Master's Department of Engineering Sciences

MSc in Automation Engineering and Robotics

Name of the unit: Multi-modal analysis of human behaviour

Lecturer: Mohamed CHETOUANI (mohamed.chetouani@sorbonne-universite.fr)
Course code:
Student workload: 10 h of lectures, 4 h of tutorials, 16 h of labs
Credits: 3 ECTS
Specialization tracks: ☒ ROBOT (elective) ☒ SI (elective) ☒ MeDH (elective)
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

Multimodal data collection: sensors, ethics and data protection Human modeling: human behavior Learning multimodal representations Visual and textual modalities, audio and temporal modalities Physiological modalities: EEG, EMG, ECG, eye tracking, pupillometry, galvanic skin response and respiration Multimodal fusion: methods and architectures Multimodal alignment and grounding Multimodal foundation models

Keywords: Multimodal machine learning, multimodal fusion, human-centered AI

Prerequisites

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

- Fundamental principles of machine learning (supervised learning, reinforcement learning)
- Fundamental principles of signal processing (frequency-domain representation)

Intended Learning Outcomes

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

1. Understand the fundamental ethical principles of experimental research involving human participants.
2. Explore AI methods and models for multimodal signal processing.
3. Analyze data from wearable sensors and conduct statistical analyses for behavior detection.
4. Develop multimodal systems for human-behavior analysis.

Indicative Teaching Sequence and Methods

Week	C/TD/TP*	Content	Preparation	Learn. outc.
S1	C (2h)	Introduction to multimodal data collection, ethical and legal framework, sensors, experiments		AAV1
S2	TD (2h)	Ethical workshop		AAV1
S3	C (2h)	Multimodal representations: audio, video, and text		AAV2
S4	C (2h)	Multimodal representations: physiological signals, EEG, EMG, ECG, eye tracking, pupillometry, GSR, and respiration		AAV3
S5	TP (4h)	Representation learning and fusion		AAV2
S6	C (2h)	Multimodal foundation models		AAV2
S7	C (2h)	Multimodal foundation models		AAV2
S8	TP (4h)	Multimodal foundation models		AAV2
S9				
S10	TP (4h)	Final Project — Multimodal Learning		AAV4
S11				
S12	TP (4h)	Final Project — Multimodal Learning		AAV4
S13				
S15	TD (2h)	Final project presentation		AAV4

* 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	Group	In-person	Other	AAV1-AAV4	80 %
S6	Individual	In-person	TP	AAV2	20 %

2nde session of exam

Session	Individ./group	In-person/remote	Type of exam	Evaluated outcomes	Scale %
2	Group	In-person	Written exam	AAV 1-10	50 %
1	Group	In-person	Quiz	AAV1-8	50 %

Date of generation of this unit description: October 6, 2026