COURSE DESCRIPTION

SPECIAL TOPICS IN DATA MINING

Academic year 2026/2027

1. Programme-related data

1.1. Higher Education Institution	Babeş-Bolyai University, Cluj-Napoca
1.2. Faculty	Faculty of Mathematics and Computer Science
1.3. Doctoral Shool	Mathematics and Computer Science
1.4. Field of study	Computer Science
1.5. Level of study	PhD

2. Course-related data

2.1. Course title	**Special topics in Data Mining**			Course code	**MDR8161**
2.2. Course coordinator	Prof. Anca Andreica				
2.3. Seminar coordinator	Prof. Anca Andreica				
2.4. Year of study	I	2.5. Semester	I	2.6. Type of assessment	Exam
2.7. Course status	Optional			2.8. Course type	Specialisation subject

3. Total estimated time (hours per semester of teaching activities)

3.1. Number of hours per week	3	of which: 3.2. course	2	3.3. seminar/ laboratory/ project	1
3.4. Total of hours in the curriculum	42	of which: 3.5. course	28	3.6. seminar/ laboratory	14
Time allocation for individual study (IS) and self-taught activities (ST)					**hours**
Learning from textbooks, course materials, bibliography, and notes (IS)					100
Additional research in the library, on subject-specific electronic platforms, and on-site					88
Preparing seminars/ laboratories/ projects, assignments, reports, portfolios, and essays					20
Tutoring (professional guidance)					
Examinations					
Other activities					
3.7. Total hours of individual study (IS) and self-taught activities (ST)				**208**	
3.8. Total hours per semester				**250**	
3.9. Number of credits				**10**	

4. Prerequisites (where applicable)

4.1. curriculum-related	-
4.2 skills-related	Programming skills in a high-level programming language

5. Specific conditions (where applicable)

5.1. course-related	-
5.2. seminar/laboratory-related	-

6. Subject-specific learning outcomes

Knowledge

1. Explains the fundamental concepts of data mining: data types, preprocessing, dimensionality reduction, and model evaluation.
2. Describes the main algorithms for classification, regression, clustering, and association rule mining.
3. Understands validation methods and performance metrics (accuracy, precision, recall, F1, AUC, etc.).
4. Explains the principles of feature selection and dimensionality reduction (PCA, filter-based methods, and learning-based methods).
5. Describes anomaly detection methods and time series analysis techniques.
6. Understands the ethical implications and limitations of data mining models (bias, overfitting, interpretability).
7. Knows the stages of a complete data analysis process (from data collection to interpretation and reporting).
Skills
1. Prepares and cleans real-world datasets for analysis.
2. Selects and applies appropriate algorithms to solve a given problem.
3. Implements data mining models using specific environments and libraries (e.g., Python/R).
4. Evaluates and compares the performance of multiple models and justifies the final choice.
5. Interprets the obtained results and draws relevant conclusions for the analyzed problem.
6. Identifies risks of overfitting and applies appropriate validation techniques.
7. Develops a structured technical report on the conducted analysis.
Responsibility and autonomy
1. Independently approaches the solution of a data analysis problem, from defining the objective to validating the solution.
2. Makes well-argued decisions regarding the selection of methods and model parameters.
3. Adheres to ethical principles concerning data use and protection.
4. Assumes responsibility for the correctness of the analysis and the interpretation of results.
5. Collaborates effectively within a team to carry out a data mining project.
6. Identifies the need for deeper study and continuous updating of knowledge in the field.

7. Contents

7.1. Course	**Teaching and learning methods**	**Remarks[1]**
1. Introduction	Presentation, conversation, debate, problematization, discovery	
2–3. Description of concepts, definitions		
4–5. Data preprocessing		
6–7. Association rules		
8–10. Classification and prediction		
11–12. Clustering		
13–14. Standards and software tools for Data Mining		
Bibliography		

[1] For example, organisational aspects, recommendations for students, specific aspects relating to the course/seminar, such as inviting experts in the field, etc.

Jiawei Han, Jian Pei, Hanghang Tong – Data Mining: Concepts and Techniques, 4th Edition, Morgan Kaufmann, 2022

Pang-Ning Tan, Anuj Karpatne, Michael Steinbach, Vipin Kumar - Introduction to Data Mining, 2nd edition, Pearson, 2019

Chengqing Zong , Rui Xia , Jiajun Zhang - Text Data Mining, Springer, 2021

International Journal of Data Warehousing and Mining, IGI Global

7.2. Seminar/ laboratory	**Teaching and learning methods**	**Remarks**
1. Data preprocessing	Providing examples	
2-3. Software tools for Data Mining		
4-7. Data Mining applications		

Bibliography

Jiawei Han, Jian Pei, Hanghang Tong – Data Mining: Concepts and Techniques, 4th Edition, Morgan Kaufmann, 2022

Pang-Ning Tan, Anuj Karpatne, Michael Steinbach, Vipin Kumar - Introduction to Data Mining, 2nd edition, Pearson, 2019

Chengqing Zong , Rui Xia , Jiajun Zhang - Text Data Mining, Springer, 2021

International Journal of Data Warehousing and Mining, IGI Global

8. Evaluation

Type of activity	8.1 Evaluation criteria[2]	8.2 Evaluation methods[3]	8.3 Percentage in the final grade
8.4. Course	The student's ability to explain the fundamental concepts and main algorithms in data mining. Understanding of validation methods, performance metrics, and dimensionality reduction techniques. Correct identification of the limitations and risks associated with the models.	Report presentation	50%
8.5. Seminar/ laboratory	Preparation, cleaning, and preprocessing of datasets. Selection and application of appropriate algorithms, model implementation, and performance evaluation. Ability to interpret results, justify chosen approaches, and write a structured report. Adherence to ethical principles and taking responsibility for the analysis.	Presentation of report experiments	50%
8.6 Minimum standard for passing			

[2] The evaluation criteria must directly reflect the learning outcomes targeted at the level of the degree programme respectively at the level of the subject. More specifically, the learning outcomes set out in the expected learning outcomes are assessed.

[3] Both final evaluation methods and ongoing evaluation strategies should be established.

Explains the essential stages of a data mining process; Knows and can describe the operation of simple algorithms; Applies these algorithms to small datasets using basic software tools; A passing grade requires achieving at least a “Satisfactory” mark.

9. SDG labels (Sustainable Development Goals)[4]

	○	Sustainable Development Generic Label

								
○	○	○	○	○	○	○	○	X
								No label applies
○	○	○	○	○	○	○	○	○

Date of entry:
13/02/2026

Signature of course coordinator

Signature of seminar coordinator

Date of approval in the department:
...

Signature of the head of department

.....

[4] Select a single label which, according to the *Implementation of SDG labels in the academic process*, best matches the subject. If the subject addresses sustainable development in a generic manner (i.e. by presenting/introducing the general framework of sustainable development, etc.), then the Sustainable Development generic label may be applied. If none of the labels describe the subject, select the last option: *"No label applies."*