SYLLABUS

Data mining in biomedicine

Academic year 2025-2026

1. Information regarding the programme

1.1. Higher education institution	Babeș-Bolyai University
1.2. Faculty	Faculty of Biology and Geology
1.3. Department	Department of Molecular Biology and Biotechnology
1.4. Field of study	Biology
1.5. Study cycle	Master, 4 semesters
1.6. Study programme/Qualification	Bioinformatics Applied in Life Sciences (English)/ Biologist
1.7. Form of education	Full-time

2. Information regarding the discipline

2.1. Name of the discipline	**Data mining in biomedicine**					Discipline code	**BME 1136**
2.2. Course coordinator				Andrei Tiberiu Alexsson, MSc Ovidiu Ilie Pavel, MSc			
2.3. Seminar coordinator				Andrei Tiberiu Alexsson, MSc Ovidiu Ilie Pavel, MSc			
2.4. Year of study	2	2.5. Semester	3	2.6. Type of evaluation	E	2.7. Discipline regime	Optional

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

3.1. Hours per week	**4**	of which: 3.2 course	**4**	3.3 seminar/laboratory	**4**
3.4. Total hours in the curriculum	56	of which: 3.5 course	28	3.6 seminar/laborator	**28**
Time allotment for individual study (ID) and self-study activities (SA)					**hours**
Learning using manual, course support, bibliography, course notes (SA)					20
Additional documentation (in libraries, on electronic platforms, field documentation)					20
Preparation for seminars/labs, homework, papers, portfolios and essays					16
Tutorship					8
Evaluations					4
Other activities: two-way communication with the course holder / tutor					2
3.7. Total individual study hours	70				
3.8. Total hours per semester	126				
3.9. Number of ECTS credits	5				

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	Online meeting platform Beamer, projection screen
5.2. for the seminar / lab activities	Attendance of a minimum of 90% of seminar classes is mandatory for granting the participation at the written exam. Computers, specific environment for developing and implementing bioinformatic pipelines/tools

6. Specific competencies acquired [1]

Professional competencies	• Use of databases methodologies and design environments for particular problems • Quality evaluation of different database management systems in terms of structure, functionality and extensibility • Implementation of database projects
Transversal competencies	• 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 • Efficient conduct of activities organized in an interdisciplinary group and development of empathic capacity of interpersonal communication, networking and collaboration with diverse groups • 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 English.

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
Types of biomedical data and specific databases	• Presentation • Explanation • Conceptualized-lerning • Active learning through case-study discussions	
Applications of data minining in biomedicine		
Data collection and preprocessing methods		
Data normalization and transformation methods		
Data analysis methods		
Quality control methods		
Genome-Wide Association Studies (GWAS) concepts		

Bibliography

Agatonovic-Kustrin, S., & Morton, D. (2016). Data mining in drug discovery and design. In Artificial neural network for drug design, delivery and disposition (pp. 181-193). Academic Press.
Jothi, N., Rashid, N. A. A., & Husain, W. (2015). Data mining in healthcare–a review. Procedia computer science, 72, 306-313.
Kolling, M. L., Furstenau, L. B., Sott, M. K., Rabaioli, B., Ulmi, P. H., Bragazzi, N. L., & Tedesco, L. P. C. (2021). Data mining in healthcare: Applying strategic intelligence techniques to depict 25 years of research development. International Journal of Environmental Research and Public Health, 18(6), 3099.
Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. Computational and structural biotechnology journal, 13, 8-17.
Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., ... & Moore, H. F. (2013). The genotype-tissue expression (GTEx) project. Nature genetics, 45(6), 580-585..
Masoodi, F., Quasim, M., Bukhari, S., Dixit, S., & Alam, S. (Eds.). (2023). Applications of Machine Learning and Deep Learning on Biological Data (1st ed.). Auerbach Publications. https://doi.org/10.1201/9781003328780
Sadee, W., Wang, D., Hartmann, K., & Toland, A. E. (2023). Pharmacogenomics: driving personalized medicine. Pharmacological reviews, 75(4), 789-814.

[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.

Tomczak, K., Czerwińska, P., & Wiznerowicz, M. (2015). Review The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemporary Oncology, 2015(1), 68-77, The references are accessible online through the branch libraries of the Central University Library of Cluj-Napoca.		
8.2 Seminar / laboratory	Teaching methods	Remarks
Searching for published studies in public databases	• Interactive discussion • Explanation • Conversation • Project-based learning • Practical demonstration	
Retrieving data using APIs		
Data visualization		
Exploratory data analysis		
Data visualization and interpretation through tables and plots		
GWAS data processing and analysis		
Final evaluation of student's project		
Bibliography Entrez Programming Utilities Help, Bethesda (MD): National Center for Biotechnology Information (US); 2010-. (https://www.ncbi.nlm.nih.gov/books/NBK25501/) A selection of updated scientific papers will be provided by the trainers. The references are accessible online through the branch libraries of the Central University Library of Cluj-Napoca.		

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 and data mining 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	Knowledge of concepts and methods from the domain of data mining and knowledge discovery	Oral exam	50%
10.5. Seminar/lab activities	Apply data mining techniques in real problems	Project implementation and presentation	50%
10.6. Minimum performance standards			
Each student should obtain a minimal grade of 5 (five) for the oral exam and for the final grade. To obtain the minimum grade 5, the student must demonstrate the mastery of the basic concepts described during the course and seminars/practical classes, respectively.			

11. Labels ODD (Sustainable Development Goals)[2]

	SDG 3. Health and well-being

Date:
08.01.2025

Signature of course coordinator

Andrei Tiberiu Alexsson

Ovidiu Ilie Pavel

Signature of seminar coordinator

Andrei Tiberiu Alexsson

Ovidiu Ilie Pavel

Date of approval:
...

Signature of the head of department

Assoc. Prof. Beatrice Kelemen, PhD

[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.*".