

Degree programme	Type	Course
Bachelor in Artificial Intelligence	FB	1

Contact lecturer

Name : Javier Vazquez Corral

Email : javier.vazquez.corral@uab.cat

Prerequisites

Students should have cursred and understood the subject of "Introduction to Programming I" and "Mathematical Foundations I".

Group languages

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

Objectives

The subject aims to provide the fundamentals of data analysis and visualization. The different stages of data analysis processes will be studied, from the collection, annotation and preparation of data, to its analysis and visualization, preparing the way for a more advanced modelling through Machine Learning (M8).

Learning outcomes

- KM15 (Describe the basic tools for manipulating different types of structured, semi-structured, and unstructured data.) Describe the basic tools for manipulating different types of structured, semi-structured, and unstructured data.
- SM17 (Develop data collection, annotation, analysis, and visualisation processes to test hypotheses about the data and to interpret and process the data needed to address an AI problem.) Develop data collection, annotation, analysis, and visualisation processes to test hypotheses about the data and to interpret and process the data needed to address an AI problem.

Contents

This subject will be divided into 11 main topics:

- Introduction.
- Basic statistics Recap.
- Numpy / Matplotlib/ Pandas
- XML/JSON
- Regular Expressions/DFA/NFA
- Data types. Missing data.

- Introduction to data analysis. Outliers.
- PCA
- kNN
- Image retrieval and recommender systems.
- K-means

Learning activities and methodology

Title	Hours	ECTS	Learning outcomes
Exercise classes	12	0.48	KM15, SM17
Theory classes	28	1.12	KM15, SM17
Project classes	12	0.48	KM15, SM17
Finishing exercises at home	16	0.64	KM15, SM17
Project labs	62	2.48	KM15, SM17
Study of the theory	20	0.8	KM15, SM17

There will be three 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.

Exercise sessions: They will be sessions that facilitate interaction. In these sessions, the aim is to reinforce the comprehension of the topics seen in the theory classes by proposing practical cases that require the design of a solution in which the methods seen in the theory classes are used.

Practical laboratory sessions: They will be sessions in which different types of activities related to the realization of the project/projects by teams of students will be carried out. During the practical sessions the project/projects to be solved will be presented and a series of activities will be carried out in teams of students in collaborative work mode. The identification of the problem, the discussion of the design, the distribution and organization of the work to be carried out, the development of the solution and the presentation of the results to the teacher and the rest of the students will be addressed.

All the information of the subject and the related documents that the students need will be found in the virtual campus.

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.

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.

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
Exam 2	0.225	0	0	KM15, SM17
class exercises	0.1	0	0	KM15, SM17
Project	0.45	0	0	KM15, SM17
Exam 1	0.225	0	0	KM15, SM17

Final mark:

The final grade is calculated using a weighted average according to the different activities that are carried out:

Final grade = $0.45 \times \text{Theory Grade} + 0.1 \times \text{Exercises Grade} + 0.45 \times \text{Laboratory Grade}$

For applying this formula, the condition is that both the theory and the laboratory grades, are higher than 5. When a student does not reach the minimum required in some one of the evaluation activities, then the final grade will be that of the element not allowing the calculation (i.e. if a student has a 6 in Theory, a 5 in Exercises, but a 2 in Laboratory, the final mark will be a 2).

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:

Theory Grade = $0.5 \times \text{Grade Exam 1} + 0.5 \times \text{Grade Exam 2}$

Exam 1 is done in the middle of the semester and serves to eliminate part of the subject if it is passed. Exam 2 is done at the end of the semester and serves to eliminate part of the subject if it is passed. These exams aim to assess the abilities of each student in an individualized manner, both in terms of solving exercises 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 partial exam grades 1 and 2 to be both higher than 4.

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. Note: The recovery exam can also be taken if the student wants to obtain a higher grade than in the previous ones. Having said this, this extra examen will not allow to obtain the "Honors" qualification.

Exercises Grade

The aim of the 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:

Exercises Grade = Portfolio evaluation

Practical Laboratory Sessions Grade

The part of laboratory based practical sessions carries an essential weight in the overall mark of the subject. Laboratory sessions aim for the student to design a solution to a problem that is set out in a contextualized way. Such problems will require the design of an integral solution, from the exploration of available techniques to data modelling. In addition, the students must demonstrate their teamwork skills and present the results to the class convincingly. Laboratory sessions are structured around project/s. The projects are evaluated through its deliverable, an oral presentation that students will make in class, and a self-evaluation process. The grade is calculated as follows:

$$\text{Project Grade} = 0.65 * \text{Grade Deliverables} + 0.3 * \text{Grade Presentation} + 0.05 * \text{Grade Self-evaluation}$$

In case of being more than 1 project, the Laboratory Grade will be the average of all the Different Project Grades. In case of not passing the project, the recovery of the part of the deliverables of the unsuccessful projects will be allowed, restricted to a maximum grade of 7/10. The oral presentation cannot be recovered.

Bibliography

Data Science from Scratch: First Principles with Python, Joel Grus, O'Reilly Media, 2015, 1st Ed.

Python Data Science Handbook, Jake Van der Plas, O'Reilly Media, 2016, 1st Ed.

Computational and Inferential Thinking: The Foundations of Data Science, Ani Adhikari and John DeNero, online: <https://ds8.gitbooks.io/textbook/content/>

Software

This subject will use Python as the 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	English	second semester	afternoon
(PAUL) Classroom practices	1	English	second semester	afternoon
(PLAB) Practical laboratories	1	English	second semester	afternoon
(PLAB) Practical laboratories	2	English	second semester	afternoon