ROBERT GORDON UNIVERSITY ABERDEEN

MODULE DESCRIPTOR

Module Title

Data Science Development

Reference	CMM535	Version	5
Created	February 2022	SCQF Level	SCQF 11
Approved	April 2015	SCQF Points	15
Amended	July 2022	ECTS Points	7.5

Aims of Module

To provide students with necessary skills for developing complete data science products using state-of-the-art techniques and tools.

Learning Outcomes for Module

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

1	Critically discuss the main concepts, lifecycle and tools for a data science project.
2	Load, explore, clean and pre-process data prior to fitting and evaluating a model using an industry standard data science development tool.
3	Report on data science analyses and results in a clear and reproducible manner.
4	Design, implement, evaluate and deploy a data science product that addresses a given data problem.

Indicative Module Content

1. Data science programming concepts 2. Exploratory Data Analysis and Visualisation 3. Data preparation, data cleaning and data pre-processing techniques 4. Feature Selection (e.g. PCA, one-hot encoding, Cluster analysis) 5. Predictive analytics (applying Classification and regression models) 6. Model evaluation. 7. Model deployment 8. Real world data problems. 9. Professional, legal, ethical, security and social issues relating to data science projects.

Module Delivery

This is a lecture based module, supplemented with practical sessions, where a data science programming language will be used to teach students how to develop a complete data science project from data preparation to advanced analytics. Online materials and online sessions will be used to support DL students.

			Module Ref:	CMM535 v5

Indicative Student Workload	Full Time	Part Time
Contact Hours	40	40
Non-Contact Hours	110	110
Placement/Work-Based Learning Experience [Notional] Hours	N/A	N/A
TOTAL	150	150
Actual Placement hours for professional, statutory or regulatory body		

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: A practical examination covering the stages of the data science lifecycle.

MODULE PERFORMANCE DESCRIPTOR

Explanatory Text

The student has to achieve Grade D in C1 to pass the module.

Module Grade	Minimum Requirements to achieve Module Grade:
A	The student needs to achieve an A in C1.
B	The student needs to achieve B in C1.
C	The student needs to achieve C in C1.
D	The student needs to achieve D in C1.
E	The student needs to achieve E in C1.
F	The student needs to achieve F in C1.
NS	Non-submission of work by published deadline or non-attendance for examination

Module Requirements

Prerequisites for Module	None.
Corequisites for module	None.
Precluded Modules	None.

INDICATIVE BIBLIOGRAPHY

1. JAMES, G., WITTEN, D., HASTIE, T.,& TIBSHIRANI, R. (2021) An introduction to statistical learning with applocations in R. New York: springer.
2. An Introduction to R, Version 4.1.2 (2021), https://cran.r-project.org/doc/manuals/R-intro.pdf
3. ZHAO, Y. (2012-2015). R and Data Mining: Examples and Case Studies, Elsevier. http://www2.rdatamining.com/uploads/5/7/1/3/57136767/rdatamining-book.pdf
4. KORDON, K. (2020) Applying Data Science: How to Create Value with Artificial Intelligence. Springer.
5. MAILUND, T. (2017) Beginning Data Science in R Data Analysis, Visualization, and Modelling for the Data Scientist. APress.