Module Specification

Clustering and Classification

Version: 2023-24, v2.0, 27 Apr 2023

Contents

Part 1: Information

Module title: Clustering and Classification

Module code: UFMFWV-15-3

Level: Level 6

For implementation from: 2023-24

UWE credit rating: 15

ECTS credit rating: 7.5

Faculty: Faculty of Environment & Technology

Department: FET Dept of Computer Sci & Creative Tech

Partner institutions: None

Field: Computer Science and Creative Technologies

Module type: Module

Pre-requisites: None

Excluded combinations: None

Co-requisites: None

Continuing professional development: No

Professional, statutory or regulatory body requirements: None

Part 2: Description

Overview: Clustering and classification techniques are fundamental statistical tools required for data science. Both approaches concentrate on the grouping of similar observations - one (classification) framework exists when the groupings are known, but a mathematical description of the commonalities that form the grouping is required; whereas clustering is a framework in which the true groupings are not known.

These methods often require dimension reduction techniques to be applied in advance, thus the module will also introduce a variety of dimension reduction techniques.

This module will combine a practical focus with a theorectical basis for these methods.

Features: This module will develop students' skills in reproducible research, a process that improves the scientific process, but also can be applied to industry - so that routine analysis can be efficiently reported.

Many "big data" problems can be categorised as either classification or clustering in nature. This module will develop the skills required to do apply, interpret and evaluated such methods.

Educational aims: Using appropriate statistical software and programming languages, students will develop skills required to solve both classification and clustering problems.

Outline syllabus: Dimension Reduction methods:

Principal Components Analysis

Factor Analysis

Fourier Decomposition

Wavelet Analysis

Clustering techniques:

Partitioning methods such as Kmeans,

Agglomorative methods such as Hierarchial clustering (using dendograms)

Model based clustering

Classification techniques:

Linear Discriminant Analysis

Classification Trees

Support Vector Machines

Part 3: Teaching and learning methods

Teaching and learning methods: The practical aspect of this module will be taught in two hour long computer labs so that students have sufficient time to engage with and explore issues within a supportive environment. One hour long seminars will supplement these practical sessions; where students will engage with the theorectical aspects of clustering and classification techniques.

Module Learning outcomes: On successful completion of this module students will achieve the following learning outcomes.

MO1 Analyse and describe structure in high-dimensional data using clustering, classification, and dimension reduction techniques.

MO2 Evaluate statistical analyses that use clustering, classification, and dimension reduction techniques.

MO3 Plan, design, and create reports of statistical results in a fully reproducible fashion using appropriate software.

MO4 Interpret and explain a wide variety of clustering and classification models, in different contexts, to both expert and non-expert audiences.

Hours to be allocated: 150

Contact hours:

Independent study/self-guided study = 114 hours

Face-to-face learning = 36 hours

Total = 150

Reading list: The reading list for this module can be accessed at readinglists.uwe.ac.uk via the following link https://rl.talis.com/3/uwe/lists/C06A83CD-174E-678F-B536-72B6C3CADF75.html

Part 4: Assessment

Assessment strategy: The coursework will be designed so that students must analyse given data, selecting appropriate statistical techniques by applying and evaluating the techniques chosen. It will allow students to demonstrate their skills in producing a piece of reproducible research. The single written report, excluding embedded tables, figures and code, will have a word limit of 2500 words.

The partially seen examination will concentrate on assessing the theorectical underpinnings of the methods taught in the module.

The resit assessment strategy will have the same format as the first sit assessment

Assessment tasks:

Examination (Online) (First Sit)
Description: Online examination
Weighting: 25 %
Final assessment: Yes
Group work: No
Learning outcomes tested: MO1, MO2, MO4

Written Assignment (First Sit)
Description: A practical assessment of the students understanding of how to apply using an appropriate programming language (such as R), report and intrepret techniques taught in the module. (2500 words plus supporting material)
Weighting: 75 %
Final assessment: No

Group work: No
Learning outcomes tested: MO1, MO2, MO3, MO4

Examination (Online) (Resit)
Description: Online Examination
Weighting: 25 %
Final assessment: No
Group work: No
Learning outcomes tested: MO1, MO2, MO4

Written Assignment (Resit)
Description: A practical assessment of the students understanding of how to apply using an appropriate programming language (such as R), report and intrepret techniques taught in the module. (2500 words plus supporting material)
Weighting: 75 %
Final assessment: No
Group work: No
Learning outcomes tested: MO1, MO2, MO3, MO4

Part 5: Contributes towards

This module contributes towards the following programmes of study:

Mathematics [Sep][FT][Frenchay][3yrs] BSc (Hons) 2021-22

Mathematics {Foundation}[Sep][FT][Frenchay][4yrs] BSc (Hons) 2020-21

Mathematics [Sep][SW][Frenchay][4yrs] BSc (Hons) 2020-21

Mathematics [Sep][SW][Frenchay][4yrs] BSc (Hons) 2020-21

Mathematics {Foundation} [Sep][SW][Frenchay][5yrs] BSc (Hons) 2019-20