Universitat Autònoma de Barcelona

2020/2021

Computational Learning

Code: 104361
ECTS Credits: 6

Degree	Type	Year	Semester
2503758 Data Engineering	OB	3	1

The proposed teaching and assessment methodology that appear in the guide may be subject to changes as a result of the restrictions to face-to-face class attendance imposed by the health authorities.

Contact

Name: Ramón Baldrich Caselles
Email: Ramon.Baldrich@uab.cat

Use of Languages

Principal working language: catalan (cat)
Some groups entirely in English: No
Some groups entirely in Catalan: Yes
Some groups entirely in Spanish: No

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

3.4 Genetic algorithms

Methodology

All the information of the subject and the related documents that the students need will be found in the page Virtual Campus (http://cv.uab.cat/).
The different activities that will be carried out in the subject are organized as follows:

Lectures

The main concepts and algorithms of each theory topic will be presented. These subjects suppose the starting point in the work of the subject.

Problem seminars

They will be classes with small groups of students, which facilitate the interaction, or of individual character, according to the cases. In these classes, practical cases will be considered that require the design of a solution in which the methods seen in the theory classes are used. It is impossible to follow the kinds of problems if the contents of the theory classes are not followed. The result of these sessions is the resolution of the problems that must be delivered on a weekly basis. The specific mechanism for the delivery, and the evaluation process, will be indicated on the web page of the subject (Charon space).

Laboratory practicum

The working groups will be formed by groups of 3-4 students and should form the second week of the course. These working groups must be maintained until the end of the course and they must self-manage: role distribution, work planning, assignment of tasks, management of available resources, conflicts, etc. Although the teacher will guide the learning process, his intervention in the management of the groups will be minimal.

At the beginning of the course, the problems to be solved will be presented and the students will define their own project. Throughout the semester, students will work in cooperative groups and should analyze the chosen problem, design and implement solutions based on different computational learning algorithms seen in class, analyze the results obtained in eachof themethods and defend their project inpublic.

To developthe project, the groups will work autonomously and the practice sessions will be devoted mainly to answer questions with the teacher who will monitor the status of the project, indicate errors to be corrected, propose improvements, etc.

Some of the sessions will be marked as control sessions in which some part of the project must be delivered. In these sessions the groups must explain the work done and the teacher will ask questions to all group members to assess the work done. Attendance at these sessions is mandatory.

In the last session of each of the internship projects, the groups will make a presentation of the project where they will explain the project developed, the solution adopted and the results obtained. In this presentation each member of the group must make a part of the presentation.

Activities

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

Assessment

Evaluation activities and instruments:

To evaluate the achievement of 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 complex problems, both group as individually.

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

- Evaluation of theoretical-practical contents

The final grade of contents will be calculated from two partial exams:

Note Contents = 1 / N * test_i

The number of tests may vary and they will be at least 2. In order to have a content grade, it will be necessary for the grades of each of the tests to be greater than 4.

The partial tests will be carried out during the course and may be of a practical nature (algorithm proposal for the resolution of a statement), or be of conceptual content where to answer different questions about the content developed in the 'theoretical' sessions.

These tests are intended to be an individualized evaluation of the student with their abilities to solve problems using the techniques explained in class as well as evaluating the level of conceptualization that the student has made of the techniques seen.

Recovery exam. In the event that the content grade does not reach the appropriate level in any of the tests, in order to obtain a sufficient final grade to consider the achievement of knowledge, students can sit for the exam to call the subject and retake an exam that evaluates the contents seen in the subject of the part / s not passed. In case of showing up to raise the grade, the high note prevails.

There are no validations in case the theoretical part had been exceeded in previous years.

- Evaluation of the practical projects

The evaluation of each of the projects will include:

- Joint evaluation of the project: a single note for all members of the working group that will assess the overall result of the project, the quality of the code, the general structure of the final presentation and the documents delivered throughout the project.
- Individual assessment to each member of the group: the individual work will be valued based on the answers to the questions in the control sessions and the final presentation of the project.
- Peer evaluation: brief confidential form rating each groupmate's contribution to the final result.

The grade of the project will be calculated according to the formula:

Project Note = (0.6 * Group Note) + (0.3 * Individual Note) + (0.1 * Av. Between equals)
Group Note = (0.6 * Program) + (0.2 * Presentation) + (0.2 * Documentation)
Individual Note = (0.5 * Individual work) + (0.5 * Presentation)

The final mark of practices will be the average of the marks of the projects, having to obtain a minimum of 3.5 in each one of them. In the event that this minimum is not reached in any of the projects, the final internship mark will have a maximum of 3.5.

Failure to pass any of the practical projects will allow the recovery of the code and memory of the failed projects, but not the oral presentation.

In case of repeating students with the practical part passed (minimum 6) the previous year exclusively, they will be able to present the practices of the previous year again, adding functionality or modifying the data according to the practical teacher if the content of the project is the same or similar the year before. These students in no case may group with first year students.

There may be group projects and individual projects. Obviously, in the second case, all group notes will become individual

- Evaluation of the work in the problem seminars

The problems are intended to cause the student to enter the contents of the subject continuously and from small problems that make them become directly familiar with the application of the theory. As evidence of this work, the presentation of a portfolio in which she will have kept the problems she has been carrying out is requested. This portfolio will have weekly digital delivery. The student will be able to self-evaluate continuously since they will have the solutions of each one of the problem sets once the delivery period has ended. Along with the hours of tutoring in case doubts appear, it is enough for each student to identify their weak points.

The final grade for the course is obtained by combining the evaluation of these 3 activities as follows:

Final Note = (0.4 * Contents) + (0.5 * Project) + (0.1 * Portfolio)

Conditions to pass the subject:

- The final grade for theory must be greater than or equal to 4 to pass the course.
- The grade for the project must be greater than or equal to 6 to pass the course.

In the event that the grade, applying the formula of the previous section ("final grade of the subject"), is higher than 5 but the minimum required in any of the parts has 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 grade 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 Caronte site (http://caronte.uab.cat/), in the space of this subject and can suffer programming changes because of adaptation to possible incidents. Cerbero.uab.cat will always be informed about these changes since it is understood that Charon will convert the usual mechanism of exchange of information between teacher and students.

For each evaluation activity, a place, date and time of revision in which the student can review the activity with the teacher will be indicated. In this context, claims may be made on the activity grade, which will be evaluated by the faculty responsible for the subject. If the student does not appear in this revision, it will not be revised later to this activity.

Notwithstanding other disciplinary measures deemed appropriate, and in accordance with the academic regulations in force, assessment activities will receive a zero whenever a student commits academic irregularities that may alter such assessment. Assessment activities graded in this way and by this procedure will not be re-assessable. If passing the assessment activity or activities in question is required to pass the subject, the awarding of a zero for disciplinary measures will also entail a direct fail for the subject, with no opportunity to re-assess this in the same academic year. Irregularities contemplated in this procedure include, among others:

• the total or partial copying of a practical exercise, report, or any other evaluation activity;

• allowing others to copy;

• presenting group work that has not been done entirely by the members of the group;

• presenting any materials prepared by a third party as one's own work, even if these materials are translations or adaptations, including work that is not original or exclusively that of the student;

• having communication devices (such as mobile phones, smart watches, etc.) accessible during theoretical-practical assessment tests (individual exams).

The numerical note of the subject will be the lower value between 3.0 and the weighted average of the marks in case the student has committed irregularities in an evaluation act (and therefore the approved by compensation will not be possible). In summary: copy, leave copy or plagiarize (or attempt to) in any of the evaluation activities is equivalent to a FAIL with a grade lower than 3.5.

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Individual test	40	7	0.28	2, 1, 3

Practicum defense (report + code + presentation + follow-up)	50%	10	0.4	1, 3, 4
Problem porfolio	10	5	0.2	2, 1, 3

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