Fundamentals of Machine Learning

Code: 106572
ECTS Credits: 6

Degree	Type	Year	Semester
2504392 Artificial Intelligence	OB	2	1

Contact

Name: Dimosthenis Karatzas

Email: dimosthenis.karatzas@uab.cat

Teaching groups languages

You can check it through this link. To consult the language you will need to enter the CODE of the subject. Please note that this information is provisional until 30 November 2023.

External teachers

Oguz Mulayim

Prerequisites

To have completed the first-year subjects of Fundamentals of Programming, Fundamentals of Mathematics, Probability and Statistics, and Data Engineering.

Objectives and Contextualisation

The objective of this subject is to provide comprehensive training in a wide range of fundamental machine learning techniques and algorithms that an engineer specializing in artificial intelligence must master.

Students will learn, implement and use a range of techniques covering both supervised and unsupervised learning. They will derive their own implementations, as well as explore existing libraries, and apply this new knowledge to solving practical problems.

Competences

- Develop critical thinking to analyse alternatives and proposals, both one's own and those of others, in a well-founded and argued manner.
- Identify, understand and apply the fundamental concepts and techniques of knowledge representation, reasoning and computational learning for the solution of artificial intelligence problems.
- Introduce changes to methods and processes in the field of knowledge in order to provide innovative responses to society's needs and demands.

- 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.
- Students must have and understand knowledge of an area of study built on the basis of general secondary education, and while it relies on some advanced textbooks it also includes some aspects coming from the forefront of its field of study.
- Work cooperatively to achieve common objectives, assuming own responsibility and respecting the role of the different members of the team.

Learning Outcomes

1. Develop critical thinking to analyse alternatives and proposals, both one's own and those of others, in a well-founded and argued manner.
2. Formulate the most appropriate cost function for a specific problem according to the characteristics of the experimental data and to the requirements/restrictions of the problem.
3. Propose new methods or informed alternative solutions.
4. 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.
5. Students must have and understand knowledge of an area of study built on the basis of general secondary education, and while it relies on some advanced textbooks it also includes some aspects coming from the forefront of its field of study.
6. Understand the concepts of bias and variance, and be able to use data preparation methods and regularisation techniques to obtain generalisable solutions from the available data.
7. Understand the gradient descent algorithm and its variants and how to apply it to solve regression and classification problems.
8. Work cooperatively to achieve common objectives, assuming own responsibility and respecting the role of the different members of the team.

Content

- Machine learning fundamentals
  - Supervised and unsupervised learning
  - Regression and classification
  - Gradient descent
  - Linear and polynomial regression
  - Logistic regression
  - Over/under-fitting and decomposition of bias and variance
  - Regularization
- Basic methods of pattern recognition
  - Nearest Neighbour models and case-based reasoning
  - Data clustering
  - Kernel Methods and Support Vector Machines
  - Decision trees and random forests
  - Experimentation and performance evaluation
- Probabilistic methods
  - Bayesian parameter estimation
  - Gaussian Mixture Models
  - Probability density estimation
  - Expectation Maximization

Methodology

Machine Learning techniques and methodologies are intimately related to the types of problems that they aim to solve. Throughout this subject it will be that typology of problems that will provide the motivation of each section and will direct the organization of the contents.

There will be two types of sessions:

Theory classes: The objective of these sessions is for the teacher to explain the theoretical background of the subject. For each one of the topics studied, the theory and mathematical formulation is explained, as well as the corresponding algorithmic solutions.

Laboratory sessions: Laboratory sessions aim to facilitate interaction, collaborative work and to reinforce the comprehension of the topics seen in the theory classes. During laboratory sessions the students will work through practical cases that require the design of solutions using the methods studied in theory classes. Problem solving will be initiated in the class and will be complemented by a weekly *set of problems* to work through at home.

The above activities will be complemented by a system of tutoring and consultations outside class hours.

Course Project: A course project will be carried out during the semester, where students will have to solve a specific problem of certain complexity. The students will work in small groups of 2-3 students, where each member of the group must equally contribute to the final solution. These working groups will be maintained until the end of the semester and must be self-managed in terms of distribution of roles, work planning, assignment of tasks, management of available resources, conflicts, etc. To develop the project, the groups will work autonomously, while some of the laboratory sessions will be used (1) for the teacher to present the projects' theme and discuss possible approaches, (2) for monitoring the status of the project and (3) for the teams to present their final results.

All the information of the subject and the related documents that the students need will be available at the virtual campus (cv.uab.cat).

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.

Activities

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Lab Sessions	21	0.84	7, 1, 6, 2, 3, 4, 8
Theory Classes	24	0.96	7, 6, 2
Type: Supervised			
Tutoring	2	0.08	7, 1, 6, 2, 3, 4
Work on problems / projects	40	1.6	7, 1, 6, 2, 3, 4, 8
Type: Autonomous			
Individual studying	20	0.8	7, 1, 6, 2
Work on problems / projects	30	1.2	7, 1, 6, 2, 3, 4, 8

Assessment

To assess the level of student learning, a formula is established that combines theoretical and practical knowledge acquisition, and the ability to solve problems.

Final Grade

The final grade is calculated weighted in the following way and according to the different activities that are carried out:

Final grade = 0.5 * Theory Grade + 0.1 * Problems Portfolio Grade + 0.4 * Project Grade

This formula will be applied as long as the theory and the practical evaluation grades are higher than 5. There is no restriction on the problems portfolio grade. If doing the calculation of the formula yields >= 5 but the theory grade or practical evaluation grade does not reach the minimum required, then a final grade of 4.5 will be given.

Theory Grade

The theory grade aims to assess the individual abilities of the student in terms of the theoretical content of the subject, this is done continuously during the course through two partial exams. The overall theory grade is the average of the grades of the two partial exams.

The mid-term exam (Exam 1) is done in the middle of the semester and serves to eliminate part of the subject if it is passed. The final exam (Exam 2) is done at the end of the semester and serves to eliminate the rest of subject if it is passed.

These exams aim to assess the abilities of each student in an individualized manner, both in terms of solving problems using the techniques explained in class, as well as evaluating the level of conceptualization that the student has made of the techniques seen. In order to obtain a final pass theory grade, it will be required for the both partial exam grades to be higher than 4. If doing the calculation of the formula yields >= 5 but the grades of any of the two partial exams do not reach the minimum required, then the grade for theory that will be used in the final calculation will be 4.5.

Recovery exam. In case the theory grade does not reach the adequate level to pass, the students can take a recovery exam, destined to recover the failed part (1, 2 or both) of the continuous evaluation process.

Problems Portfolio Grade

The aim of the problems (exercises) is for the student to train with the contents of the subject continuously and become familiar with the application of the theoretical concepts. As evidence of this work, the presentation of a portfolio is requested in which the exercises worked out will be collated. To obtain a problems grade it is required that the student submits a minimum of 70% of the problem sets, in the opposite case, the problems grade will be 0.

Project Grade

The course project carries an essential weight in the overall mark of the subject. Developing the project requires that the students work collaboratively and design an integral solution to the defined challenge. In addition, the students must demonstrate their teamwork skills and present the results to the class.

The project is evaluated through its deliverables, an oral presentation that students will make in class, and a self-evaluation process. The participation of students in all three activities (preparing the deliverable, presentation and auto evaluation) is necessary in order to obtain a projects grade. The grade is calculated as follows:

Project Grade = 0.6 * Grade Deliverables + 0.3 * Grade Presentation + 0.1 * Grade Self-evaluation

If performing the above calculation yields >= 5 but the student did not participate in any of the activities (deliverable, presentation, auto evaluation), then a final grade of 4.5 will be given to the course project.

In case the minimum grade is not achieved, the group will be able to recover the project, restricted to a maximum grade of 7/10.

In case of not submitting the project deliverables, not presenting the project, or considering the contents copied or synthetically generated, the course project will be considered failed.

Important notes

Notwithstanding other disciplinary measures deemed appropriate, and in accordance with the academic regulations in force, evaluation activities will be suspended with zero (0) whenever a student commits any academic irregularities that may alter such evaluation (for example, plagiarizing, copying, letting copy, ...). The evaluation activities qualified in this way and by this procedure will not be recoverable. If you need to pass any of these assessment activities to pass the subject, this subject will be failed directly, without opportunity to recover it in the same course.

In case there the student does not deliver any exercise solutions, does not participate in any practical evaluation, and does not take any exam, the corresponding grade will be a "non-evaluable". In any other case, the "no shows" count as a 0 for the calculation of the weighted average.

In order to pass the course with honours, the final grade obtained must be equal or higher than 9 points. Because the number of students with this distinction cannot exceed 5% of the total number of students enrolled in the course, it is given to whoever has the highest final marks. In case of a tie, the results of the partial exams will be taken into account.

More details about the evaluation process will be given at class during the first weeks of the semester. In case of any discrepancy between this guide and the information given at class, the information provided at class takes precedence.

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Practical examination	40%	4	0.16	1, 3, 5, 4, 8
Problem deliverables	10%	5	0.2	7, 1, 6, 2
Written exams	50%	4	0.16	7, 6, 2, 5

Bibliography

- Data Science from Scratch: First Principles with Python, Joel Grus, O'Reilly Media, 2015, 1st Ed.
- Pattern Recognition and Machine Learning, Christopher Bishop, Springer, 2011
- Model-Based Machine Learning, J. Winn, C. Bishop, early access: http://mbmlbook.com/
- Introduction to Data Science. L. Igual, S. Seguí. Springer, 2017
- Pattern Classification, Duda, Hart, and Stork, 2nd Ed., 2002

Software

For the problems and projects of the course we will use Python, and the Python libraries NumPy, MatPlotLib, SciKit Learn, Pandas