

Module Code	ACCA7015
Title	Data Mining and Visualisation
Date	11/04/2019
Level	7
Credits	20
Pre-Requisites	None

AIM(S)

- To develop the student's ability to apply suitable statistical and data mining techniques using a variety of supervised and unsupervised data mining methods.

LEARNING OUTCOMES

Upon the successful completion of this module, the student should be able to:

- Apply and justify the application of a range of statistical methods and data mining techniques to the extraction of information from data.
- Critically evaluate the validity of the techniques employed with respect to the relevant data and to their intended use.
- Interpret the intelligence provided by the data mining process within the context of the data set investigated.
- Effectively communicate the data mining process and its results to a decision maker.

INDICATIVE CONTENT

- Data mining methodologies such as CRISP-DM and SEMMA
- Data preparation, visualisation and exploration
- Basic techniques: statistical modelling, decision trees, rulesets, association rules, linear models, regression, instance-based learning, clustering, multi-instance learning, text mining, sentiment analysis
- Credibility of data mining solutions; training and testing, information loss, cost curves; minimum description length principle
- Machine learning; classification, association, perceptrons, numeric prediction, Bayesian networks, semi-supervised and unsupervised learning
- Data transformations; attribute selection, discretisation, projection, sampling, cleansing
- Applications of data mining
- Software tools for data mining

Study Time Allocation

The study time for the module is divided as follows:

Activity type	Hours	Percentage
Scheduled learning	48	24%
Independent learning	152	76%
Placement learning	0	0%
TOTAL	200	100%

ASSESSMENT

Assessment Component 1

☐ Written exam ☐ Coursework ☒ Practical

Mode of assessment	Volume	Weighting
Presentation	8 minutes	25%

Assessment Component 2

☐ Written exam ☒ Coursework ☐ Practical

Mode of assessment	Volume	Weighting
Report	4,500 words	75%

Assessment Summary

Activity type	Percentage
Written exam	0%
Coursework	75%
Practical	25%
TOTAL	100%

INDICATIVE READING LIST

Module Code	ACCA7015	
Module Title	Data Mining and Visualisation	
Classification	Reading List entry	Electronic*
Essential	North, M. (2018) <i>Data Mining for the Masses, 3rd Ed.: With Implementations in RapidMiner and R</i> . CreateSpace Independent Publishing	☒
Essential	Witten, I. H., Frank, E., Hall, M. A. and Pal, C. J. (2016) <i>Data Mining: Practical Machine Learning Tools and Techniques, 4th Edition</i> , Burlington: Morgan Kaufmann	☒
Further	Tuffery, S. (2011) <i>Data Mining and Statistics for Decision Making</i> , Hoboken: Wiley	☒
Further	Kranzler, J.H. (2017) <i>Statistics for the Terrified</i> , Pearson	☒
Further	Few, S. (2012), <i>Show Me The Numbers</i> , Analytics Press, Burlingame, CA.	☒
Further	Websites: http://www.kdnuggets.com/ http://www.kdd.org/	☒