



| Teaching Guide      |   |        |  |         |
|---------------------|---|--------|--|---------|
| Identifying Data    |   |        | 2021/22  |         |
| Subject (*)         | Machine Learning I  | Code   | 614G02019  |         |
| Study programme     | Grao en Ciencia e Enxeñaría de Datos  |        |  |         |
| Descriptors         |   |        |  |         |
| Cycle               | Period  | Year   | Type   | Credits |
| Graduate            | 2nd four-month period   | Second | Obligatory   | 6       |
| Language            | Spanish   |        |  |         |
| Teaching method     | Face-to-face  |        |  |         |
| Prerequisites       |   |        |  |         |
| Department          | Ciencias da Computación e Tecnoloxías da Información  |        |  |         |
| Coordinador         | Rabuñal Dopico, Juan Ramon  | E-mail | juan.rabunal@udc.es  |         |
| Lecturers           | Álvarez González, Sara<br>Molares Ulloa, Andrés<br>Rabuñal Dopico, Juan Ramon<br>Rivero Cebrián, Daniel<br>Rodríguez Tajés, Álvaro  | E-mail | sara.alvarezg@udc.es<br>andres.molares@udc.es<br>juan.rabunal@udc.es<br>daniel.rivero@udc.es<br>a.tajes@udc.es |         |
| Web                 |   |        |  |         |
| General description | This course presents an overview of automatic learning. The syllabus explains the different techniques and methods, including supervised and unsupervised learning. In the practical part, real cases will be solved. |        |  |         |



|                         |   |
|-------------------------|---|
| <b>Contingency plan</b> | <p>1. Modifications to the contents</p> <p>No changes are made.</p> <p>2. Methodologies</p> <p>Methodologies are maintained.</p> <p>The written assessment test is changed to a non-attendance assessment test. This final test is necessary to make an individual evaluation of each student, since the practices and works are made in groups.</p> <p>3. Mechanisms for personalized attention to students</p> <p>Use of Moodle to provide material to students.</p> <p>Use of the Moodle forum to communicate all those events of the subject (modifications, deliveries of practices, etc.)</p> <p>Synchronous teaching in class time and asynchronous teaching through Teams.</p> <p>Tutoring through the Teams chat on a continuous basis.</p> <p>Tutorials through email on a continuous basis.</p> <p>4. Modifications in the evaluation</p> <p>The evaluation mechanisms are maintained, with the above-mentioned change of the written test to a non-attendance test.</p> <p>*Evaluation observations:</p> <p>The need to obtain a minimum score in the theory exam is eliminated. The rest of the evaluation observations are maintained.</p> <p>5. Modifications to the bibliography or webgraphy</p> <p>No changes are made.</p> |
|-------------------------|---|

| Study programme competences / results |  |
|---------------------------------------|--|
| Code                                  | Study programme competences / results  |
| A24                                   | CE24 - Comprensión e dominio das principais técnicas básicas e avanzadas de aprendizaxe automática, incluíndo as dedicadas ao tratamento de grandes volumes de datos.  |
| A25                                   | CE25 - Capacidade para identificar a adecuación de cada unha das técnicas de aprendizaxe automática á resolución dun problema, incluíndo os aspectos relacionados coa súa complexidade computacional ou a súa capacidade explicativa, de acordo aos requisitos establecidos. |
| A26                                   | CE26 - Coñecemento das ferramentas informáticas actuais no campo da aprendizaxe automática, e capacidade para seleccionar a máis adecuada para a resolución dun problema.  |
| B2                                    | CB2 - Que os estudantes saiban aplicar os seus coñecementos ao seu traballo ou vocación dunha forma profesional e posúan as competencias que adoitan demostrarse por medio da elaboración e defensa de argumentos e a resolución de problemas dentro da súa área de estudo   |
| B3                                    | CB3 - Que os estudantes teñan a capacidade de reunir e interpretar datos relevantes (normalmente dentro da súa área de estudo) para emitir xuízos que inclúan unha reflexión sobre temas relevantes de índole social, científica ou ética                                    |
| B7                                    | CG2 - Elaborar adecuadamente e con certa orixinalidade composicións escritas ou argumentos motivados, redactar plans, proxectos de traballo, artigos científicos e formular hipóteses razoables.   |
| B8                                    | CG3 - Ser capaz de manter e estender formulacións teóricas fundadas para permitir a introdución e explotación de tecnoloxías novas e avanzadas no campo.   |



|     |   |
|-----|---|
| B9  | CG4 - Capacidade para abordar con éxito todas as etapas dun proxecto de datos: exploración previa dos datos, preprocesado, análise, visualización e comunicación de resultados.       |
| B10 | CG5 - Ser capaz de traballar en equipo, especialmente de carácter multidisciplinar, e ser hábiles na xestión do tempo, persoas e toma de decisións.                                   |
| C1  | CT1 - Utilizar as ferramentas básicas das tecnoloxías da información e as comunicacións (TIC) necesarias para o exercicio da súa profesión e para a aprendizaxe ao longo da súa vida. |

| Learning outcomes   |                                       |                       |    |
|---|---------------------------------------|-----------------------|----|
| Learning outcomes   | Study programme competences / results |                       |    |
| Understand the relationship between the complexity of learning models, training data features and overfitting, and know the mechanisms to avoid it.   | A24<br>A25                            |                       |    |
| Develop skills to design the stages of a complete data analysis process based on automatic learning techniques.   |                                       | B2<br>B7<br>B9<br>B10 | C1 |
| Know how to correctly apply automatic learning techniques to obtain reliable and significant results.   | A24                                   | B3<br>B8              |    |
| Know the most representative and current techniques of unsupervised, semi-supervised and supervised learning, with and without reinforcement.   | A24                                   | B8                    |    |
| Know the most representative learning techniques for the classic problems of classification, regression and clustering, and other less classic ones such as sorting problems, one class problems or multitasking. | A24                                   | B8                    |    |
| Identify appropriate data analysis techniques according to the problem.   | A25                                   | B3<br>B8              |    |
| Manage the most current tools and work environments in the field of machine learning.   | A26                                   | B2<br>B10             |    |

| Contents                          |   |
|-----------------------------------|---|
| Topic                             | Sub-topic   |
| 1. Introduction                   | 1.1. Introduction to Machine Learning<br>1.2. Inductive Learning  |
| 2. Supervised learning            | 2.1. Introduction<br>2.2. Artificial Neural Networks<br>2.3. Support Vector Machines<br>2.4. Decision trees<br>2.5. Regression trees and regression model trees<br>2.6. Instance-based learning |
| 3. Evolutionary Computation       | 3.1. Genetic Algorithms<br>3.2. Genetic Programming<br>3.3. Swarms and other Evolutionary Computation techniques  |
| 4. Methodologies in data analysis | 4.1. Training, evaluation and model selection methodologies<br>4.2. Methodologies of a data analysis project  |
| 5. Unsupervised learning          | 5.1. Clustering methods<br>5.2. Self-organised networks   |

| Planning                       |                        |                                      |                               |             |
|--------------------------------|------------------------|--------------------------------------|-------------------------------|-------------|
| Methodologies / tests          | Competencies / Results | Teaching hours (in-person & virtual) | Student?s personal work hours | Total hours |
| Guest lecture / keynote speech | A24 A25 B3 B8 B9       | 30                                   | 60                            | 90          |



|                        |                  |    |    |    |
|------------------------|------------------|----|----|----|
| Laboratory practice    | A26 B2 B3 B10 C1 | 18 | 18 | 36 |
| Supervised projects    | B2 B3 B7 B9 B10  | 10 | 10 | 20 |
| Objective test         | A24 A25 B8 B9    | 2  | 0  | 2  |
| Personalized attention |                  | 2  | 0  | 2  |

(\*)The information in the planning table is for guidance only and does not take into account the heterogeneity of the students.

| Methodologies                  |   |
|--------------------------------|---|
| Methodologies                  | Description   |
| Guest lecture / keynote speech | Theoretical teaching of the subject matter of the course  |
| Laboratory practice            | Solve practical problems by using the different techniques that will be explained in the theory classes.  |
| Supervised projects            | Writing, under the supervision of the teacher, of the reports explaining the resolution of the problems carried out in the laboratory practices and the results obtained. |
| Objective test                 | This is a written assessment test in which the student must demonstrate the knowledge acquired from the subject.  |

| Personalized attention |  |
|------------------------|--|
| Methodologies          | Description  |
| Laboratory practice    | Practical work carried out with the advice of the teacher.         |
| Supervised projects    | Writing of the explanatory report under the teacher's supervision. |

| Assessment          |                        |   |               |
|---------------------|------------------------|---|---------------|
| Methodologies       | Competencies / Results | Description   | Qualification |
| Laboratory practice | A26 B2 B3 B10 C1       | Resolution of real world problems using the methodology, for which several techniques explained in theory will be used, and the student will be stimulated to generate new ideas for the resolution of these problems.  | 25            |
| Supervised projects | B2 B3 B7 B9 B10        | Writing of the report on the resolution of the real problems carried out in the laboratory practices. The writing of the report will include a bibliographic review of the most important works related, written in English for the most part, documentation on the problem to be solved, methodology used, and comparison of the results found in the application of the different techniques, as well as a critical evaluation of both the results obtained and the information used. | 25            |
| Objective test      | A24 A25 B8 B9          | Test questions about the contents of the course, based on the different machine learning techniques and their applications.   | 50            |

| Assessment comments  |
|--|
| <p>In order to pass the subject, the student must obtain a minimum score of 5 out of 10 in the result of combining the grades of the objective test, the laboratory practices and the supervised works. In addition, the student must obtain a minimum score of 2 out of 5 points in the objective test. If the student does not obtain this minimum grade, the grade of the subject will be that corresponding to the grade of the objective test.</p> <p>In the second opportunity, the grade obtained in the laboratory practices and supervised works will be maintained, not being able to obtain again a grade since it results from the continuous evaluation of the work during the credits of practice of the subject. The student can retake the examination of the objective test, the criteria for obtaining the total score being those indicated at the beginning of this section.</p> <p>Part-time students must deliver in their reports on the same dates as full-time students, and attend the RGTs in which they will be corrected. Similarly, it is recommended that they attend the practice classes.</p> |

| Sources of information |
|------------------------|
|------------------------|



|                      |   |
|----------------------|---|
| <b>Basic</b>         | <ul style="list-style-type: none"><li>- D. Borrajo, J. González, P. Isasi (2006). Aprendizaje automático. Sanz y Torres</li><li>- T.M. Mitchell (1997). Machine Learning. McGraw Hill</li><li>- Basilio Sierra Araujo (2006). Aprendizaje automático: conceptos básicos y avanzados. Aspectos prácticos utilizando el software WEKA. Pearson Education</li><li>- Saso Dzeroski, Nada Lavrac (). Relational Data Mining. Springer</li><li>- David Aha (). Lazy Learning. Kluwer Academics Publishers</li><li>- Andrew Webb (2002). Statistical Pattern Recognition. Wiley</li><li>- Ethem Alpaydin (2004). Introduction to Machine Learning. MIT Press</li><li>- Marcos Gestal, Daniel Rivero, Juan Ramón Rabuñal, Julián Dorado, Alejandro Pazos (2010). Introducción a los Algoritmos Genéticos y a la Programación Genética. Servicio de Publicaciones de la Universidade da Coruña</li></ul> <p>&lt;br /&gt;</p> |
| <b>Complementary</b> |   |

## Recommendations

### Subjects that it is recommended to have taken before

Design and Analysis of Algorithms/614G02011  
Regression Models/614G02012  
Statistical Modeling of High Dimensional Data/614G02013  
Signals and Systems/614G02014  
Fundamentals of Programming II/614G02009  
Fundamentals of Programming I/614G02004  
Statistical Inference/614G02007

### Subjects that are recommended to be taken simultaneously

Information Theory/614G02018  
Mathematical Optimisation/614G02020

### Subjects that continue the syllabus

Large Scale Machine Learning/614G02032  
Numerical Methods for Data Science/614G02033  
Machine Learning III/614G02026  
Image, Video and Audio Processing/614G02028  
Machine Learning II/614G02021

### Other comments

(\*)The teaching guide is the document in which the URV publishes the information about all its courses. It is a public document and cannot be modified. Only in exceptional cases can it be revised by the competent agent or duly revised so that it is in line with current legislation.