SYLLABUS
Data Mining and Knowledge Discovery
2025-2026

1. Information regarding the programme

1.1 Higher education institution	**Babeş Bolyai University**
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	**Master**
1.6 Study programme / Qualification	**Data science for industry and society**

2. Information regarding the discipline

2.1 Name of the discipline (en) (ro)			**Data Mining and Knowledge Discovery** **Data Mining și descoperirea cunoștințelor**				
2.2 Course coordinator			**Lect. Dr. Ioan-Coroiu Adriana Mihaela**				
2.3 Seminar coordinator			**Lect. Dr. Ioan-Coroiu Adriana Mihaela**				
2.4. Year of study	**1**	2.5 Semeste	**2**	2.6. Type of evaluation	**E**	2.7 Type of discipline	**Optional**
2.8 Code of the discipline		**MME8183**					

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	1 lab+1 project
3.4 Total hours in the curriculum	56	Of which: 3.5 course	28	3.6 seminar/laboratory	28

Time allotment:	hours
Learning using manual, course support, bibliography, course notes	36
Additional documentation (in libraries, on electronic platforms, field documentation)	36
Preparation for seminars/labs, homework, papers, portfolios and essays	36
Tutorship	3
Evaluations	8
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
4.2. competencies	• Average programming skills

5. Conditions (if necessary)

5.1. for the course	• Videoprojector
5.2. for the seminar /lab activities	• Computers, specific development environment

6. 6.1. Specific competencies acquired

Professional competencies	**C5.3** Use of databases methodologies and design environments for particular problems **C5.4** Quality evaluation of different database management systems in terms of structure, functionality and extensibility **C5.5** Implementation of database projects
Transversal competencies	**CT1.** Application of efficient work rules and responsible attitudes towards the scientific domain, for the creative exploitation of one's own potential according to the principles and rules of professional ethics **CT2.** Efficient conduct of activities organized in an interdisciplinary group and development of empathic capacity of interpersonal communication, networking and collaboration with diverse groups **CT3.** Use of efficient methods and techniques for learning, information, research and development of abilities for knowledge exploitation, for adapting to the needs of a dynamic society and for communication in a widely used foreign language.

6.2. Learning outcomes

Knowledge	The student knows: • Collection, representation, analysis and visualization of data from various fields (eg economics, finance, biology, natural sciences) • Classification of structured and unstructured data and information • Evaluation of data quality
Skills	The student is able to: • Data management and preparation for analysis by sampling, normalizing, cleaning and extracting features / information from raw data • Management of data science projects involving real-world information • Exploitation, discovery and innovation of analysis methods through interdisciplinary investigations
Responsibility and autonomy:	The student has the ability to work independently to obtain: • Applying scientific and analytical methods of data analysis to solve problems in different fields, developing creative thinking about data and making decisions • Collaboration within the team • Providing reproducible data analysis

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

7.1 General objective of the discipline	• To learn data mining and knowledge discovery concepts, methods and techniques
7.2 Specific objective of the discipline	• The students will learn various data analysis techniques and will apply these techniques for solving data mining problems using special software systems and tools.

8. Content

8.1 Course	Teaching methods	Remarks
1. Introduction	• Interactive exposure • Presentation • Explanation • Practical examples • Case-study discussions	
2. Concept description; Definitions		
3. Data Preparation		
4. Discovering, Ingesting, and Exploring Data		
5. Transforming Data into Analytics-Ready Data		
6. Data understanding		
7. Visualising data for explaratory analysis		
8-9. Unsupervised models for data visualisation		
10-12 Model Assessment and Validation		
12-14. Student presentations		

Bibliography

1. S. Chakrabarti et al, Data Mining. Know It All, Morgan Kaufmann, 2009.
2. K. Cios, W. Pedrycz, R. Swiniarski, L. Kurgan, Data Mining. A Knowledge Discovery Approach, Springer, 2007.
3. J. Han, M. Kamber, Data Mining: Concepts and Techniques, 2nd Edition, Morgan Kaufmann, 2006.
4. P. Tan, M. Steinbach, V. Kumar, Introduction to Data Mining, Addison Wesley, 2006.
5. D. Larose, Discovering Knowledge in Data. An Introduction to Data Mining, John Wiley & Sons, 2005.
6. Han, J., Kamber, M., Data Mining: Concepts and Techniques, 1st Edition, Morgan Kaufmann, 2000.
Weka system and documentation (http://www.cs.waikato.ac.nz/ml/weka/). Weka is a suite of machine learning / data mining software. It contains Java implementation for various mining algorithms, data preprocessing filters, and experimentation capabilities. Weka is free open-source software under the GNU General Public License (GPL).

8.2 Seminar / laboratory	Teaching methods	Remarks
1-2. Data preprocessing	• Interactive exposure • Explanation • Conversation • Didactical demonstration	
3. Association Rules		
4. Clustering		
5-6. Classification		
7. Students project presentations		

Bibliography

1. S. Chakrabarti et al, Data Mining. Know It All, Morgan Kaufmann, 2009.
2. K. Cios, W. Pedrycz, R. Swiniarski, L. Kurgan, Data Mining. A Knowledge Discovery Approach, Springer, 2007.
3. J. Han, M. Kamber, Data Mining: Concepts and Techniques, 2nd Edition, Morgan Kaufmann, 2006.
4. P. Tan, M. Steinbach, V. Kumar, Introduction to Data Mining, Addison Wesley, 2006.
5. D. Larose, Discovering Knowledge in Data. An Introduction to Data Mining, John Wiley & Sons, 2005.
6. Han, J., Kamber, M., Data Mining: Concepts and Techniques, 1st Edition, Morgan Kaufmann, 2000.
7. Weka system and documentation (http://www.cs.waikato.ac.nz/ml/weka/). Weka is a suite of machine learning / data mining software. It contains Java implementation for various mining algorithms, data preprocessing filters, and experimentation capabilities. Weka is free open-source software under the GNU General Public License (GPL).

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

- The course exists in the curriculum of many universities in the world.
- The results of course are considered by software companies particularly useful and topical,

developing needed abilities in modelling and visualization of data.

10. Evaluation

Type of activity	10.1 Evaluation criteria	10.2 Evaluation methods	10.3 Share in the grade (%)
10.4 Course	Know concepts and methods from the domain of data mining and knowledge discovery	Report presentation	40%
		Written exam	10%
10.5 Seminar/lab activities	Apply data mining techniques in real problems	Projects implementation and presentation	50%
10.6 Minimum performance standards			
Each student should obtain minimum 5 for the final grade. In order to obtain the minimum grade 5, the student must demonstrate the mastery of the basic concepts of data preparation in order to analyze them.			

Date
10.06.2024

Signature of course coordinator

Lect. Dr. Ioan-Coroiu Adriana Mihaela

Signature of seminar coordinator

Lect. Dr. Ioan-Coroiu Adriana Mihaela

Date of approval

.....

Signature of the head of department

Assoc. prof. dr. Sterca Adrian