

Degree programme	Type	Course
Applied Statistics	OB	3

Contact lecturer

Name : Rosario Delgado De la Torre

Email : rosario.delgado@uab.cat

Teaching staff

David Moríña Soler

Anabel Blasco Moreno

Group languages

You can consult this information at the [end](#) of the document.

Prerequisites

This course assumes that the student has acquired the knowledge taught in various subjects on the following topics:

- Linear Algebra and Calculus
- Probability and Statistical Inference
- Computer Tools for Statistics and Introduction to Programming
- Linear Models

Objectives

This course aims to introduce students to the field of Supervised Machine Learning by presenting various methodologies and basic concepts.

Learning outcomes

- CM11 (Create new machine learning models, running experiments to demonstrate their feasibility and improved performance compared to the state of the art.) Create new machine learning models, running experiments to demonstrate their feasibility and improved performance compared to the state of the art.
- CM12 (Assess the existence of inequalities on the grounds of gender in databases, to avoid bias in automatic (algorithmic) decision-making.) Assess the existence of inequalities on the grounds of gender

- in databases, to avoid bias in automatic (algorithmic) decision-making.
- KM16 (Recognise supervised and unsupervised, profound and generic machine learning models, fostering innovation in the field of statistics.) Recognise supervised and unsupervised, profound and generic machine learning models, fostering innovation in the field of statistics.

Contents

- Introduction to Supervised Machine Learning.
- Reinforcement Learning
- Support Vector Machines
- K-Nearest Neighbors
- Decision Trees and Random Forests
- Validation, confusion matrices, and performance metrics (binary case)
- Bias and fairness in Machine Learning

Learning activities and methodology

Title	Hours	ECTS	Learning outcomes
Theory sessions	50	2	
Lab sessions	30	1.2	
Personal study of the subject	46	1.84	

Teaching will combine face-to-face lectures with hands-on problem-solving and computer-based practical activities.

In all aspects of teaching/learning activities, the best efforts will be made by teachers and students to avoid language and situations that can be interpreted as sexist. To achieve continuous improvement in this subject, everyone should collaborate in highlighting them.

Annotation: within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

Assessment

Continuous assessment activities

Title	Weight	Hours	ECTS	Learning outcomes
PAC2. Deliveries of problems and exercises with computer (classroom)	10%	12	0.48	CM11, CM12, KM16
Computed-based practical exam (classroom)	15%	3	0.12	CM11, CM12, KM16
Exam	50%	3	0.12	CM12, KM16
Resit exam	75%	3	0.12	CM12, KM16
PAC1 (problem solving test)	25%	3	0.12	KM16

Continuous Assessment

The assessment of the course consists of four components: the grade obtained in PAC1 (a mid-semester problem-solving test), the grades obtained from assignments submitted during problem-solving sessions and/or computer-based practical activities (PAC2), the grade obtained in the computer-based practical exam held during the last practical session (Eprac), and the grade obtained in the final problem-solving examination (Ex).

The course grade will be calculated as

$$N = 0.25 \cdot \text{PAC1} + 0.10 \cdot \text{PAC2} + 0.15 \cdot \text{Eprac} + 0.50 \cdot \text{Ex},$$

provided that each component is at least 3.5 out of 10. Otherwise, any component with a grade below 3.5 will be counted as 0 in the calculation of N.

If $N \geq 5$, the course will be considered passed and the final grade (FG) will be $\text{FG} = N$.

Otherwise, students may take a resit examination (ExRec). In this case, the final grade will be calculated as

$$\text{FG} = 0.75 \cdot \text{ExRec} + 0.10 \cdot \text{PAC2} + 0.15 \cdot \text{Eprac}.$$

Therefore, neither the PAC2 assignments nor the practical examination are recoverable.

Under no circumstances may the resit examination be used to improve the grade of a course that has already been passed.

A student will be considered assessable if they have participated in at least one assessment activity. Otherwise, the final record will state "Not Assessed".

Single Assessment

For students who opt for the single-assessment modality, the final grade will be based on the final examination (75%) and a computer-based practical examination (25%).

USE OF ARTIFICIAL INTELLIGENCE

For this course, the use of Artificial Intelligence (AI) technologies is permitted exclusively for support tasks, such as literature and information searches, text revision, or translation.

Students must clearly identify, where applicable, which parts of their submitted work have been generated with the assistance of AI technologies, specify the tools used, and include a critical reflection on how these tools have influenced both the process and the final outcome of the activity.

Failure to disclose the use of AI in an assessment activity will be considered a breach of academic integrity and may result in a partial or total reduction of the grade for that activity, or in more severe disciplinary measures in serious cases.

Any irregularity committed in an assessment activity (academic fraud, plagiarism, or improper use of Artificial Intelligence, unless such use is expressly authorized in the course syllabus) that may lead to a significant alteration of the grade will result in that assessment activity being graded with a 0 (fail).

If the course syllabus establishes that obtaining a minimum grade in that assessment activity is an essential requirement to pass the course, or if multiple irregularities are detected in the assessment activities of the same course, the final grade for the course will be 0 (fail).

In addition, disciplinary proceedings may be initiated against any student who commits any of these irregularities.

Bibliography

Basic bibliography

James, G., Witten, D., Hastie, T., & Tibshirani, R. *An Introduction to Statistical Learning: with Applications in R* (and *An Introduction to Statistical Learning: with Applications in Python*). Copies are available through the university library, both in print and online.

Additional bibliography

- Géron, A. (2019, 2nd ed.). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*.
- Hastie, T., Tibshirani, R., & Friedman, J. (2016, 2nd ed.). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Series in Statistics.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Information Science and Statistics Series. Springer.
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill Series in Computer Science: Artificial Intelligence.
- Sutton, R. S., & Barto, A. G. (2018, 2nd ed.). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.
- Barocas, S., Hardt, M., & Narayanan, A. (2019). *Fairness and Machine Learning*. Available at: fairmlbook.org

Software

Python and RStudio will be used, the latter being an IDE (Integrated Development Environment) specifically designed for the R programming language.

Course groups and languages

The information provided is provisional until November 30. After this date, you will be able to consult the language of each group through this [link](#). To access the information, you will need to enter the course CODE

Type of teaching	Group	Language	Semester	Shift
(TE) Theory	1	Catalan	first semester	afternoon
(PLAB) Practical laboratories	1	Catalan	first semester	afternoon
(PLAB) Practical laboratories	2	Catalan	first semester	afternoon