

Degree programme	Type	Course
Computational Mathematics and Data Analytics	OB	3

Contact lecturer

Name : Antonio Lozano Bagen

Email : antonio.lozano.bagen@uab.cat

Prerequisites

A good knowledge of the contents of the subjects studied during the first year such as Probability and Calculus and of the second year such as Artificial Intelligence and Modeling and Inference is considered very important. Although some theoretical part of the contents has already been treated in those subjects from a mathematical point of view, in this subject we will focus on their implementation in Python and their application in multiple real cases.

Group languages

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

Objectives

The main objective of this course is to provide students with a solid understanding of the fundamental principles of machine learning, with an emphasis on the computational and algorithmic aspects that enable the construction of models from datasets.

Knowledge:

- Describe the basic techniques of computer learning.
- List the essential steps of different machine learning algorithms.
- Identify the advantages and disadvantages of the learning algorithms.
- Solve problems by applying different machine learning techniques to find the optimal solution.
- Understand the results and limitations of each learning technique in different case studies.
- Know how to choose the most appropriate learning algorithm to solve contextualized problems.

Skills:

- Recognize situations in which the application of machine learning algorithms may be adequate.
- Analyze the problem to solve and design the optimal solution applying the learned techniques.
- Write technical documents related to the analysis and solution of a problem.

- Program the basic algorithms to solve the proposed problems.
- Evaluate the results of the implemented solution and propose possible improvements.
- Defend and argue the decisions taken in the solution of proposed problems.

Learning outcomes

- CM22 (Apply the most suitable learning techniques to solve computational problems in different case studies.) Apply the most suitable learning techniques to solve computational problems in different case studies.
- CM23 (Assess the results and limitations of the most common learning techniques.) Assess the results and limitations of the most common learning techniques.
- CM24 (Apply deep learning mechanisms based on neural networks to be able to design the most suitable architecture for a given problem, checking that fundamental rights and duties and democratic values are not violated.) Apply deep learning mechanisms based on neural networks to be able to design the most suitable architecture for a given problem, checking that fundamental rights and duties and democratic values are not violated.
- KM20 (Identify the human knowledge representation techniques.) Identify the human knowledge representation techniques.
- KM21 (Define computational solutions in multiple domains to make decisions based on the exploration of alternatives, uncertain reasoning and task planning.) Define computational solutions in multiple domains to make decisions based on the exploration of alternatives, uncertain reasoning and task planning.
- SM19 (Develop optimum search schemes for different problems through knowledge representation and classification.) Develop optimum search schemes for different problems through knowledge representation and classification.

Contents

- Introduction to machine learning
- Linear Models
- Statistical Learning
- Support Vector Machines
- K-nearest neighbors
- Naive Bayes
- Decision Trees
- Random Forest
- Gradient Boosting
- Text mining
- Model Explainability

Learning activities and methodology

Title	Hours	ECTS	Learning outcomes
Project development	67	2.68	
Exams preparation	30	1.2	
Attend the theoretical and practical classes	49	1.96	

Teaching will combine classroom lessons by the teacher and hands-on work by students with a computer.

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
Theory	60%	4	0.16	CM22, CM23, KM20
Practicals	40%	0	0	CM22, CM23, CM24, KM20, KM21, SM19

Continuous Assessment

- The assessment of the course will consist of two parts: the theory part, NT, and the practical part, NP. The final grade of the course will be $N = 0.6 \cdot NT + 0.4 \cdot NP$.
- The assessment of the theory part will consist of two exams: a partial exam, NEP, and a final exam, NEF.
- The final grade for the theory part will be $NT = \max(NEF, 0.3 \cdot NEP + 0.7 \cdot NEF)$, provided that NEF is greater than 3.5; otherwise, $NT = NEF$.
- The assessment of the practical part will be carried out through the development of a project, with the possibility of a practical exam to validate the authorship of the project (in case the authorship is not validated, the practical grade NP will be 0).
- On the date of the course recovery exam, only the theory part can be recovered. If a student takes the recovery exam, the theory grade, NT, will be $NT = \min(6, NER)$, where NER is the grade of the recovery exam.
- It is mandatory to obtain a grade equal to or greater than 3.5 in each of the two sections (Theory and Practical) to pass the course. If NT or NP do not exceed 3.5, then the final grade of the course will be $N = \min(NT, NP)$.
- A student will be considered evaluable if they have submitted activities totaling at least 40% of the course weight. Otherwise, they will appear on the exam record as Non evaluable.

Single Assessment

- The assessment for students who have opted for the single assessment mode will be based on the final exam grade (60%) and the practical grade (40%).
- The same recovery system as for continuous assessment will apply.

Use of AI

In this course, the use of Artificial Intelligence (AI) technologies is allowed as an integral part of the work development, provided that the final result reflects a significant contribution from the student in terms of analysis and personal reflection. The student must clearly identify which parts were generated with this technology, specify the tools used, and include a critical reflection on how these have influenced the process and the final result of the activity. Non-transparency regarding the use of AI will be considered academic dishonesty and may result in a penalty in the activity grade or heavier sanctions in serious cases.

Bibliography

- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). An introduction to statistical learning with

- applications in Python (2nd ed.). Springer. <https://www.statlearning.com/>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer. <https://hastie.su.domains/ElemStatLearn/>

Software

The software required will be the Python programming language with the Visual Studio Code programming environment and several libraries needed for data analysis such as numpy or sklearn.

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	morning-mixed
(PLAB) Practical laboratories	1	Catalan	first semester	morning-mixed