Course guide
2500036 - GECAPAUTC - Machine Learning and Data Science

Unit in charge: Barcelona School of Civil Engineering
Teaching unit: 751 - DECA - Department of Civil and Environmental Engineering.
Degree: BACHELOR’S DEGREE IN CIVIL ENGINEERING (Syllabus 2020). (Optional subject).
Academic year: 2022 ECTS Credits: 4.5 Languages: English

LECTURER

Coordinating lecturer: IRENE ARIAS VICENTE
Others: IRENE ARIAS VICENTE, PEDRO DIEZ MEJIA, ALEJANDRO TORRES SÁNCHEZ

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
14406. Ability to analyze the problem of safety and health in construction sites. (Common module to the Civil branch)
14410. Knowledge of the typology and calculation bases of prefabricated elements and their application in manufacturing processes. (Specific technology module: Civil Construction)
14411. Knowledge about the project, calculation, construction and maintenance of building works in terms of structure, finishes, facilities and own equipment. (Specific technology module: Civil Construction)
14413. Capacity for the construction and conservation of roads, as well as for the dimensioning, the project and the elements that make up the basic road equipment. (Specific technology module: Civil Construction)
14414. Capacity for the construction and conservation of railway lines with knowledge to apply specific technical regulations and differentiating the characteristics of the mobile material. (Specific technology module: Civil Construction)
14415. Ability to apply construction procedures, construction machinery and construction planning techniques. (Specific technology module: Civil Construction)
14416. Capacity for the construction of geotechnical works. (Specific technology module: Civil Construction)

TEACHING METHODOLOGY

The course consists of 1.5 hours per week of classroom activity (large size group) and 1.5 hours weekly with half the students (medium size group).

The 1.5 hours in the large size groups are devoted to theoretical lectures, in which the teacher presents the basic concepts and topics of the subject, shows examples and solves exercises.

The 1.5 hours in the medium size groups is devoted to solving practical problems with greater interaction with the students. The objective of these practical exercises is to consolidate the general and specific learning objectives.

Support material in the form of a detailed teaching plan is provided using the virtual campus ATENEA: content, program of learning and assessment activities conducted and literature.

Although most of the sessions will be given in the language indicated, sessions supported by other occasional guest experts may be held in other languages.
LEARNING OBJECTIVES OF THE SUBJECT

Knowledge about machine learning algorithms and data science.

1. Understand and apply the main machine learning algorithms.
2. Understand the life cycle phases of data science: data mining processes.

Supervised (regression and classification), unsupervised (clustering) and semi-supervised learning. Linear regression methods (regression error functions, least squares, notion of regularization, generalized regression). Linear methods by classification (error functions by classification, Bayesian classifiers). Hierarchical methods, general construction of decision trees. Neural networks. Kernel-based methods. Kernelized regularized linear regression, basic kernel functions. Explore the life cycle of data science: questioning, data collection, analysis, visualization, statistical inference, prediction, and decision making. It focuses on quantitative critical thinking and key principles and techniques: languages to transform, query, and analyze data; algorithms for machine learning methods: regression, classification and grouping; principles of informational visualization; measurement and prediction error; and techniques for scalable data processing.

STUDY LOAD

<table>
<thead>
<tr>
<th>Type</th>
<th>Hours</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hours medium group</td>
<td>22,5</td>
<td>20.00</td>
</tr>
<tr>
<td>Hours large group</td>
<td>22,5</td>
<td>20.00</td>
</tr>
<tr>
<td>Guided activities</td>
<td>4,5</td>
<td>4.00</td>
</tr>
<tr>
<td>Self study</td>
<td>63,0</td>
<td>56.00</td>
</tr>
</tbody>
</table>

Total learning time: 112.5 h

CONTENTS

Introduction to machine learning and theory of decision making

Description:

Full-or-part-time: 9h 36m  
Theory classes: 2h  
Practical classes: 2h  
Self study : 5h 36m

Unsupervised learning

Description:
Principal component analysis  
Principal Component Analysis  
Principal component analysis

Full-or-part-time: 24h  
Theory classes: 4h  
Practical classes: 4h  
Laboratory classes: 2h  
Self study : 14h
### Bayesian Inference

**Description:**
Bayesian model update. Prior and posterior.

**Full-or-part-time:** 9h 36m  
Theory classes: 2h  
Laboratory classes: 2h  
Self study : 5h 36m

### Supervised learning

**Description:**
Least squares, error functions for regression, probabilistic approach, sum of squares error as maximum likelihood, model selection, the curse of dimensionality, generalized regression.

Linear models for regression  
Discriminant functions, connection to maximum likelihood. Model selection. Bayesian logistic regression.  
Linear classification models

**Full-or-part-time:** 19h 12m  
Theory classes: 4h  
Laboratory classes: 4h  
Self study : 11h 12m

### Artificial Neural Networks

**Description:**
Basic concepts of ANN  
The multilayer perceptron  
Network training  
Regularization in ANN

**Full-or-part-time:** 28h 47m  
Theory classes: 4h  
Laboratory classes: 8h  
Self study : 16h 47m

### Simulation

**Description:**
Monte-Carlo Method and Stochastic Finite elements  
Assignment of Stochastic finite elements

**Full-or-part-time:** 7h 11m  
Theory classes: 1h  
Laboratory classes: 2h  
Self study : 4h 11m
Course Project Presentations

Full-or-part-time: 9h 36m
Laboratory classes: 4h
Self study : 5h 36m

GRADING SYSTEM

The student will be assessed using two types of tests. On the one hand, during the classes there will be several evaluable tests, which the student must deliver at the end of the session or shortly after; these tests will weigh 30% in the final grade and may include an oral presentation. On the other hand, towards the end of the course the student will have to carry out an evaluation exam. This exam will weigh 70% of the final grade. Criteria of qualification and of admission to the re-evaluation: The students suspended to the ordinary evaluation that have presented regularly in the proofs of evaluation of the asignatura suspended will have option to realize a proof of re-evaluation in the period fixed in the academic calendar. Students who have already passed it or students who have qualified as not presented will not be able to take the re-assessment test for a subject. The maximum grade in the case of taking the re-assessment exam will be five (5.0). The non-attendance of a student summoned to the re-evaluation test, held in the set period may not lead to the performance of another test with a later date. Extraordinary assessments will be carried out for those students who, due to accredited force majeure, have not been able to take any of the continuous assessment tests. These tests must be authorized by the corresponding head of studies, at the request of the teacher responsible for the subject, and will be carried out within the corresponding teaching period.

EXAMINATION RULES.

Failure to perform a laboratory or continuous assessment activity in the scheduled period will result in a mark of zero in that activity.

BIBLIOGRAPHY

Basic: