Module Specification

Module Summary Information

1	Module Title	Data Mining
2	Module Credits	20
3	Module Level	7
4	Module Code	CMP7206

5	Module Overview

Data mining is the non-trivial process of finding patterns and building models from data stored in data repositories such as databases and data warehouses. At the heart of Big Data Analytics and business intelligence, data mining algorithms provide readily available solutions to many Big Data problems. Data mining is an established field that provides both predictive and descriptive analytics solutions. Such solutions are often generic and can be applied to a wide range of applications from business to scientific and governmental applications.

In this module, you will be taught the internal mechanisms of developing descriptive and predictive data mining methods. Also you will be taught how to use modern data mining tools to build and numerically validate models and patterns extracted from data. You also will be able to critically evaluate current trends in data mining.

Learning and Teaching

This module will introduce data mining techniques through practice based tasks associated with real world problem scenarios. You will gain knowledge of the application of data mining techniques by applying theory to practice and by exploring 'real world' issues via a range of data mining software tools.

The module is designed as a series of pre-reading and discussions, interactive-taught lectures with breakout sessions and workshop/lab exercises using a variety of tools and techniques. You are expected to investigate the topics before sessions. This approach will enable you to gain an insight into how specific data mining techniques are being used in organisations before you learn underpinning theory and practical application through the use of data mining software. You are expected to come to sessions prepared and having completed all the exercises and activities set.

This module is very practical and you will be expected to complete all set exercises in the order which they are presented. This approach will enable you to build your knowledge, skills and practice the embedded transferable employability skills.

6	Indicative Content

Please note that this schedule is indicative and is subject to change for operational and/or educational reasons. Academic staff constantly monitor and review student progress during the teaching period and will make changes to the schedule as appropriate. Any changes will be notified fully to students.

This course discusses basics of data mining and the knowledge discovery process. The course will include the following topics:

Introduction to data mining. Identify new possible applications of data mining with justification
Data and pre-processing. Practise using different pre-processing methods
Association Rule Mining. Practise using association rule mining methods
Kmeans clustering method and hierarchical clustering.
DBSCAN clustering.
Decision tress classification methods.
Naïve Bayes and k-NN classification methods.
Genetic Algorithms.
Guest lecturer will be invited to speak about modern data mining methods.
Tools used in the learning sessions include Oracle R Distribution and Oracle R Enterprise

7	Module Learning Outcomes
	On successful completion of the module, students will be able to:
1	Critically evaluate different data mining techniques for suitability to a given knowledge discovery problem.
2	Apply, compare and numerically validate data mining methods using modern data mining tools.
3	Professionally report data mining results providing clear solutions to identified knowledge discovery problems.
4	Critically review recent trends in data mining literature.

8	Module Assessment		
Learning Outcome			
	Coursework	**Exam**	**In-Person**
1, 2, 3	x		
4			x

9	Breakdown Learning and Teaching Activities	
Learning Activities	**Hours**	
Scheduled Learning (SL) includes lectures, practical classes and workshops, peer group learning, Graduate+, as specified in timetable	48	
Directed Learning (DL) includes placements, work-based learning, external visits, on-line activity, Graduate+, peer learning, as directed on VLE	60	
Private Study (PS) includes preparation for exams	92	
Total Study Hours:	200	