Swansea University
Prifysgol Abertawe

FACULTY OF SCIENCE AND ENGINEERING

POSTGRADUATE STUDENT HANDBOOK

MSC (FHEQ LEVEL 7)

APPLIED DATA SCIENCE

DEGREE PROGRAMMES

SUBJECT SPECIFIC
PART TWO OF TWO
MODULE AND COURSE STRUCTURE
2023-24

DISCLAIMER

The Faculty of Science and Engineering has made all reasonable efforts to ensure that the information contained within this publication is accurate and up-to-date when published but can accept no responsibility for any errors or omissions.

The Faculty of Science and Engineering reserves the right to revise, alter or discontinue degree programmes or modules and to amend regulations and procedures at any time, but every effort will be made to notify interested parties.

It should be noted that not every module listed in this handbook may be available every year, and changes may be made to the details of the modules. You are advised to contact the Faculty of Science and Engineering directly if you require further information.

The 23-24 academic year begins on 25 September 2023

Full term dates can be found here

DATES OF 23-24 TERMS

25 September 2023 – 15 December 2023

8 January 2024 – 22 March 2024

15 April 2024 – 07 June 2024

SEMESTER 1

25 September 2023 – 29 January 2024

SEMESTER 2

29 January 2024 – 07 June 2024

SUMMER

10 June 2024 – 20 September 2024

IMPORTANT

Swansea University and the Faculty of Science of Engineering takes any form of **academic misconduct** very seriously. In order to maintain academic integrity and ensure that the quality of an Award from Swansea University is not diminished, it is important to ensure that all students are judged on their ability. No student should have an unfair advantage over another as a result of academic misconduct - whether this is in the form of **Plagiarism**, **Collusion** or **Commissioning**.

It is important that you are aware of the **guidelines** governing Academic Misconduct within the University/Faculty of Science and Engineering and the possible implications. The Faculty of Science and Engineering will not take intent into consideration and in relation to an allegation of academic misconduct - there can be no defence that the offence was committed unintentionally or accidentally.

Please ensure that you read the University webpages covering the topic – procedural guidance here and further information here. You should also read the Faculty Part One handbook fully, in particular the pages that concern Academic Misconduct/Academic Integrity.

Welcome to the Faculty of Science and Engineering!

Whether you are a new or a returning student, we could not be happier to be on this journey with you.

At Swansea University and in the Faculty of Science and Engineering, we believe in working in partnership with students. We work hard to break down barriers and value the contribution of everyone.

Our goal is an inclusive community where everyone is respected, and everyone's contributions are valued. Always feel free to talk to academic, technical and administrative staff, administrators - I'm sure you will find many friendly helping hands ready to assist you. And make the most of living and working alongside your fellow students.

During your time with us, please learn, create, collaborate, and most of all – enjoy yourself!

Professor David Smith
Pro-Vice-Chancellor and Executive Dean
Faculty of Science and Engineering

Faculty of Science and Engineering	
Pro-Vice-Chancellor and Executive Dean	Professor David Smith
Director of Faculty Operations	Mrs Ruth Bunting
Associate Dean – Student Learning and Experience (SLE)	Professor Laura Roberts
School of Mathematics and Computer Science	
Head of School	Professor Elaine Crooks
School Education Lead	Dr Neal Harman
Head of Mathematics	Professor Vitaly Moroz
Mathematics Programme Director	Dr Kristian Evans
Year Coordinators	Year 0 – Dr Zeev Sobol Year 1 – Dr Nelly Villamizar Year 2 – Professor Chenggui Yuan Year 3 – Professor Grigory Garkusha Year 4 – Professor Grigory Garkusha MSc – Dr Guo Liu

STUDENT SUPPORT

The Faculty of Science and Engineering has two **Reception** areas - Engineering Central (Bay Campus) and Wallace 223c (Singleton Park Campus).

Standard Reception opening hours are Monday-Friday 8.30am-4pm.

The **Student Support Team** provides dedicated and professional support to all students in the Faculty of Science and Engineering. Should you require assistance, have any questions, be unsure what to do or are experiencing difficulties with your studies or in your personal life, our team can offer direct help and advice, plus signpost you to further sources of support within the University. There are lots of ways to get information and contact the team:

Email: studentsupport-scienceengineering@swansea.ac.uk (Monday–Friday, 9am–5pm)

Call: +44 (0) 1792 295514 (Monday-Friday, 10am–12pm, 2–4pm).

Zoom: By appointment. Students can email, and if appropriate we will share a link to our Zoom calendar for students to select a date/time to meet.

The current student **webpages** also contain useful information and links to other resources:

https://myuni.swansea.ac.uk/fse/

READING LISTS

Reading lists for each module are available on the course Canvas page and are also accessible via http://ifindreading.swan.ac.uk/. We've removed reading lists from the 23-24 handbooks to ensure that you have access to the most up-to-date versions.
We do not expect you to purchase textbooks, unless it is a specified key text for the course.

THE DIFFERENCE BETWEEN COMPULSORY AND CORE MODULES

Compulsory modules must be **pursued** by a student.

Core modules must not only be **pursued**, but also **passed** before a student can proceed to the next level of study or qualify for an award. Failures in core modules must be redeemed.

Further information can be found under "Modular Terminology" on the following link - https://myuni.swansea.ac.uk/academic-life/academic-regulations/taught-guidance/essential-

info-taught-students/your-programme-explained/

MSc (FHEQ Level 7) 2023/24
Applied Data Science (MSc)
MSc Applied Data Science

Coordinator: Dr G Liu

Semester 1 Modules	Semester 2 Modules
CSCM53 **Computer System Concepts** **15 Credits** **Dr S Yang**	**CSCM35** **Big Data and Data Mining** **15 Credits** **Dr S Yang**
CSCM59 **Relational and Object-Oriented Database Systems** **15 Credits** **Dr W Macinnes**	**CSCM37** **Data Visualisation** **15 Credits** **Dr B Mora/Dr JF Maestre Avila**
MA-M06 **Mathematics for Data Science** **15 Credits** **Prof V Moroz/Prof T Brzezinski**	**MA-M07** **Further Mathematics for Data Science** **15 Credits** **Dr NY Villamizar/Dr AY Pachon**
MA-M16 **Probability and Statistics for Data Science** **15 Credits** **Dr DL Finkelshtein/Dr K Evans**	**MA-M17** **Modelling and Machine Learning** **15 Credits** **Prof B Lucini/Prof GG Powathil**
Dissertation	
MA-D02 **Applied Data Science Masters Dissertation** **60 Credits** **Prof C Yuan**	
Total 180 Credits	

CSCM35 Big Data and Data Mining

Credits: 15 Session: 2023/24 January-June

Pre-requisite Modules:

Co-requisite Modules:

Lecturer(s): Dr S Yang

Format: 20 hours lectures, 10 hours lab

Delivery Method: Primarily on campus

Module Aims: This module introduces students to the fundamental topics of data mining, including data pre-processing techniques, applied probability and statistics, data mining algorithms (incl. associate rule, classification, clustering, outlier detection and probabilistic graphical model), and big data frameworks.

Module Content: Basic knowledge in machine learning and mathematics are required, where we students are strongly encouraged to take CSCM45 and CSCM70 in the first semester. For those who have not taken those two modules, we will revisit some key concepts in the first three weeks, i.e. probability and statistics. Practical work will be done through the medium level of Python programming. You will also get to discover key scientific libraries of Python, i.e. NumPy, SciPy, Scikit-Learn, Matplotlib, Tensorflow, Hadoop, Spark. Classes in the first 3 weeks will include support time for Python programming and the use of these libraries.

Course Overview and Python Programming for Data Science
Mathematics Background: Numerical Analysis, Applied Probability and Statistics
Frequent Pattern, Association, Correlations
Naïve Bayes Classifier, Quantitative Evaluation
Decision Tree, Random Forest, AdaBoost
Maximum Likelihood Estimation, Expectation Maximisation
Clustering, DBSCAN, High Dimensional Data Analysis
Text Data Analysis, Word2Vec, Skip-Through, CBOW
Time-Series Data Analysis, Regression, Hidden Markov Model
Scalability and Efficiency Big Data Analysis

Intended Learning Outcomes: After completing this module students will:
- be able to manipulate complex, large, heterogeneous datasets, from storage to processing
- be able to extract information from complex, large, heterogeneous datasets
- be able to critically evaluate and select data mining algorithms and techniques, and be able to apply them in real world applications.

Assessment:	Coursework 1 (15%) Coursework 2 (15%) Examination (70%)
Resit Assessment:	Examination (Resit instrument) (100%)

Assessment Description: Coursework 1 (February): Examination of Python programming in Data Science
Coursework 2 (March): Examination of Data Mining and Machine Learning algorithm
Exam (May/June): Standard Format Computer Science Exam

Moderation approach to main assessment: Moderation by sampling of the cohort

Assessment Feedback: Grades, individual and collective written feedback for coursework.

Failure Redemption: Redemption of failure via resit instrument.

Additional Notes: Available to visiting and exchange students. Basic knowledge in machine learning and mathematics are required and Computer Science students are strongly encouraged to take CSCM45 and CSCM70 in the first semester. For those who have not taken those two modules, some key concepts will be revisited in the first three weeks, i.e. probability and statistics. NOTE: students on the MSc Applied Data Science will have covered the relevant material in MA-M06 and MA-M16

CSCM37 Data Visualisation

Credits: 15 Session: 2023/24 January-June

Pre-requisite Modules:

Co-requisite Modules:

Lecturer(s): Dr B Mora, Dr JF Maestre Avila

Format: 20 hours lectures,
10 hours practicals

Delivery Method: On-campus and lab sessions.

Module Aims: Data Visualisation is concerned with the automatic or semi-automatic generation of digital images that depict data in a meaningful way(s). It is a relatively new field of computer science that is rapidly evolving and expanding. It is also very application oriented, i.e., real tools are built in order to help scientists from other disciplines.

Module Content: Introductory topics include: purposes and goals of visualisation, applications, challenges, the visualisation pipeline, sources of data: data dimensionality, data types, and grid types.

Information visualisation topics include: abstract data, hierarchical data, tree maps, cone trees, focus and context techniques, hyperbolic trees graphs and graph layouts, multi-dimensional data, scatter plots, scatter plot matrices, icons, parallel coordinates, interaction techniques, linking and brushing.

Volume visualisation topics include: slicing, surface vs. volume rendering, transfer functions, interpolation schemes, direct volume visualisation, ray casting, shear-warp factorisation, image order vs. object order algorithms, gradients, filtering, interpolation, and isosurfacing.

Flow visualisation topics include: simulation, measured, and analytical data, steady and time-dependent (unsteady) flow, direct and indirect flow visualisation, applications, hedgehog plots, vector glyphs, numerical integration schemes, streamlines, streamline placement, geometric flow visualisation techniques, line integral convolution (LIC), texture-based techniques, feature-based flow visualisation.

Intended Learning Outcomes: Students will be able to:
- identify problems that can be addressed with visualisation.
- comprehensively explain data visualisation techniques and be able to critically appraise their suitability to particular situations.
- choose, evaluate and apply visualisation techniques to effectively reveal insights into complex and potentially-incomplete data.

Assessment: Examination 1 (60%)
Coursework 1 (20%)
Coursework 2 (20%)

Resit Assessment: Examination (Resit instrument) (100%)

Assessment Description: Standard Computer Science format unseen examination, duration 2hrs.
Two practical courseworks.

Moderation approach to main assessment: Moderation by sampling of the cohort

Assessment Feedback: Outline solutions provided along with group and individual analytical feedback for courseworks.
Examination feedback summarising strengths and weaknesses of the class.

Failure Redemption: Resit examination and/or resubmit coursework(s) as appropriate.

Additional Notes: Available to visiting and exchange students.

CSCM53 Computer System Concepts

Credits: 15 Session: 2023/24 September-January

Pre-requisite Modules:

Co-requisite Modules:

Lecturer(s): Dr S Yang

Format: 30 Lectures

Delivery Method: On-campus and lab sessions.

Module Aims: This module gives an overview of some of the main principles underlying computers and computing from both a theoretical and an applied point of view. It includes a brief history of computers and software, an introduction to the representation of data and the basic components of a computer, the basic features of operating systems, file systems, computer networks, the world wide web, and some basic issues of computer security. A brief discussion on the use of formal methods in specifying computer systems is also given.

Module Content: Brief history of computers and software
Binary values and number systems
Data representation
Logic, gates and circuits
Computing components
Low level programming
Operating system concepts
File systems and directories
Computer networks
The world wide web
Overview of computer security
Introduction to formal methods

Intended Learning Outcomes: After completing this module students will have demonstrated:
- an understanding of how software and hardware interact in a computer system.
- an understanding of the principles behind modern computer architecture, operating systems, file systems, networks and the world wide web.
- an understanding of the critical issues in computer security.
- an ability to apply the fundamental mathematical tools of Computer Science and be able to apply them to writing formal specifications and documentation.

Assessment:	Examination 1 (80%)
	In class test (non-invigilated) (20%)
Resit Assessment:	Examination (Resit instrument) (100%)

Assessment Description: Examination: duration 2hrs.
Weekly in-class Canvas quizzes.

Moderation approach to main assessment: Moderation by sampling of the cohort

Assessment Feedback: Outline solutions provided along with analytical individual feedback for coursework's. Examination feedback summarising strengths and weaknesses of the class.

Failure Redemption: Resit examination

Additional Notes:

Only available as part of the MSc Computer Science and MSc Applied Data Science degree programmes.

CSCM59 Relational and Object-Oriented Database Systems

Credits: 15 Session: 2023/24 September-January

Pre-requisite Modules:

Co-requisite Modules:

Lecturer(s): Dr W Macinnes

Format: 20 lectures, 2 x 3 hours lab class,
4 hours problem sessions

Delivery Method: Primarily on campus

Module Aims: This module gives an appreciation of of the complexity of real-world databases. It considers some of the problems that can occor in multi-user, multi-transactions situations. It discusses relational and object-oriented databases and covers their design and implementation. Distributed databases and databases linked to the web will also be discussed, as will data warehousing and data mining. Students will gain practical experience in designing and implementing a database.

Module Content: A review of the nature of data and databases and an overview of database management and
database system architecture.
Data models: relational databases, object databases.
Relational databases: the structure of the relational model, integrity constraints, relational algebra and calculus, normalisation.
Transaction management, data security and recovery, optimisation, distributed databases, concurrency control.
Object-oriented databases, type inheritance, active databases, temporal databases, logic-based databases.
Data warehouses and data mining, data visualisation.
Web technology and databases.

Intended Learning Outcomes: After completing this module students will have demonstrated:
- a comprehensive understanding of the principles of relational databases and object-oriented databases.
- the ability design and implement databases according to best practice.
- the ability to normalise a database and explain why it is necessary.
- the ability to explain the additional problems associated with distributed databases and the need to ensure integrity and how to control concurrency.
- the ability to explain how data warehouses operate and the principles of data mining.
- how to link databases to web applications.

Assessment: Examination 1 (70%)
Coursework 1 (20%)
Laboratory work (10%)

Resit Assessment: Examination (Resit instrument) (100%)

Assessment Description: Standard Computer Science format unseen examination, duration 2hrs. All questions should be attempted.
Practical assignment. Students will be required to build and interrogate a database according to specifications.
Laboratory sessions. Students will demonstrate key skills.

Moderation approach to main assessment: Moderation by sampling of the cohort

Assessment Feedback: Outline solutions provided along with group and individual analytical feedback for courseworks.
Examination feedback summarising strengths and weaknesses of the class.
Individual feedback on submissions from lecturer and/or demonstrators in laboratory sessions.

Failure Redemption: Resit exam or resubmit assignment as appropriate

Additional Notes: Available to visiting and exchange students

MA-D02 Applied Data Science Masters Dissertation

Credits: 60 Session: 2023/24 September-June

Pre-requisite Modules:

Co-requisite Modules:

Lecturer(s): Prof C Yuan

Format: 12 - mixture of introductory lectures and individual supervisions.

Delivery Method: Individual project supervision. Topics are usually suggested by supervisors, however there are opportunities to modify/negotiate the suggested topics. Students can discuss topics with their supervisor to try to tailor their dissertation to their interests.

Students are supported by supervision meetings (a minimum of 4 and a maximum of 8) and introductory lectures.

Module Aims: A research project selected from an area of Applied Data Science. It will enable the students to develop an enquiring, analytical critical and creative approach to problem identification and solution. The project will typically focus on a subject area related to one of the taught modules in the MSc scheme.

Module Content: The student will study an area of Applied Data Science related to the content of the taught part of the MSc scheme, drawing together material from different published sources, and their own investigations, to product a substantial written report on the topic.

Intended Learning Outcomes: At the end of this module, the student will be able to:
plan and undertake a significant project with a high degree of independence,
synthesize information and ideas;
independently evaluate alternative approaches to a given problem;
accurately report on and evaluate their own work, and work of others;
perform and present research in an ethical manner;
communicate complex mathematical, statistical or computational ideas and concepts.

Assessment: Dissertation (100%)

Assessment Description: Written dissertation, including presentation, subject to the regulations for an MSc dissertation.

Moderation approach to main assessment: Universal Non-Blind Double Marking of the whole cohort

Assessment Feedback: Students will receive ongoing feedback during the meetings with their supervisor. The official result of the MSc dissertation will be communicated to the student by the University. The student can receive individual feedback on their dissertation from their supervisor, informed by the views of both markers.

Failure Redemption: Resubmission of the dissertation, in accordance with the University regulations.

Additional Notes: Dissertation for the MSc Applied Data Science programme. Only available to students on the MSc Applied Data Science programme.

Each student will be assigned a member of staff as a supervisor who will help to choose the topic to be studied, and guide the student. Each student will be required to produce a word-processed thesis of 15,000-20,000 words, based on topics covered in the taught part of the MSc scheme, and on published research in the field.

Students will also be required to present their work.

MA-M06 Mathematics for Data Science

Credits: 15 Session: 2023/24 September-January

Pre-requisite Modules:

Co-requisite Modules:

Lecturer(s): Prof V Moroz, Prof T Brzezinski

Format: There be weekly delivery, with each week have 3 lectures and 1 examples class.

Delivery Method: The module will be delivered on Bay Campus, with a traditional mix of lectures and example classes underpinned with weekly assessments of a formative/summative mix.

Module Aims: The module provides a basic introduction to mathematical foundations for data science and machine learning. It covers essential concepts from linear algebra and calculus that are of direct utility to analysis of large data sets. The module includes basic problems in one-dimensional Calculus and Linear Algebra that illustrate concepts of differentiation and integration of functions of one variable, operations with matrices and methods for solving linear systems of equations. It also highlights the use of Python in the numerical solution of these problems, as well as in symbolic computation.

Module Content: - Continuous functions;
- Derivatives, finding maxima and minima;
- Integration, Fundamental Theorem of Calculus;
- Numerical integration;
- Vectors: real vector spaces and the visualization of vectors;
- Matrices: forming matrices; linear combination as a matrix-vector product; examples;
- Linear mappings: lineal functionals; inner product; matrix-matrix product;
- Matrix decompositions;
- Determinant and trace;
- Eigenvalues and eigenvectors;
- Eigen-decomposition and diagonalisation.

Intended Learning Outcomes: Upon completion of this module students will be able to:
- Explain and analyse what a continuous function is;
- Demonstrate knowledge of basic examples of polynomial, rational and trigonometric functions;
- Calculate 1st and higher order derivatives of a function;
- Demonstrate an efficient use of calculus algorithms for finding maxima and minima of functions;
- Demonstrate an understanding of the concept of integration and its connection with differentiation;
- Calculate integrals of the basic polynomial and trigonometric functions both analytically and using symbolic computational tools;
- Demonstrate an understanding of the numerical integration methods and to be able to implement numerical integration in programming;
- Identify the basic tools from linear algebra and their potential as descriptors of large sets of data.

Assessment:	Examination (60%)
	Assignment 1 (40%)
Resit Assessment:	Examination (Resit instrument) (100%)

Assessment Description: Component 1 is a written, closed-book examination at the end of the module. Component 2 is formed of a number of coursework assignments during the semester.

Moderation approach to main assessment: Moderation of the entire cohort as Check or Audit

Assessment Feedback: For the homework assignments, students will receive feedback in the form of marks, model solutions, overall feedback on the cohort performance, and some individual comments on their work.
For the exam, students will receive feedback in the form of marks and overall feedback on the cohort performance.
Further, individualised feedback, can be provided upon request.

Failure Redemption: Supplementary Examination.

Additional Notes: Only available to MSc Applied Data Science students.

MA-M07 Further Mathematics for Data Science

Credits: 15 Session: 2023/24 January-June

Pre-requisite Modules:

Co-requisite Modules:

Lecturer(s): Dr NY Villamizar, Dr AY Pachon

Format: There be weekly delivery, with each week have 3 lectures and 1 examples class.

Delivery Method: The module will be delivered on Bay Campus, with a traditional mix of lectures and example classes underpinned with weekly assessments of a formative/summative mix.

Module Aims: The module develops mathematical foundations for data science and machine learning. The module covers basic problems of Vector Calculus, Linear Algebra and Optimization including concepts of partial derivatives, gradient and Hessian, eigenvalues and eigenvectors of matrices and their use in optimisation problems, method of gradient descent and its application in Machine Learning. It also highlights the use of Python in the numerical solution of these problems, as well as in symbolic computation.

Module Content: - Vectors and vector fields, functions of several variables;
- Partial derivatives, critical points;
- Finding maxima and minima, Hessian test;
- Method of gradient descent;
- Vector spaces: formal rules; vector subspaces; linear dependence; basis and dimension;
- The Singular Value Decomposition (SVD): orthogonal matrices; change of coordinates; SVD solution of linear algebra problems;
- Least squares: projection; Gram-Schmidt orthonormalization; QR solution of linear algebra problems; general orthogonal bases; least-squares problems.

Intended Learning Outcomes: Learning Outcomes:
At the end of this module students should be able to:
- Explain and critically analyse ideas of Vector Calculus, such as vector, vector field, function of several variable;
- Demonstrate an ablity to calculate 1st and 2nd order partial derivatives of a function;
- Demonstrate a comprehensive understanding of gradient vector fields;
- Systematically apply variational algorithms for finding maxima and minima of functions;
- Demonstrate an in-depth understanding of the method of gradient descent;
- Develop and demonstrate an understanding of methods for constructing low-dimensional approximations of data sets, including projections, least squares and the singular value decomposition.

Assessment:	Examination (60%) Coursework 1 (10%) Coursework 2 (10%) Coursework 3 (10%) Coursework 4 (10%)
Resit Assessment:	Examination (Resit instrument) (100%)

Assessment Description: Examination: A closed book examination to take place at the end of the module.
Courseworks 1-4: This coursework will develop students' skills in problem solving, and developing and writing logical arguments.

Moderation approach to main assessment: Moderation of the entire cohort as Check or Audit

Assessment Feedback: For the homework assignments, students will receive feedback in the form of marks, model solutions, overall feedback on the cohort performance, and some individual comments on their work.
For the exam, students will receive feedback in the form of marks and overall feedback on the cohort performance.
Further, individualised feedback, can be provided upon request.

Failure Redemption: Supplementary Examination.

Additional Notes: Only available to MSc Applied Data Science students.

MA-M16 Probability and Statistics for Data Science

Credits: 15 Session: 2023/24 September-January

Pre-requisite Modules:

Co-requisite Modules:

Lecturer(s): Dr DL Finkelshtein, Dr K Evans

Format: There be weekly delivery, with each week have 3 lectures and 1 lab class.

Delivery Method: The module will be delivered on Bay Campus, with a traditional mix of lectures, examples and lab classes underpinned with weekly assessments of a formative/summative mix.

Module Aims: This module serves as an introduction to the concepts of probability and statistics required for Data Science. Both the theoretical concepts and practical examples will be explored throughout the module. Particular attention will be paid to applications using a programming language.

Module Content: - Basic concepts in probability;
- Rules of probability;
- Discrete probability distributions;
- Basic concepts in statistics;
- Linear regression and correlation;
- Logistic regression;
- Continuous probability distributions;
- Maximum likelihood estimation;
- Law of large numbers and the central limit theorem;
- Hypothesis testing;
- Data reduction;
- Time series.

Intended Learning Outcomes: At the end of this module students will be able to:
1) demonstrate a comprehensive knowledge of the basic concepts and rules of probability;
2) demonstrate a comprehensive understanding of discrete and continuous distributions;
3) demonstrate a comprehensive understanding of the basic concepts of statistics;
4) fit and systematically analyse models for linear and logistic regression;
5) demonstrate an understanding of maximum likelihood estimation, the law of large numbers and the central limit theorem;
6) analyse data using hypothesis testing, data reduction and time series.

Assessment:	Examination (60%) Assignment 1 (20%) Assignment 2 (20%)
Resit Assessment:	Examination (Resit instrument) (100%)

Assessment Description: Component 1 is a written, closed-book examination at the end of the module.
Component 2 is formed of a number of coursework assignments during the semester.
Component 3 is formed of a lab test during the semester.

Moderation approach to main assessment: Moderation of the entire cohort as Check or Audit

Assessment Feedback: For the homework assignments, students will receive feedback in the form of marks, model solutions, overall feedback on the cohort performance, and some individual comments on their work.
For the lab test, students will receive feedback in the form of marks, model solutions, overall feedback on the cohort performance, and some individual comments on their work.
For the exam, students will receive feedback in the form of marks and overall feedback on the cohort performance.
Further, individualised feedback, can be provided upon request.

Failure Redemption: Supplementary Examination.

Additional Notes: Only available to MSc Applied Data Science students.

MA-M17 Modelling and Machine Learning

Credits: 15 Session: 2023/24 January-June

Pre-requisite Modules:

Co-requisite Modules:

Lecturer(s): Prof B Lucini, Prof GG Powathil

Format: There be weekly delivery, with each week have 3 lectures and 1 lab class.

Delivery Method: The module will be delivered on Bay Campus, with a traditional mix of lectures, examples and lab classes underpinned with weekly assessments of a formative/summative mix.

Module Aims: The module provides a basic introduction to mathematical modelling and computational methods. The module covers basic problems in computation including error analysis and methods for solving equations. It also highlights the use of Python in the numerical solution of problems, as well as symbolic computation. The module will also introduce some of the most widely used and state-of-the-art artificial intelligence and machine learning techniques and their applications to big data problems.

Module Content: - Introduction to modelling;
- Introduction to computational methods;
- Numerical and symbolic computation;
- Representing and approximating numbers. Algorithms;
- Evaluating functions;
- Numerical differentiation and integration;
- Numerical Solutions to ODE;
- Numerical optimisation methods;
- Introduction to Machine Learning;
- Machine Learning and optimisation;
- Classification and regression in Machine Learning: linear regression, logistic regression, support vector machine, linear perceptron.
- Unsupervised learning methods: Principal Component Analysis, clustering methods.
- Introduction to Deep Learning.

Intended Learning Outcomes: Learning Outcomes:
At the end of this module students should be able to:
- explain and critically analyse how an error can propagate through a calculation;
- select and implement an appropriate algorithm for a given problem;
- critically analyse and interpret mathematical models and underlying dynamics;
- interpret and implement relevant algorithms;
- describe, explain, and critique the fundamental techniques of analysing complex and heterogeneous data;
- describe and explain machine learning techniques and their applications to problems involving collected or generated data sets, including big data problems;
- discuss and contrast both conventional and state-of-the-art machine learning techniques and be able to choose the best method for a given problem;
- demonstrate a deep understanding of the mathematics of machine learning, the regime of applicability of different machine learning algorithm and potential limitations in specific scenarios;
- implement and apply machine learning techniques to synthesise solutions;
- analyse data sets, gathering insights on the underlying production mechanism in order to optimise the latter or solve problems related to the production process.

Assessment:	Examination (60%) Coursework 1 (6%) Coursework 2 (7%) Coursework 3 (7%) Laboratory work (20%)
Resit Assessment:	Examination (Resit instrument) (100%)

Assessment Description: Examination: A closed book examination to take place at the end of the module. Courseworks 1-3: This coursework will develop skills in problem solving, applying techniques to real world problems and understanding the use of computers to investigate problems. Lab Assessment: Computing based controlled assessment to assess skills in the use of computers to investigate and analyse real world problems.
Moderation approach to main assessment: Moderation of the entire cohort as Check or Audit
Assessment Feedback: For the homework assignments, students will receive feedback in the form of marks, model solutions, overall feedback on the cohort performance, and some individual comments on their work. For the lab test, students will receive feedback in the form of marks, model solutions, overall feedback on the cohort performance, and some individual comments on their work. For the exam, students will receive feedback in the form of marks and overall feedback on the cohort performance. Further, individualised feedback, can be provided upon request.
Failure Redemption: Supplementary Examination.
Additional Notes: Only available to MSc Applied Data Science students.