedback on Your Experience

UEL values student feedback and we take action based on the feedback we receive.

(Please note any changes you have made to your module this year based on student feedback)

Key Links

- Academic Appeals
- Academic Integrity
- Access and Participation Plan
- Accreditation of Experiential Learning
- Assessment and Feedback Policy
- Bus Timetable
- Complaints procedure (UEL login required)
- Wellbeing Support (UEL login required)
- Disability support
- Attendance and Engagement Policy
- Procedure for problems completing an assessment (UEL login required)
- IT Support (UEL login required)
- Library Archives and Learning Services (UEL login required)
- Manual of General Regulations
- Mentoring (UEL login required)
- Office for Institutional Equity (UEL login required)
- Referencing guidelines (UEL login required)
- Student Engagement, Retention and Success (UEL login required)
- Student Protection Plan
- Suitability Procedure
- Track My Future