COURSE SPECIFICATION DOCUMENT

Academic School/Department:	Science, Innovation & Technology
Programme:	Math & Data Science
***FHEQ* Level:**	6
Course Title:	Advanced Computational Methods in Data Science
Course Code:	MATH 6101
Total Hours:	160
Timetabled Hours:	45
Guided Learning Hours:	15
Independent Learning Hours:	100
Credits:	16 UK CATS credits 8 ECTS credits 4 US credits

Course Description:

This course provides a deep dive into advanced computational techniques used in data science. It is designed to equip students with a solid foundation in the computational methods necessary for processing and analysing large-scale datasets. Students will learn about high-performance computing, optimization algorithms, advanced numerical methods, and specialized techniques for data analysis, visualization, and interpretation. Emphasis will be placed on both theoretical understanding and practical implementation of computational algorithms, including parallel computing, optimization strategies, and their applications in big data analytics and complex problem-solving.

Prerequisites:

70 credits, MATH 5104 Linear Algebra and Multivariable Calculus or

COMP 5102 Cyber Security or SENG 5101 Intermediate Software development with AI.

Aims and Objectives:

Aims:

- To develop students' understanding of advanced computational methods and their applications in Data Science.
- To provide practical experience in implementing algorithms that address the challenges of large-scale data processing and analysis.
- To introduce students to high-performance computing environments and parallelization techniques.
- To cultivate the ability to critically evaluate and select appropriate computational techniques for specific data science problems.
- To explore how computational methods are used in a variety of fields, including machine learning, optimization, and statistical modeling.

Objectives:

- Provide an understanding of advanced data structures and algorithms for high-performance data processing.
- Develop the ability to design and implement efficient computational solutions to large-scale data science problems.
- Explore various optimization algorithms and understand their convergence properties.
- Introduce parallel and distributed computing frameworks and tools.
- Encourage students to analyse and solve complex problems using computationally intensive methods.

Programme Outcomes

L6 AI, AII, BIII, CII, DIII

A detailed list of the programme outcomes is found in the Programme Specification. This is maintained by Registry and located at: https://www.richmond.ac.uk/programme-and-course-specifications/

Learning Outcomes:

By the end of the course, students will be able to:

1. Demonstrate an understanding of key computational methods and algorithms used in Data Science.
2. Critically evaluate different computational approaches and determine their suitability for specific data science tasks.
3. Design and implement efficient algorithms to solve large-scale data processing challenges, with a focus on performance and scalability.
4. Analyse, visualize, and interpret the results of advanced computational methods in a variety of Data Science applications.
5. Implement and optimize algorithms in various programming environments, using languages such as Python, R, and C++ for scientific computing.
6. Engage in problem-solving activities that involve handling large-scale datasets and real-world computational challenges in Data Science.

Indicative Content:

- Data Structures and Algorithms for Large-Scale Data Processing
- Numerical Methods and Linear Algebra
- Optimization Algorithms
- MapReduce and Apache Spark
- Monte Carlo Methods and Simulation
- Computational Methods in Machine Learning
- Big Data Analysis
- Scientific Computing Tools Python (NumPy, SciPy) for large-scale computations
- Ethical & Sustainable Considerations in Computational Data Science

Assessment:

This course conforms to the University Assessment Norms approved at Academic Board and located at: https://www.richmond.ac.uk/university-policies/

Teaching Methodology:

This course will be delivered face to face through a combination of lectures and interactive sessions. In addition to classroom activities, there are guided learning elements that are tutor led and arranged through Blackboard. These activities can be asynchronous online sessions, flipped classrooms, set readings with discussion boards or set guest lectures for example. Set activities are monitored by the instructor to ascertain student engagement. Students are encouraged to prepare for class and to play an active part, to raise questions, following-up ideas and interact with a wide range of provided material.

Indicative Text(s):

Bruce, P. and Bruce, A. (2020) *Practical Statistics for Data Scientists, 2e: 50+ Essential Concepts Using R and Python*. Sebastopol, CA: O'Reilly Media

Kleppmann, M. (2017) *Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems*. Sebastopol, CA: O'Reilly Media

Liao, S. (2020) *Ethics of artificial intelligence.* Oxford: Oxford University Press.

Schäfer, M. and van Es, K. (2017) *Datafied society: studying culture through data*. Amsterdam: Amsterdam University Press.

Vaughan, D. (2020) *Analytical Skills for AI and Data Science: Building Skills for an AI-driven Enterprise.* Sebastopol, CA: O'Reilly Media.

Journals

The Journal of data Science. Available at**:** Journal of data science

Intelligent Data Analysis. Available at: Intelligent Data Analysis

Online Resources:

Python Documentation (NumPy, SciPy, Dask).

See syllabus for complete reading list.

Change Log for this CSD:

Nature of Change	Date Approved & Approval Body (School or LTPC)	Change Actioned by Academic Registry
First Edition	Nov 2024	