

## MODULE DESCRIPTOR

### Module Title

Data Mining

Reference	CMM510	Version	9
Created	February 2022	SCQF Level	SCQF 11
Approved	April 2005	SCQF Points	15
Amended	July 2022	ECTS Points	7.5

### Aims of Module

To provide students with an understanding of the main principles underlying Data Mining and Machine Learning techniques and the ability to apply current Data Mining and Machine Learning tools to real datasets.

### Learning Outcomes for Module

On completion of this module, students are expected to be able to:

- 1 Critically discuss, compare and contrast the advantages and disadvantages of applying a specific data mining technique to a given learning task.
- 2 Use industry standard tools to develop a data mining application tailored to a given learning task and evaluate the results obtained.
- 3 Effectively interpret the results of learning through an understanding of the strengths and limitations of data mining technology and the selection of an appropriate evaluation technique.
- 4 Demonstrate knowledge of the state-of-the-art in data mining.

### Indicative Module Content

Data mining concepts. Implementation of fundamental learning approaches to classification, regression, clustering and association rules models. Bias-variance trade-off. Incorporating domain knowledge in learning. Advanced techniques for evaluating learned concepts. Calculation of confidence intervals for predictive performance. Comparison of data mining schemes. Boosting, bagging and stacking techniques. Applications. Legal, ethical, social and professional issues in data mining.

### Module Delivery

This is a lecture based course, supplemented with laboratory sessions, where industry standard data mining software is applied to varied learning tasks and tutorials where additional understanding is gained through practical exercises which supplement the lectures.

**Indicative Student Workload**

	Full Time	Part Time
Contact Hours	30	30
Non-Contact Hours	120	120
Placement/Work-Based Learning Experience [Notional] Hours	N/A	N/A
TOTAL	150	150
<i>Actual Placement hours for professional, statutory or regulatory body</i>		

**ASSESSMENT PLAN**

If a major/minor model is used and box is ticked, % weightings below are indicative only.

**Component 1**

Type:	Practical Exam	Weighting:	100%	Outcomes Assessed:	1, 2, 3, 4
Description:	Practical examination applying data mining techniques to a given dataset.				

**MODULE PERFORMANCE DESCRIPTOR****Explanatory Text**

To achieve a pass in this module requires a minimum of grade D in Assessment Component 1

Module Grade	Minimum Requirements to achieve Module Grade:
<b>A</b>	A in Component 1
<b>B</b>	B in Component 1
<b>C</b>	C in Component 1
<b>D</b>	D in Component 1
<b>E</b>	E in Component 1
<b>F</b>	F in Component 1
<b>NS</b>	Non-submission of work by published deadline or non-attendance for examination

**Module Requirements**

Prerequisites for Module	None except for course entry requirements.
Corequisites for module	None.
Precluded Modules	None.

**INDICATIVE BIBLIOGRAPHY**

- 1 OLSON, DL. (2019), Descriptive Data Mining, 2nd Edition. Springer Nature.
- 2 WITTEN, IH. FRANK, E. HALL, MA, PAL, C. (2017) Data Mining - Practical Machine Learning Tools and Techniques, 4th ed. Morgan Kaufman.
- 3 NISBET, R. MINER, G. YALE, KP. (2018), Handbook of Statistical Analysis and Data Mining Applications, 2nd Edition, Academic Press.
- 4 AGGARWAL, CC. (2015). Data Mining: the Textbook. Springer.
- 5 CICHOSZ, P. (2015) Data Mining Algorithms: Explained Using R, Wiley.
- 6 JAMES, G. WITTEN, D. HASTIE, T. TIBSHIRANI, R. (2021) An introduction to statistical learning : with applications in R. Springer