1. **Title of the module**

COMP8320 (CO832) – Data Mining and Knowledge Discovery

1. **Division or partner institution which will be responsible for management of the module**

Division of Computing, Engineering, Mathematical Sciences (CEMS)

1. **The level of the module (Level 4, Level 5, Level 6 or Level 7)**

Level 7

1. **The number of credits and the ECTS value which the module represents**

15 credits (7.5 ECTS)

1. **Which term(s) the module is to be taught in (or other teaching pattern)**

Autumn or Spring

1. **Prerequisite and co-requisite modules**

Pre-requisite: knowledge of programming such as that provided by

COMP5200 Further Object-Oriented Programming,

COMP8710 Advanced Java for Programmers,

COMP8820 Advanced Object-Oriented Programming,

COMP3590 Programming for Artificial Intelligence or

COMP8210 Programming for Data Handling

1. **The course(s) of study to which the module contributes**

Portfolio of Taught Postgraduate courses in Computing

Portfolio of BSc courses in Computing

BSc Data Science, both with and without Year in Industry.

1. **The intended subject specific learning outcomes.  
   On successfully completing the module students will be able to:**
   1. Explain the differences between the major data mining tasks, in terms of their assumptions, requirement for a specific kind of data, and the different kinds of knowledge discovered by algorithms performing different kinds of task.
   2. Describe data mining algorithms for the major data mining tasks.
   3. Identify which data mining task and which algorithm is the most appropriate for a given data mining project, taking into account both the nature of the data to be mined and the goals of the user of the discovered knowledge.
   4. Use a state-of-the-art data mining tool in a principled fashion, being aware of the strengths and weaknesses of the algorithms implemented in the tool.
   5. Evaluate the quality of discovered knowledge, taking into account the requirements of the data mining task being solved and the goals of the user.
   6. Describe the main tasks and algorithms involved in the preprocessing and postprocessing steps of the knowledge discovery process.
   7. Utilize the library and exploit web sites to support investigations into these areas.
2. **The intended generic learning outcomes.  
   On successfully completing the module students will be able to:**
   1. Understand the major kinds of data mining tasks and the main kinds of algorithms that are often used to solve these tasks.
   2. Understand the strengths and weaknesses of some data mining algorithms, identifying the kind of algorithm that is most appropriate for each data mining problem.
   3. Understand the process of knowledge discovery, involving not only data mining but also preprocessing and post-processing steps.
3. **A synopsis of the curriculum**

This module explores a range of different data mining and knowledge discovery techniques and algorithms. You learn about the strengths and weaknesses of different techniques and how to choose the most appropriate for any particular task. You use a data mining tool, and learn to evaluate the quality of discovered knowledge.

1. **Reading list (Indicative list, current at time of publication. Reading lists will be published annually)**

Witten, IH, Frank, E, Hall, MA, Pal, CJ (2016). *Data Mining: practical machine learning tools and techniques*, 4rd edition. Morgan Kaufmann.

Tan, P-N, Steinbach, M, Karpatne, A, Kumar, V (2018) Introduction to Data Mining, Pearson, 2nd edition.

1. **Learning and teaching methods**

Total contact hours: 22 hours

Private study hours: 128 hours

Total study hours: 150 hours

1. **Assessment methods**
   1. Main assessment methods

One exercise with a data mining tool 10%

One Short Essay (about 1,000 words) 10%

Examination 80%

13.2 Reassessment methods

Like for like.

1. ***Map of module learning outcomes (sections 8 & 9) to learning and teaching methods (section12) and methods of assessment (section 13)***

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Module learning outcome** | 8.1 | 8.2 | 8.3 | 8.4 | 8.5 | 8.6 | 8.7 | 9.1 | 9.2 | 9.3 |
| **Learning/ teaching method** |  |  |  |  |  |  |  |  |  |  |
| Lectures | x | x | x | x | x | x | x | x | x | x |
| Private study | x | x | x | x | x | x | x | x | x | x |
| **Assessment method** |  |  |  |  |  |  |  |  |  |  |
| Exercise with a data mining tool |  |  |  | x | x |  |  | x | x |  |
| Short Essay |  |  |  |  |  |  | x | x | x | x |
| Examination | x | x | x |  | x | x |  | x | x | x |

1. **Inclusive module design**

The Division recognises and has embedded the expectations of current equality legislation, by ensuring that the module is as accessible as possible by design. Additional alternative arrangements for students with Inclusive Learning Plans (ILPs)/declared disabilities will be made on an individual basis, in consultation with the relevant policies and support services.

The inclusive practices in the guidance (see Annex B Appendix A) have been considered in order to support all students in the following areas:

a) Accessible resources and curriculum

b) Learning, teaching and assessment methods

1. **Campus(es) or centre(s) where module will be delivered**

Canterbury

1. **Internationalisation**

The topics addressed by this module relate to a field which is of international importance, given the global role of computers in today’s technological innovation. The topics covered by this module are international in nature, being identical worldwide and independent of traditional spoken language.

**DIVISIONAL USE ONLY**

**Revision record – all revisions must be recorded in the grid and full details of the change retained in the appropriate committee records.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Date approved | Major/minor revision | Start date of delivery of revised version | Section revised | Impacts PLOs (Q6&7 cover sheet) |
| 10/11/2020 | Minor |  | 6, 7, 16 | No |