

Name of the unit: Natural language processing

Lecturer: Laure Soulier (laure.soulier@sorbonne-univ.fr)

Course code: UM4RBI20-Langage

Student workload: 10h of lectures, h of tutorial sessions, 16h of lab sessions@

Credits: 3 ECTS

Specialization tracks:

- E3A :	☐ CIMES	☐ Syscom	☐ IPS
- AR :	☐ ROBOT	☐ SI	☐ MeDH
- MECA :	☐ MS2	☐ MF2A	☐ CompMech
	☐ ACOU	☐ EE	☐ EE APP

Semester offered: ☐ S1 ☒ S2 ☐ S3 ☐ S4

Language of instruction: ☐ French ☐ English

Targeted audience: ☐ Engineering Sciences students ☐ Other (specify):

Localization : ☒ PMC Campus ☐ Other (specify):

Course Overview

This course presents the major issues surrounding text processing, particularly automatic language processing. The aim of this course is therefore to present and manipulate the main models for analyzing, synthesizing, exploiting/querying, and producing documents. This course focuses on the notion of text element semantics (theme extraction, latent and contextual representations, etc.). Text classification and mining models (sentiment detection, etc.) as well as text enrichment using knowledge bases are also introduced. Finally, the latest advances in the field from a deep learning perspective are presented.

Keywords: Automatic text processing, automatic language processing, semantics, classification, deep learning.

Prerequisites

Students should have previously acquired the following prerequisites to follow this course:

- Python programming
- Concepts of statistics and machine learning

Intended Learning Outcomes

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

1. Text data manipulation: encoding and utilization
2. Extracting semantics from a text
3. Leveraging deep learning techniques for text data manipulation

Indicative Teaching Sequence and Methods

Week	C/TD/TP*	Content	Preparation	Learn. outc.
S1	C1-5	Presentation of technical methods for automatic language processing		
S2	TP1-5	Practical implementation of the course		

* C/TD/TP respectively corresponds to lectures, tutorials and lab sessions.

Indicative Assessment of Intended Learning Outcomes (1st session)

Week	Individ./group	In-person/remote	Type of exam	Evaluated outcomes	Scale %
S15	Individual	In-person	Labs	All	50 %
S15	Individual	In-person	Written	All	50 %

2nde session of exam

Session	Individ./group	In-person/remote	Type of exam	Evaluated outcomes	Scale %
2	Individual	In-person	Written		50 %
1	Individual	In-person	Labs		50 %