

Module Code	Pre-requisite Module codes	Co-Requisite Modules code(s)	ISCED Code	Subject Code	ECTS Credits	NFQ Level (CPD)#
DATA9900	None	None	489		5	9
Module Title	Data Mining					

School Responsible: School of Computing

Module Overview:

Data mining is the name typically used to describe the full life-cycle of deploying advanced analytical solutions within an organisation. This module will take students through a typical life-cycle (for example CRISP-DM), exploring each of the stages and what tasks and technologies are needed for each. All data mining projects should start with a business objective or question and then use a variety of data discovery techniques and algorithms to find patterns in the data. Building upon these patterns, solutions can be built and deployed in various parts of the organisation. Some of these will be deployed in production. As with most life-cycles the data mining process is iterative. The issues and challenges around this will be explored and how various solutions can be used for this.

Learning Outcomes (LO):

On Completion of this module, the learner will be able to

1	Explain the role of data mining in an organisation
2	Compare the life-cycle stages and what is involved in each both technical and non-technical
3	Devise and apply various data discovery methods
4	Assess, combine and construct data for use in projects
5	Select, apply and evaluate various data mining algorithms for different types of problems
6	Select, evaluate and interpret patterns and knowledge discovered as a result of applying data mining solutions to specified data mining problems.
7	Evaluate various approaches for deploying data mining solutions
8	Assess and contrast various legal, ethical and management issues with projects

Indicative Syllabus:

Module content will be broadly as follows:

- Overview
 - o Introduction to data mining and applications of data mining
 - o Data, Information, Knowledge
 - o Framing a business model
 - o How Data Mining fits within the organisation
- Data Mining Life Cycle
 - o Stages of a DM project
 - o Explore various data mining life-cycles
 - o Evolving nature of roles and responsibilities of people involved in data mining projects
- Data Preparation
 - o Extracting and loading data mining
 - o Data transformations
 - o Data sampling
 - o Data aggregation
 - o Feature engineering
- Exploring Data and Gaining Insights
 - o Using a variety of analytic methods to gain data insights
 - o Role of visualisations in pattern discovery
 - o Time-series forecasting
 - o Exploring and mining text
- Data Mining Techniques. Explore the use of various techniques for structured and unstructured data including:
 - o Classification
 - o Regression
 - o Association rule analysis

Module Code	Pre-requisite Module codes	Co-Requisite Modules code(s)	ISCED Code	Subject Code	ECTS Credits	NFQ Level (CPD)#
DATA9900	None	None	489		5	9
Module Title		Data Mining				

<ul style="list-style-type: none"> ○ Data cluster analysis ○ Anomaly Detection ○ Other data mining techniques within a business context - Understanding and evaluating the outputs and determine what to use - Deploying Data Mining Solutions <ul style="list-style-type: none"> ○ Issues around deployment of data mining solutions ○ Combining multiple algorithms and models ○ Creating pipelines for deployment ○ Various deployment architectures including, API, Docker, Function as a Service, etc ○ Model management and when to retrain models and solutions - Topics on the Management of the Data Mining Process and Life-Cycle <ul style="list-style-type: none"> ○ Legal Issues ○ Ethical Issues ○ Biases in data ○ Using and managing different technologies
--

Learning and Teaching Methods:	
Lectures, tutorials and computer laboratory sessions	
Total Teaching Contact Hours	39
Total Self-Directed Learning Hours	61

Module Delivery Duration:
One Semester.

Assessment		
Assessment Type	Weighting (%)	LO Assessment (No.)
<p>Continuous Assessment will require the student to undertake a data mining problem, to prepare the data for data mining, to select the appropriate techniques, build models, evaluate the models, and recommend which one to use for their problem.</p> <p>Data sets will be provided and may be based on data sets used in various competitions from various data mining related conferences or from Kaggle or other similar events.</p> <p>A report will be submitted containing the details of their data mining work. In this report the students should discuss and justify all decisions made, and be able to explain the outcomes a meaningful way as if they are presenting the results to business user/manager in a work environment.</p> <p>When the assignment is based on previous competitions, the students will be able to evaluate their work with previously published results.</p>	50	3, 4, 5, 6
End of semester examination	50%	1, 2, 3, 6, 7, 8
Module Specific Assessment Arrangements (if applicable)		
(a) Derogations from General Assessment Regulations		
(b) Module Assessment Thresholds		
(c) Special Repeat Assessment Arrangements		

Essential Reading: (author, date, title, publisher)
--

Module Code	Pre-requisite Module codes	Co-Requisite Modules code(s)	ISCED Code	Subject Code	ECTS Credits	NFQ Level (CPD)#
DATA9900	None	None	489		5	9
Module Title	Data Mining					

Kelleher & Tierney, 2017, Essentials of Data Science, MIT Press

Jiawei Han and Micheline Kamber, 2011, Data Mining : Concepts and Techniques, Morgan Kaufmann.

Pang-Ning Tan, Michale Steinbach and Vipin Kumar, 2013, Introduction to Data Mining, Pearson Education.

Michael J. A. Berry, Gordon Linoff, 2011, Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management, Wiley

Supplemental Reading: (author, date, title, publisher)

A range of supplemental reading will be used to support the topics being covered throughout the semester.

These will include various websites, white pages, research papers, reports by industrial analysts, and reports, tutorials, etc from data and database management companies.

Version No:		Amended By	
Commencement Date		Associated Programme Codes	

Modules that are to be offered as Stand-Alone CPD Programmes must have an NFQ level assigned

*Details of the assessment schedule should be contained in the student handbook for the programme stage.

Date of Academic Council approval