Module code: MOD007892	**Version:** 3 **Date Amended:** 12/Jul/2023

1. Module Title

Principles of Data Mining and Machine Learning

2a. Module Leader

Faraz Janan

2b. School

School of Computing and Information Sciences

2c. Faculty

Faculty of Science and Engineering

3a. Level

7

3b. Module Type

Standard (fine graded)

4a. Credits

30

4b. Study Hours

300

5. Restrictions

Type	Module Code	Module Name	Condition
Pre-requisites:	None		
Co-requisites:	None		
Exclusions:	None		
Courses to which this module is restricted:			

LEARNING, TEACHING AND ASSESSMENT INFORMATION

6a. Module Description

This module presents an insight into the principles of data mining and machine learning. It equips you with essential skills in analysing data and drawing actionable insights for informed decision-making, leveraging mathematical and statistical models to tackle real-world problems. This module is designed to benefit students with no prior exposure to data mining and machine learning topics, by covering the fundamental principles and core concepts of data mining and the flow of the data in the machine learning pipeline. You will learn how to start looking at data from the perspective of the data scientist and apply machine learning techniques to various types of data with the goal of extracting valuable intelligence. You will explore and learn various data mining and machine learning tools and algorithms through real-world case studies from various industrial domains such as healthcare, finance, retail and IT. An in-depth understanding of various probabilistic and statistical models along with supervised and unsupervised machine learning techniques, including regression, ensemble learning, support vector machines, tree-based methods and the latest frontiers of machine learning, such as neural networks will be covered in this module.

6b. Outline Content

This module covers a wide spectrum of topics, ranging from preliminary core data mining concepts to state- of-the-art deep artificial neural networks. The module starts with a comprehensive overview of the core concepts and techniques in data mining and machine learning to reinforce students' understanding of the fundamentals. Statistical analysis techniques such as mathematical moments, correlation, regression, analysis of variance will be covered in this module, allows students to better grasp the underlying statistical concepts in data mining and machine learning. Additionally, students will be familiarized with several supervised and unsupervised machine learning techniques including logistic regression, decision trees, support vector machines, k-nearest neighbours and ensemble techniques. This module also covers the basics of optimization and reinforcement learning using walkthrough real-world examples and case studies. Finally, this module presents a comprehensive insight into artificial neural networks (ANN) and deep learning, the cutting edge of the technology in the field of data mining and machine learning.

6c. Key Texts/Literature

The reading list to support this module is available at: https://readinglists.aru.ac.uk/

6d. Specialist Learning Resources

Computer with a processor speed of 1.8GHz (or better), at least 8GB RAM and 256MB hard disk. Access to a programming environment including Anaconda and Pycharm.

7. Learning Outcomes (threshold standards)		
No.	Type	On successful completion of this module the student will be expected to be able to:
1	Knowledge and Understanding	Demonstrate in-depth understanding of the main machine learning and knowledge discovery approaches.
2	Knowledge and Understanding	Demonstrate operational knowledge of data mining and machine learning principles, including statistical analysis, supervised, unsupervised learning techniques and artificial neural networks.
3	Intellectual, practical, affective and transferrable skills	Implement practical data mining software solutions based on real-world case studies from various businesses and industrial domains such as healthcare, finance, retail and IT.
4	Intellectual, practical, affective and transferrable skills	Design and develop end-to-end data mining and machine learning pipelines, including cleansing and normalization, data analysis, modelling and evaluation.

8a. Module Occurrence to which this MDF Refers				
Year	Occurrence	Period	Location	Mode of Delivery
2025/6	ZZF	Template For Face To Face Learning Delivery		Face to Face

8b. Learning Activities for the above Module Occurrence			
Learning Activities	Hours	Learning Outcomes	Details of Duration, frequency and other comments
Lectures	36	1,2,3,4	3 hours lecture per week
Other teacher managed learning	24	1,2,3,4	2 hours tutorial per week
Student managed learning	240	1,2,3,4	Self-directed learning
TOTAL:	300		

9. Assessment for the above Module Occurrence					
Assessment No.	Assessment Method	Learning Outcomes	Weighting (%)	Fine Grade or Pass/Fail	Qualifying Mark (%)
010	Coursework	1-4	100 (%)	Fine Grade	30 (%)
2 Components - Project (2500 words equivalent and In-Class Test (75 minutes)					

Assessment components for Element 010				
Component No.	Assessment Title	Submission Method	Weighting (%)	Components needed for Mark Calculation?
010/1	Project report (2500 words equivalent)	Canvas	70 (%)	All
010/2	In-class test (75 minutes)	Scheduled Activity: Timetabled assessment task	30 (%)	

In order to pass this module, students are required to achieve an overall mark of 40% (for modules at levels 3, 4, 5 and 6) or 50% (for modules at level 7*).

In addition, students are required to:

(a) achieve the qualifying mark for each element of fine graded assessment as specified above

(b) pass any pass/fail elements

[* the pass mark of 50% applies for all module occurrences from the academic year 2024/25 – see Section 3a of this MDF to check the level of the module and Section 8a of this MDF to check the academic year]