Computational Learning

Code: 104361
ECTS Credits: 6

2024/2025

Degree	Type	Year
2503758 Data Engineering	OB	3

Contact

Name: Ramon Baldrich Caselles

Email: ramon.baldrich@uab.cat

Teaching groups languages

You can view this information at the end of this document.

Prerequisites

It is essential to have acquired a good mathematical basis as well as to have a good level of programming, mainly in Python.

Objectives and Contextualisation

The course aims to introduce the concepts of artificial intelligence that is based on obtaining knowledge, concepts and trends from the data. It is about training the student to be a "data engineer," and it is one of the professions with the most future and most in demand today by large companies and technology start-ups. In fact, the growth in demand for this data engineering professional is expected to be exponential at the European level, mainly due to the growth in mass data generation. Thus, the main objective of the subject is that the student knows how to find a good solution (sometimes the best is impossible) to problems in different contexts of the treatises, from identifying the needs of representation of the knowledge and, according to this, apply the most appropriate technique (s) to automatically generate good mathematical models that explain the data with an acceptable error.

The contents chosen for this subject are the techniques and concepts that are used extensively in the industry, understanding it in its broadest concept. The algorithmic basis will be fundamental during the development of the subject that wants to have an eminently engineering approach, focusing on the use of the proposals without leaving aside the understanding of the mathematical foundations that support them. The algorithms and techniques shown are the fundamental basis for 'traditional' computational learning without which one cannot understand the techniques that will be developed in future courses. Not because they are basic, they are obsolete, on the contrary, they cover a wide range of applications and problems where they are fundamental. The student must be aware that this knowledge that is the spearhead of the state of the art has an inherent difficulty, involving considerable study and dedication, quantified in hours in the section of formative activities of this guide. . This is because in this subject not only some of the most important contents in the field of machine learning to become a data engineer are taught, but also a curriculum line is worked that allows to expand the range of jobs to which you can access after the degree, as well as lay the methodological bases necessary to do a Master in data engineering or artificial intelligence.

The objectives of the subject can be summarized in:

Knowledges:

- Describe basic computer learning techniques.

- List the essential steps of the different learning algorithms

- Identify the advantages and disadvantages of the learning algorithms that are explained.

- Solve computational problems applying different learning techniques to find the optimal solution.

- Understand the result and limitations of learning techniques in different case studies.

- Know how to choose the most appropriate learning algorithm to solve contextualized problems.

Abilities:

- Recognize situations in which the application of computational learning algorithms may be appropriate to solve a problem

- Analyze the problem to be solved and design the optimal solution applying the techniques learned

- 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 assess possible improvements

- Defend and argue the decisions taken in solving the proposed problems

Competences

- Analyse data efficiently for the development of smart systems with the capacity for autonomous learning and/or data mining.
- Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
- Work cooperatively in complex and uncertain environments and with limited resources in a multidisciplinary context, assuming and respecting the role of the different members of the group.

Learning Outcomes

1. Choose the search algorithm and programming paradigm for a problem of optimisation of parameters or states
2. Decide on the most suitable data-learning method depending on the characteristics of the data to be analysed
3. Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
4. Work cooperatively in complex and uncertain environments and with limited resources in a multidisciplinary context, assuming and respecting the role of the different members of the group.

Content

UNIT 1: INTRODUCTION

1.1 Basic concepts and bioinspired paradigms

1.2 History of computer learning

UNIT 2: REGRESSION AND CLASSIFICATION

2.1 Regression of numerical data: gradient descent

2.2 Regularization and logistic regression

2.3 Classification of numerical data: support vector machines

2.4 Decission trees, Radom forest

2.5 Bayesian Classification

UNIT 3: CLUSTERING AND SEARCH

3.1 Memorization: lazy learning

3.2 Recommender systems: Content-based vs. Collaborative filtering

3.3 Clustering: k-means and Expectation-Maximization

Activities and Methodology

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Theoretical content	22	0.88	1, 2, 3
Type: Supervised			
lab practicums	16	0.64	1, 2, 3, 4
serminars	10	0.4	1, 2
Type: Autonomous			
Setup an ddevelopment of practical projects	52	2.08	1, 2, 3, 4
study	28	1.12	1, 2, 3

All the information on the subject and related documents that students need can be found on the Virtual Campus page (http://cv.uab.cat/).

The different activities that will be carried out in the subject are organized as follows:

Theory classes:
The main concepts and algorithms of each theory topic will be exposed. These topics are the starting point for the work of the subject.

Problem seminars:
They will be classes with small groups of students, which facilitate interaction, or of an individual nature, as the case may be. In these classes, practical cases will be presented that require the design of a solution in which

the methods seen in the theory classes are used. It is impossible to follow the problem classes if the contents of the theory classes are not followed. The result of these sessions is the resolution of the problems that will have to be delivered weekly. The specific delivery mechanism, as well as the evaluation mechanism, will be indicated on the subject's website.

This problem seminar will all be practical and will include the programming of a solution to the problem posed. This solution will have to be delivered weekly by creating a portfolio of work. Deliveries are not refundable.

Group project:
a minimum of three weeks will be available to carry out a larger project in which to put into practice what will have been practiced in the problem sessions.

Both the theory assessment and the group work will be recoverable.

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

Continous Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Individual test	30	7	0.28	1, 2, 3
Problem porfolio	10	5	0.2	1, 2, 3
Project defence	20%	5	0.2	1, 3, 4
Team project	40	5	0.2	1, 2, 3, 4

Activities and evaluation instruments:

To evaluate the achievement of the knowledge and skills associated with the subject, an evaluation mechanism is established that combines the assimilation of knowledge, the ability to solve problems, and significantly, the ability to generate computational solutions to problems complex, both as a group and individually.

With this objective, the evaluation is divided into three parts:

- Evaluation of contents

The final content grade will be calculated from several partial exams:

Note Contents = 1/N * Test_i

The number of tests may vary and will be set at the beginning of the course. In order to be able to have a content grade, the grades for each of the tests must be greater than 4.

The partial tests will be done during the course and will mainly be of conceptual content where you will answer different questions about the content developed in the 'theoretical' sessions.

These tests aim to be an individualized evaluation of the student with his abilities to understand the techniques

explained in class as well as to evaluate the level of conceptualization that the student has made of the techniques seen.

Recovery tests. In the event that the content grade does not reach the appropriate level in any of the tests, in order to obtain a final grade sufficient to consider the achievement of the knowledge, students can appear in the examination of the call for the subject and return to take an exam that evaluates the content seen in the subject of the part(s) not passed.

There are no validations if the theoretical part had been passed in previous years.

- Evaluation of the work in the problem seminars

The aim of the problems is to cause the student to enter the contents of the subject continuously and from small problems that make them familiar directly with the application of the theory. As evidence of this work, the presentation of a portfolio is requested in which you will have kept the problems you have been working on. This portfolio will have weekly digital delivery. The student will be able to self-assess continuously since he will have the solutions to each of the sets of problems once the delivery period has ended. Together with the hours of tutoring in case doubts arise, it is enough for each student to identify their weak points.

- Group project evaluation.

In the last weeks of the semester, a larger project will be carried out than the exercises done during the course in the problem sessions. This project will be evaluated both individually and as a group. The evaluation mechanisms will be code, report, presentation and monitoring of the project in the assigned class sessions.

The final mark of the subject is obtained by combining the evaluation of these 3 activities as follows:

Final Grade = (0.3 * Contents) + (0.1 * problem portfolio) + (0.6 * Project)
the project will have a mark for its defense and another for its development and depth of the solution

Conditions to pass the subject:

The final content grade must be greater than or equal to 4 in order to pass the subject.
The grade for the project and its defense must be greater than or equal to 6 in order to pass the subject.

In the event that the grade, applying the formula of the previous section ("final grade of the subject"), was higher than 5 but the minimum required in any of the parts had not been exceeded, the final grade in the record will be a 4.5.

As many honors registrations will be assigned as the current regulations allow as long as the mark is higher than 9.0. The assignment of the registrations will be done following the order of notes. In case there are multiple candidates with the same evaluation likely to receive Md'H, additional activities will be proposed to determine the best candidate (s).

The student will be graded as "Not Evaluable" if he / she has no evaluated part of either the theoretical or practical contents.

Important notices:

- The dates of continuous evaluation and delivery of works, as well as all the teaching material will be published in the virtual campus (http://cv.uab.cat/), in the space of this subject and can be subject to changes of programming for reasons of adaptation to possible incidents. Cerbero.uab.cat will always be informed about these changes as it is understood that CV will become the usual mechanism for the exchange of information between teachers and students.
- For each assessment activity, a place, date and time of review will be indicated in which the student will be able to review the activity with the teacher. In this context, claims may be made on the grade of the

activity, which will be evaluated by the teacher responsible for the subject. If the student does not appear for this review, this activity will not be reviewed later. Without prejudice to other disciplinary measures

- Without prejudice to other disciplinary measures deemed appropriate, and in accordance with current academic regulations, irregularities committed by a student that may lead to a variation in the grade in an assessable activity will be graded with a zero (0). Assessment activities qualified in this way and by this procedure will not be recoverable. If it is necessary to pass any of these assessment activities to pass the course, this course will be suspended directly, without the opportunity to retake it in the same course. These irregularities include, but are not limited to: - the total or partial copy of a practice, report, or any other evaluation activity;

- let copy;

- present a group work not entirely by the members of the group (applied to all members, not only those who have not worked);

- present as their own materials prepared by a third party,even if they are translations or adaptations, and in general works with non-original and exclusive elements of the student;

- have digital and / or communication devices (such as mobile phones, smart watches, camera pens, etc.) accessible during individual theoretical-practical assessment tests (exams).

- talk with classmates during individual theoretical-practical assessment tests (exams). - observe / look at the theoretical-practical evaluation tests (exams) of other classmates during the performance of the same, even if the copy has not been carried out.

- observe / look at the table, sheets, wall, etc. writings related to the subject during the realization of the theoretical-practical evaluation tests (exams) even if the copy has been proceeded.

The numerical mark of the transcript will be the lower value between 3.0 and the weighted average of the marks in case the student has committed irregularities in an act of evaluation (and therefore it will not be possible to pass it for compensation). In short: copying, copying or plagiarizing (or attempting to) in any of the assessment activities is equivalent to a SUSPENSION with a grade below 3.5.

No unique evaluation planned

Bibliography

Web links

- Subject web page: http://cv.uab.cat
- Artificial Intelligence: A Modern Approach. http://aima.cs.berkeley.edu/

Basic Bibliography

- S. Russell, P. Norvig. Artificial Intelligence: A Modern Approach. Ed. Prentice Hall, Second Edition, 2003. (Existeix traducció al castellà: Inteligencia artificial: Un Enfoque Moderno)
- T. Mitchell. Machine Learning. McGraw Hill. 1997.

Additional Bibliography

- C. Bishop. Pattern Recognition and Machine Learning. Springer-Verlag New York, Inc. 2006

Software

No special software other than the usual ones will be used in these studies.

Language list

Name	Group	Language	Semester	Turn
(PAUL) Classroom practices	811	Catalan	first semester	afternoon
(PAUL) Classroom practices	812	Catalan	first semester	afternoon