

Degree	Type	Year
2504392 Artificial Intelligence	OB	2

## Contact

Name: Alicia Fornes Bisquerra

Email: [alicia.fornes@uab.cat](mailto:alicia.fornes@uab.cat)

## Teachers

Pau Torras Coloma

## Teaching groups languages

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

## Prerequisites

There are no official prerequisites but it is recommended to have completed the subjects of Fundamentals of Programming I and II, Fundamentals of Mathematics I and II, Probability and Statistics, Data Engineering and Fundamentals of Machine Learning.

## Objectives and Contextualisation

This course provides an overview of the fundamentals techniques for natural language processing (NLP), covering classical approaches for text processing and parsing, language and sequence modelling and text representation, showing their application to usual NLP problems. The course also covers an introduction to the application of recent deep learning techniques to NLP. The content of the course will be expanded in subsequent optional courses, where deep learning-based approaches will be covered in greater depth, along with other more advanced topic such as semantic analysis, language generation or speech processing.

By the end of this course, students will be able to:

- Understand the fundamental concepts and techniques used in NLP.
- Implement and evaluate various NLP techniques using Python and popular NLP libraries.
- Apply NLP methods to real-world problems and interpret the results.

## Competences

- Analyse and solve problems effectively, generating innovative and creative proposals to achieve objectives.
- Develop critical thinking to analyse alternatives and proposals, both one's own and those of others, in a well-founded and argued manner.
- Develop strategies to formulate and solve different learning problems in a scientific, creative, critical and systematic way, knowing the capabilities and limitations of the different existing methods and tools.
- 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.
- Know and apply the techniques of natural language processing for the exploitation of data of linguistic nature and the creation and evaluation of the components of language-based AI systems.
- Know, understand, use and apply appropriately the mathematical foundations necessary to develop systems for reasoning, learning and data manipulation.
- Students can apply the knowledge to their own work or vocation in a professional manner and have the powers generally demonstrated by preparing and defending arguments and solving problems within their area 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. Analyse a situation and identify areas for improvement.
2. Analyse and solve problems effectively, generating innovative and creative proposals to achieve objectives.
3. Develop critical thinking to analyse alternatives and proposals, both one's own and those of others, in a well-founded and argued manner.
4. Develop solutions for specific natural language processing projects.
5. Students can apply the knowledge to their own work or vocation in a professional manner and have the powers generally demonstrated by preparing and defending arguments and solving problems within their area of study.
6. Understand and apply fundamental natural language and speech modelling techniques.
7. Understand and use algebraic representations of alphabets, words and languages by means of formal languages such as automata and grammars.
8. 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.
9. Understand, apply and adapt methodologies for evaluating and analysing natural language processing systems.
10. Understand, use and apply the mathematical foundations necessary for natural language processing.
11. Weigh up the risks and opportunities of both your own and others' proposals for improvement.
12. Work cooperatively to achieve common objectives, assuming own responsibility and respecting the role of the different members of the team.

## Content

1. Introduction to linguistics and NLP
2. Basic text processing
3. Syntactic parsing
4. Language modelling
5. Sequence labelling
6. Text embeddings
7. Deep learning for language processing

## Activities and Methodology

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Problems sessions	16	0.64	1, 2, 3, 4, 5, 6, 7, 9, 10, 11
Project sessions	4	0.16	1, 2, 3, 4, 5, 6, 9, 10, 12
Theory classes	25	1	1, 3, 5, 6, 7, 8, 9, 10
Type: Supervised			
Work on the project	50	2	1, 2, 4, 5, 6, 10, 12
Type: Autonomous			
Individual studying	24	0.96	5, 6, 7, 8, 9, 10
Problem solving (individual)	25	1	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11

There will be three types of teaching activities: theory classes, solving practical exercises individually (problems) and developing a project in small groups of 2-3 students.

1. Theory classes: Presentation of the theoretical content of the subject. For each of the topics studied, the main theoretical concepts and mathematical formulation are exposed, as well as the corresponding algorithmic solutions.
2. Laboratory sessions: The laboratory sessions aim to facilitate interaction and reinforce the understanding of the topics covered in the theory classes. During the laboratory sessions, we will address two types of activities: the resolution of practical exercises (problems) and the monitoring and presentation of projects.

2.1. Problems: A set of problems to work through will be used, provided in Jupyter notebooks that exemplifies the coding details of the concepts exposed during theory classes. Work on the problems will begin in class and must be completed at home. Students will be required to make regular submissions of their work, which will comprise the problems portfolio.

2.2. Project: A project will be carried out during the semester, where students will have to solve a specific problem of certain complexity. The project will be solved in small groups of 2-3 students, where each member of the group must contribute a part and put it together with the rest to obtain the final solution. These working groups must 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 the practical sessions will be used (1) for the teacher to present the project theme and discuss possible approaches, (2) for monitoring the status of the project and (3) for the teams to present their final results.

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

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

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	Weighting	Hours	ECTS	Learning Outcomes
Exams	40	4	0.16	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11
Problem solving	20	0	0	1, 2, 4, 5, 6, 8, 9
Project	40	2	0.08	1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12

The evaluation will be continuous, so there is no unique evaluation modality. To assess the level of student learning, a formula is established that combines knowledge acquisition, the ability to solve problems and the ability to work as a team, as well as the presentation of the results obtained.

#### Final grade

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

$$\text{Final grade} = 0.4 * \text{Theory Grade} + 0.2 * \text{Problems Grade} + 0.4 * \text{Project Grade}$$

This formula will be applied as long as the theory and the Project grades, are higher than 5. There is no restriction on the problems grade. If doing the calculation of the formula yields  $\geq 5$  but the student does not reach the minimum required in any of the evaluation activities, 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:

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

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.

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 4.5 and the average to be above 5.0.

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 Grade

The aim of the problems is for the student to become familiar with the practical implementation of the theoretical concepts. The regular submission of problem solutions will be used as evidence of this work.

In order to obtain a grade for exercises, it is necessary that more than 50% of the exercises are submitted during the semester. In the contrary, the problems grade will be 0.

In each of the two partial exams there will be some questions about the problems of that part of the subject. The final problems grade will be the combination of the problems portfolio and these questions in the exam.

$$\text{Problems Grade} = 0.5 * \text{Portfolio evaluation} + 0.5 * \text{Exam questions}$$

## Project Grade

The project carries an essential weight in the overall mark of the subject. Developing the project requires that the students work in groups 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 deliverable, an oral presentation that students will make in class, and an individual-evaluation process. The participation of students in all three activities (preparing the deliverable, presentation and individual 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 Individual evaluation

If performing the above calculation yields  $\geq 5$  but the student did not participate in any of the activities (deliverable, presentation, individual evaluation), then a final grade of 4.5 will be given to the corresponding project.

In case the deliverable is presented, but the final project grade does not reach the minimum of 5, there will be a recovery of the project. In case of not presenting the deliverable or considering it copied, there will be no recovery and the subject will be considered failed. The maximum project grade that can be obtained in case of recovery is 7.

## 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 year.

In case the student does not deliver any problems solutions, does not attend any project presentation session during the laboratory sessions and does not take any exam, the corresponding grade will be a "non-evaluable". In another 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.

## Bibliography

- D. Jurafsky, JH Martín. *Procesamiento del habla y el lenguaje* . Tercera edición. 2021 < <https://web.stanford.edu/~jurafsky/slp3/> >
- J. Eisenstein. *Procesamiento del lenguaje natural* . 2018. Prensa del MIT
- H. Lane, C. Howard, HM Hapke. *Procesamiento del lenguaje natural en acción* . 2019. Publicaciones de Manning
- Kenny, Dorothy, ed. *Traducción automática para todos* . Prensa científica lingüística, 2022. < <https://langsci-press.org/catalog/book/342> > \_ \_
- Rowe, Bruce M. y Diane P. Levine. *Una breve introducción a la lingüística* . Routledge, 2018.

## Software

For the problems and projects of the course we will use Python, along with some Python libraries for NLP that will be specified during the course.

## Language list

Name	Group	Language	Semester	Turn
(PAUL) Classroom practices	711	English	second semester	afternoon
(PLAB) Practical laboratories	711	English	second semester	afternoon
(PLAB) Practical laboratories	712	English	second semester	afternoon
(TE) Theory	71	English	second semester	afternoon