1. **Title of the module**

MAST6060 – Statistical Machine Learning

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 6

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**

Prerequisite: MAST5001 (Applied Statistical Modelling 1), MAST4009 (Probability), MAST4011 (Statistics), MAST4006 (Mathematical Methods)

Co-requisite: None

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

BSc Data Science (including course with a Year in Industry), BSc Data Science with a Foundation Year

1. **The intended subject specific learning outcomes.  
   On successfully completing the module students will be able to:**
   1. demonstrate systematic understanding of key aspects of machine learning;
   2. demonstrate the capability to deploy established approaches accurately to analyse and solve problems using a reasonable level of skill in calculation and manipulation of the material in the following areas: Supervised learning and prediction, regression tree-based methods and support vector machines;
   3. apply key aspects of machine learning in well-defined contexts, showing judgement in the selection and application of tools and techniques;
   4. show judgement in the selection and application of machine learning approaches.
2. **The intended generic learning outcomes.  
   On successfully completing the module students will be able to:**
   1. manage their own learning and make use of appropriate resources;
   2. understand logical arguments, identifying the assumptions made and the conclusions drawn;
   3. communicate straightforward arguments and conclusions reasonably accurately and clearly;
   4. manage their time and use their organisational skills to plan and implement efficient and effective modes of working;
   5. solve problems relating to qualitative and quantitative information;
   6. make use of R, online resources (Moodle), internet communication;
   7. communicate technical and non-technical material competently;
   8. demonstrate an increased level of skill in numeracy and computation;
   9. demonstrate the acquisition of the study skills needed for continuing professional development.
3. **A synopsis of the curriculum**

Statistical machine learning deals with the problem of finding a predictive function based on data, and focuses on the intersection of statistics and machine learning. It involves the development of algorithms that learn from observed data by constructing stochastic models, which can be used for making predictions and decisions. In this module, students study statistical machine learning methods and how they are implemented in R. Both theoretical and practical aspects are covered.

Indicative content: Classification and prediction; K-Nearest Neighbours; cross-validation and bootstrap; classification and regression trees; bagging; random forests; boosting; support vector classifiers; support vector machine (SVM); regression SVM; semisupervised learning.

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

Bishop, C. M. (2006), Pattern Recognition and Machine Learning. Springer, New York

James, G, Witten, D., Hastie, T., Tibshirani, R. (2013) Introduction to Statistical Learning. Springer, New York.

Brett, L. (2019) Machine Learning with R: Expert techniques for predictive modeling, 3rd Edition. Packt Publishing, Birmingham.

Berry, M., and Linoff, G. (2012). Data Mining Techniques: For Marketing, Sales and Customer Relationship Management.

1. **Learning and teaching methods**

Total contact hours: 44

Private study hours: 106

Total study hours: 150

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

Assessment 1 Exercises, requiring on average between 10 and 15 hours to complete 20%

Assessment 2 Exercises, requiring on average between 10 and 15 hours to complete 20%

Examination 2 hours 60%

The coursework mark alone will not be sufficient to demonstrate the student’s level of achievement on the module, and the coursework must be passed overall in order to pass the module.

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 | 9.1 | 9.2 | 9.3 | 9.4 | 9.5 | 9.6 | 9.7 | 9.8 | 9.9 |
| **Learning/ teaching method** |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Private Study | **X** | **X** | **X** | **X** | **X** | **X** | **X** | **X** | **X** | **X** | **X** | **X** | **X** |
| Lectures/Exercise Classes | **X** | **X** | **X** |  |  | **X** | **X** |  | **X** |  | **X** | **X** |  |
| Computer sessions | **X** | **X** | **X** | **X** | **X** | **X** | **X** |  | **X** | **X** | **X** | **X** |  |
| **Assessment method** |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Examination | **X** | **X** | **X** |  | **X** | **X** | **X** | **X** | **X** |  | **X** | **X** | **X** |
| Coursework | **X** | **X** | **X** | **X** | **X** | **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**

Data science, as a discipline, uses internationally recognised techniques developed and refined by statisticians and computer scientists across the globe. Mastery of the subject-specific learning outcomes, 8.1 to 8.4, will equip students to apply the theories and techniques of this module in a wide range of international contexts. The module team is drawn from the School of Mathematics, Statistics and Actuarial Science/School of Computing, which includes many members of staff with international experience of teaching and research collaboration.

In compiling the reading list, consideration has been given to the range of texts that are available internationally and a selection of texts has been identified to complement the delivery of the material.

Examples with an international dimension are included in the module where appropriate.

The support SMSAS/Computing provides to its students is also internationally attuned given our international student body.

**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) |
| July 2023 | Minor | September 2023 | 13 |  |
|  |  |  |  |  |