



## **Module Specification**

### **Data Mining [TSI]**

Version: 2021-22, v1.0, 26 Oct 2021

#### **Contents**

<b>Module Specification .....</b>	<b>1</b>
<b>Part 1: Information .....</b>	<b>2</b>
<b>Part 2: Description .....</b>	<b>2</b>
<b>Part 3: Teaching and learning methods .....</b>	<b>4</b>
<b>Part 4: Assessment.....</b>	<b>5</b>
<b>Part 5: Contributes towards .....</b>	<b>7</b>

## Part 1: Information

**Module title:** Data Mining [TSI]

**Module code:** UFCEA1-12-M

**Level:** Level 7

**For implementation from:** 2021-22

**UWE credit rating:** 12

**ECTS credit rating:** 6

**Faculty:** Faculty of Environment & Technology

**Department:** FET Dept of Computer Sci & Creative Tech

**Partner institutions:** Transport and Telecommunication Institute

**Delivery locations:** Transport and Telecommunication Institute Latvia

**Field:** Computer Science and Creative Technologies

**Module type:** Standard

**Pre-requisites:** None

**Excluded combinations:** None

**Co-requisites:** None

**Continuing professional development:** No

**Professional, statutory or regulatory body requirements:** None

## Part 2: Description

**Overview:** This module will introduce students to various data mining concepts and technologies.

**Features:** Not applicable

**Educational aims:** This module provides students with practical and applied knowledge of how to conduct data mining activities. This includes key concepts in

data mining as well as the statistical and modelling techniques necessary to analyse large data sets to generate meaningful intelligence. The aims of the module are as follows:

Appreciating the value of data mining

Understanding the foundation concepts of data mining

Exploring algorithms commonly used in data mining tools

Ability to apply data mining tools to real-world problems

**Outline syllabus:** This module will introduce and cover the following areas:

Introduction. Basic data mining tasks and techniques review. Data Mining Process (CRISP-DM).

Types of data: interval-scaled, binary, nominal, ordinal, mixed types

Data pre-processing: data cleaning, integration and transformation, dimensionality reduction.

Data Exploration and Visualisation using software

Unsupervised and supervised learning algorithms review. Examples. Measuring Similarity between Objects.

Cluster analysis. Hierarchical algorithms.

Partitional Methods. Distribution and density-based methods. Cluster analysis validity.

Discriminant analysis as predictive technologies. Canonical Analysis. Classification rule classification functions; Mahalanobis distance-based; Posterior probabilities.

Stepwise discriminant analysis.

Computer intensive methods and application for testing classification quality. Main concepts of bootstrap method: the confidence interval estimation.

Essence of dimensionality reduction problem and different methods of its solution.

Principal Components Analysis. Factor analysis.

Multiple regression: model building, transformation of variables, interpretation of results. Spatial and Temporal Data. Time Series. ARIMA. Short-term forecast of transportation volume. ANOVA, MANOVA

Business Intelligence and Data Mining

Advanced Methods and Applications. Text Mining. Social Networks. Big Data.

Web-mining: content, usage, structure

### **Part 3: Teaching and learning methods**

**Teaching and learning methods:** Learning and teaching will be provided to students in two forms: lectures and practical classes (laboratory works in computer classes). During lectures, theoretical aspects of the course will be provided to students by the teaching staff. Lectures will be supported by presentation published and available to the students on e.tsi.lv under the module section. Also, additional materials, like publications on the internet, videos, case-studies etc will be presented in e.tsi.lv.

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

**MO1** To understand data lifecycle, CRISP stages and main approaches of data collection, data quality evaluation

**MO2** Evaluate and apply different data mining algorithms to transform data into useful and actionable information

**MO3** Analyse domain related research process model, identify and analyse available data to identify research questions and/or organisational objectives, to suggest new data required for analytics tasks to deliver maximum insight and formulate sound hypothesis

**MO4** Verify data quality and veracity, to recognise value of data, to apply analytics and statistics methods for data preparation, pre-processing

**MO5** Use effective visualization and storytelling methods to create data analytics reports

**MO6** Apply effectively appropriate data analytics and statistical techniques on available data to discover new relations and deliver insights into research problem or organisational processes and support decision-making

**Hours to be allocated:** 120

**Contact hours:**

Independent study/self-guided study = 112 hours

Face-to-face learning = 48 hours

Total = 160

**Reading list:** The reading list for this module can be accessed at [readinglists.uwe.ac.uk](https://rl.talis.com/3/uwe/lists/48377AAB-25C5-0850-1028-7D4AE165ACBE.html?lang=en-gb&login=1) via the following link <https://rl.talis.com/3/uwe/lists/48377AAB-25C5-0850-1028-7D4AE165ACBE.html?lang=en-gb&login=1>

## **Part 4: Assessment**

**Assessment strategy:** The assessment strategy for this module consists of three parts.

Component A – written examination

Component B

Laboratory work – during this module students will participate with a series of labs (completed individually) where they will complete specific data process tasks.

Component B

Written Report and Presentation – students will be provided with a list of topics for individual student, they are required to produce an essay and presentation on their selected topic.

For referral Component A will be the similar to the main sit. For Component B, the approach will be similar to the main sit, with the volume of the work appropriately scaled down.

**Assessment components:**

**Examination - Component A (First Sit)**

Description: Exam (2 hours)

Weighting: 50 %

Final assessment: Yes

Group work: No

Learning outcomes tested: MO1, MO2, MO3, MO4

**Written Assignment - Component B (First Sit)**

Description: Essay and presentation

Weighting: 25 %

Final assessment: No

Group work: No

Learning outcomes tested: MO1, MO2, MO3, MO4, MO5, MO6

**Laboratory Report - Component B (First Sit)**

Description: A series of practical assignments

Weighting: 25 %

Final assessment: No

Group work: No

Learning outcomes tested: MO2, MO4, MO5, MO6

**Examination - Component A (Resit)**

Description: Exam (2 hours)

Weighting: 50 %

Final assessment: Yes

Group work: No

Learning outcomes tested: MO1, MO2, MO3, MO4

**Written Assignment - Component B (Resit)**

Description: Essay and presentation

Weighting: 25 %

Final assessment: No

Group work: No

Learning outcomes tested: MO1, MO2, MO3, MO4, MO5, MO6

**Laboratory Report - Component B (Resit)**

Description: Resit a single lab which satisfies the learning outcomes

Weighting: 25 %

Final assessment: No

Group work: No

Learning outcomes tested: MO2, MO4, MO5, MO6

**Part 5: Contributes towards**

This module contributes towards the following programmes of study:

Computer Science (Data Analytics and Artificial Intelligence) {Double Degree}

[Feb][FT][TSI][2yrs] MSc 2021-22