

SYLLABUS

Intelligent techniques for processing structured and large data

University year 2025-2026

1. Information regarding the programme

1.1. Higher education institution	Babeş-Bolyai University Cluj-Napoca
1.2. Faculty	Faculty of Mathematics and Computer Science
1.3. Department	Department of computer science
1.4. Field of study	Computer science
1.5. Study cycle	Bachelor
1.6. Study programme/Qualification	Artificial Intelligence

2. Information regarding the discipline

2.1. Name of the discipline	Intelligent techniques for processing structured and large data				Discipline code	MLE5210	
2.2. Course coordinator	Research assistant PhD Limboi Sergiu-George						
2.3. Seminar coordinator	Research assistant PhD Limboi Sergiu-George						
2.4. Year of study	3	2.5. Semester	6	2.6. Type of evaluation	E	2.7. Discipline regime	Compulsory

3. Total estimated time (hours/semester of didactic activities)

3.1. Hours per week	4	of which: 3.2 course	2	3.3 seminar/laboratory/project	2L
3.4. Total hours in the curriculum	56	of which: 3.5 course	28	3.6 seminar/laboratory/project	28
Time allotment for individual study (ID) and self-study activities (SA)					hours
Learning using manual, course support, bibliography, course notes (SA)					35
Additional documentation (in libraries, on electronic platforms, field documentation)					35
Preparation for seminars/labs, homework, papers, portfolios and essays					40
Tutorship					4
Evaluations					5
Other activities:					
3.7. Total individual study hours					119
3.8. Total hours per semester					175
3.9. Number of ECTS credits					7

4. Prerequisites (if necessary)

4.1. curriculum	Algorithms, data structures, statistics, Machine learning
4.2. competencies	Programming in Python Knowledge of using Machine learning techniques for complex problems

5. Conditions (if necessary)

5.1. for the course	Videoprojector
5.2. for the seminar /lab activities	Computers

6.1. Specific competencies acquired ¹

¹ One can choose either competences or learning outcomes, or both. If only one option is chosen, the row related to the other option will be deleted, and the kept one will be numbered 6.

Professional/essential competencies	<ul style="list-style-type: none"> • programming in high-level languages • using artificial intelligence concepts and techniques to solve real-world problems
Transversal competencies	<ul style="list-style-type: none"> • applying principles of organized and efficient work, adopting responsible attitudes toward the academic and scientific field, in order to creatively leverage one's own potential, while respecting the principles and norms of professional ethics

6.2. Learning outcomes

Knowledge	<p>The student knows how to:</p> <ul style="list-style-type: none"> • collect, analyze, and visualize data from diverse domains (financial, biological, medical) • understand how to define models for large and structured data • apply data mining techniques • evaluate the quality of the algorithms used
Skills	<p>The student is able to:</p> <ul style="list-style-type: none"> • carry out data exploitation, discovery, and the innovation of analysis methods through interdisciplinary investigations • perform data preparation and data analysis
Responsibility and autonomy:	<p>The student has the ability to work independently for:</p> <ul style="list-style-type: none"> • collaboration within a team project • data collection • applying different data analysis methodologies

7. Objectives of the discipline (outcome of the acquired competencies)

7.1 General objective of the discipline	<ul style="list-style-type: none"> • learning data mining concepts for processing large and unstructured data
7.2 Specific objective of the discipline	<ul style="list-style-type: none"> • The student will learn various data analysis techniques and will apply these techniques to solve problems using specialized software systems and tools.

8. Content

8.1 Course	Teaching methods	Remarks
1. Introduction into data mining, data analysis and knowledge discovery.	Explanations Case study discussions	

2. Representing large & structured data	Practical examples	
3. Machine learning algorithms I		
4. Machine learning algorithms II		
5. Large-scale data mining		
6. Data stream mining		
7. Anomaly & rare pattern detection		
8. Multi-view & heterogeneous data		
9. Recommender systems		
10. Text mining		
11. Ethics in AI		
12. Invited speaker. Exam review.		

Bibliography

1. S. Chakrabarti et al, Data Mining. Know It All, Morgan Kaufmann, 2009
2. Han, Jiawei, Jian Pei, and Hanghang Tong. *Data mining: concepts and techniques*. Morgan kaufmann, 2022.
3. D. Larose, Discovering Knowledge in Data. An Introduction to Data Mining, John Wiley & Sons, 2005.
4. P. Tan, M. Steinbach, V. Kumar, Introduction to Data Mining, Addison Wesley, 2006.
5. Zong, Chengqing, Rui Xia, and Jiajun Zhang. *Text data mining*. Vol. 711. Singapore: Springer, 2021.
6. Leskovec, Jure, Anand Rajaraman, and Jeffrey David Ullman. *Mining of massive data sets*. Cambridge university press, 2020.
7. Hassani, Hossein, et al. "Text mining in big data analytics." *Big Data and Cognitive Computing* 4.1 (2020): 1.
8. Jassim, Mustafa Abdalrassual, and Sarah N. Abdulwahid. "Data mining preparation: process, techniques and major issues in data analysis." *IOP conference series: materials science and engineering*. Vol. 1090. No. 1. IOP Publishing, 2021.

8.2 Laboratory	Teaching methods	Remarks
1. Data collection and pre-processing	Explanations Case study discussions Practical examples	
2. Data vizualization		
3. Processing large datasets		
4. Machine learning algorithms		
5-10. Work on the project.		
11-12. Project presentations		

Bibliography

1. S. Chakrabarti et al, Data Mining. Know It All, Morgan Kaufmann, 2009.
2. J. Han, M. Kamber, Data Mining: Concepts and Techniques, 2nd Edition, Morgan Kaufmann, 2006.
3. Hector Cuesta, Dr. Sampath Kumar, Practical Data Analysis.
4. D. Larose, Discovering Knowledge in Data. An Introduction to Data Mining, John Wiley & Sons, 2005

9. Corroborating the content of the discipline with the expectations of the epistemic community, professional associations and representative employers within the field of the program

- This course is included in the study programs of all major universities in Romania and abroad.
- The content of this course is considered important by IT companies.

10. Evaluation

Activity type	10.1 Evaluation criteria	10.2 Evaluation methods	10.3 Percentage of final grade
10.4 Course	Knowledge of concepts and	Written exam	40%

	methods in the field of data mining		
10.5 Seminar/laboratory	Apply data analysis techniques to real-world problems	Work laboratory and project implementation	60%
10.6	Minimum standard of performance		
	<ul style="list-style-type: none"> • Each student must obtain a minimum grade of 5 for the written exam, the project, and the final grade. • To obtain the minimum grade of 5, the student must demonstrate mastery of the data mining techniques for large and structured data 		

11. Labels ODD (Sustainable Development Goals)²

Not applicable.

Date:
14.02.20265

Signature of course coordinator
PhD Limboi Sergiu-George

Signature of seminar coordinator
PhD Limboi Sergiu-George

Date of approval:
...

Signature of the head of department
Assoc.prof.phd. Adrian STERCA

² Keep only the labels that, according to the [Procedure for applying ODD labels in the academic process](#), suit the discipline and delete the others, including the general one for *Sustainable Development* – if not applicable. If no label describes the discipline, delete them all and write „*Not applicable.*”.