



ADVANCED BIOMEDICAL MACHINE LEARNING

Enrollment year	2020/2021
Academic year	2021/2022
Regulations	DM270
Academic discipline	ING-INF/06 (ELECTRONIC AND INFORMATION BIOENGINEERING)
Department	DEPARTMENT OF ELECTRICAL, COMPUTER AND BIOMEDICAL ENGINEERING
Course	COMPUTER ENGINEERING
Curriculum	Data Science
Year of study	2°
Period	2nd semester (07/03/2022 - 17/06/2022)
ECTS	6
Lesson hours	46 lesson hours
Language	English
Activity type	WRITTEN TEST
Teacher	DAGLIATI ARIANNA (titolare) - 3 ECTS ABU-HANNA AMEEN - 3 ECTS
Prerequisites	Since this is an advanced course, it is advisable to have some basic knowledge on machine learning
Learning outcomes	In this one-semester course, we will explore a variety of advanced machine learning methods for mining clinical data. Much of this exploration will be a hands-on experience, using time in class to expose the principles of each method. During the course we will be using two softwares: R and KNIME. R is a state of the art environment for statistical computing, which includes a variety of packages for machine learning. KNIME is a well-known, freely available knowledge discovery software suite. It runs on Macs, Linux, and Windows, and supports many of the methods we will use in the course. We will use these software tools to explore machine methods you have already learned about and then we will focus on new methods,

including naturally-inspired computational approaches, natural language processing, and temporal data mining, as well as evaluation methods and ethical considerations.

Learning objectives for the course:

1. Demonstrate familiarity with the literature on advanced data mining methods
2. Present and discuss the application of advanced methods for mining biomedical data
3. Perform analyses of biomedical data using advanced data mining methods and tools

Course contents

Biomedical data: specific characteristics
Methods for dealing with missing Values
Dimensionality Reduction techniques
Ensemble classifiers: Random Forests, AdaBoost, Gradient Boosting
Natural Language Processing
Naturally Inspired algorithms: introduction, genetic algorithms, evolution-based machine learning, optimization
Electronic Phenotyping
Evaluation of Machine Learning methods

Teaching methods

The course is structured with a series of lectures and several workshops, during which the instructors show the application of the presented methodologies presented to real case studies, using the R and KNIME softwares.

Recommeneded or required readings

Slides, recorded lectures, and references available on the Kiro page of the course

Assessment methods

Journal club (25% of the final grade)
The "journal club" is commonly used to engage in a presentation and lively discussion of a given article from the scientific literature. A journal club will be held at the beginning of each lab, for 30 minutes. Here is the procedure:

- We will select one paper for each lab, related to the topics of the lab.
- One (or more) students will be called upon to present the article's methods and results for 15 minutes.
- Two students will be asked to read the paper as well and ask specific questions for 15 minutes. Also the other students can ask questions if they want to.

Final Project (75% of the final grade)
The final project comprises two deliverables- a paper and an in-class presentation. The goal of the project is to demonstrate competency in identifying a public-use biomedical or public health dataset, posing a general research question (NOT a hypothesis), proposing and defending an analytic approach to mining the data in order to address the research question, applying the method, evaluating the results, and proposing new directions for further investigation.
The project can be done by yourself, or working with another student or in a small team. Each student (or team) will present the final project on the last days of class. Papers will be submitted to the instructors at the

	end of the course.
Further information	NA
Sustainable development goals - Agenda 2030	\$lbl_legenda_sviluppo_sostenibile