MD-2068, CHISINAU, 9/7 STUDENTILOR STR, PHONE: 022 50-99-63, www.utm.md

Data Science. Semester Project

1. Course/module information

Facultaty	Computers, Informatics, and Microelectronics				
Chiar/department	Computer Science and Systems Engineering				
Cycle of studies	Bachelor's degree, cycle II				
Study program	**Data Science**				
Year of study	**Semester**	**Type of evaluation**	**Training category**	**Optionality category**	**ECTS credits**
II (full-time education)	3	E	S – specialized course unit	O - Obligatory course unit	5

2. Total estimated time

Total hours in the curriculum	Including				
	Auditory hours		Individual work		
	Lecture	Laboratory/seminar	Term paper	Study of theoretical material	Application development
300		40	80	220	

3. Prerequisites for access to the course/module

According to the curriculum plan	To achieve the module objectives, students will use skills acquired in related courses of the semester: Applied Sciences, Mathematics, Mathematical Analysis 1, 2, Data Structures and Algorithms, Object-Oriented Programming, C++ Programming, Python, Operating Systems.
According to competencies	The use of domain-specific theories and tools (algorithms, methods, techniques, etc.) for analyzing basic algorithms in statistics, mathematical analysis, and data visualization through graphs.

4. Conditions for conducting the educational process

Lecture	A blackboard, projector, and computer are required to present the theoretical material in the classroom.
Laboratory/seminar	Students will perfect reports according to the conditions set by the methodological guidelines. The deadline for submitting the laboratory work is 2 weeks after its completion. Late submission will result in a 1-point deduction per week of delay. A blackboard, projector, and computer are required to conduct seminars in the study auditorium.

5. Specific competencies acquired

Professional competencies	**CPM 1.** System architecture design and development K1 Architecture models, methodologies, and system design tools K2 System architecture requirements: performance, maintainability, extensibility, scalability, availability, security, and accessibility K3 Costs, benefits, and risks of a system architecture K4 Enterprise architecture and the company's internal standards K5 Emerging new technologies (e.g., distributed systems, virtualization models, data sets, mobile systems) S1 Provide expertise to help solve complex technical problems and ensure the implementation of the best architectural solutions S2 Utilizes technological knowledge from different fields to develop and implement enterprise architecture

	S3 Understands the company's objectives that impact architecture components (data, applications, security, development, etc.) S4 Helps communicate the enterprise architecture, along with standards, principles, and objectives, to the various teams involved S5 Develops design models and architectural patterns to assist system analysts in designing coherent applications **CPM 2.** Monitoring technological trends. Innovation. Sustainable development K1 Existing and emerging technologies and their relevant applications in the market K2 Business, society, and research objectives, trends, and needs K3 Relevant information sources (e.g., journals, conferences and events, opinion leaders, online forums, etc.) K4 Concrete approaches to applied research programs K5 Innovation process techniques K6 Sustainable development criteria and indicators K7 Corporate social responsibility (CSR) of stakeholders within the information system infrastructure S1 Monitors information sources and continuously tracks the most promising ones S2 Identifies vendors and suppliers of the most promising solutions; evaluates, justifies, and proposes the most suitable ones S3 Identifies the advantages and improvements brought by adopting emerging technologies S4 Thinks without preconceived ideas S5 Applies recommendations in projects that support the latest sustainable development strategies **CPM 3.** Application development. Component integration. Systems engineering. K1 Appropriate programs/modules, DBMS, and suitable programming languages. Cutting-edge technologies. K3 The impact of system integration on the organization or existing system. K4 Interface techniques between modules, systems, and components. K5 Integration testing techniques. K7 Hardware components, tools, and hardware architectures. K8 Functional and technical design. K9 Information security fundamentals. K10 Prototyping. S1 Applies suitable software and/or hardware architectures. S2 Measures system performance before, during, and after system integration. S3 Identifies and records activities, issues, and corrective actions related to maintenance. S4 Adapts client needs to existing products. S5 Secures and backs up data to ensure its integrity during data or system integration. S6 Explains and communicates with the client regarding design/development. S7 Launches and evaluates test results according to product specifications. S9 Applies data models and processes to develop efficiently and productively. **CPM 4.** Staff development. K1 Methods for developing competencies K2 Methodologies for analyzing competency needs K3 Methods for supporting learning and development (e.g., coaching, teaching) K4 Relevant information technologies and processes K5 Techniques for empowerment and autonomy S1 Identifies competency and qualification gaps S2 Identifies and recommends development opportunities based on work activities S3 Incorporates competency development opportunities into routine work processes

	CPM 5. Process improvement. K1 Research, comparison, and measurement methods K2 Evaluation, design, and implementation methods K3 Existing internal processes K4 Relevant developments/evolutions in the ICT field (e.g., virtualization, open data, etc.) and their potential impact on processes K5 Specificity of web, cloud, and mobile technologies S1 Drafts, documents, and catalogs essential processes and procedures S2 Proposes process modifications to facilitate and streamline improvements
Transversal competencies	**CTM1.** Autonomy and responsibility **CTM2.** Social interaction **CTM3.** Personal and professional development

6. Course/Module objectives

General objective	• Knowledge and application of advanced techniques in data science for developing software solutions and innovative technologies. • Optimizing data analysis and processing processes using machine learning methods, artificial intelligence algorithms, and Big Data technologies. • Understanding information systems and distributed computing platforms used for managing and processing large data sets. • Developing practical projects based on real data, providing students with hands-on experience in solving data analysis and processing problems applicable across various industries. • Developing skills in interpreting and communicating analytical results, ensuring that proposed solutions are clear and practically implementable.
Specific objectives	• Developing practical projects based on real data, providing students with hands-on experience in solving data analysis and processing problems, applicable across various industries. • Developing skills in interpreting and communicating analytical results, ensuring that proposed solutions are clear and practically implementable.

7. Course/Module content

Syllabus of teaching activities	Number of hours Full-time education
Course topics	
Topic 1: Introduction to Data Science	4
Topic 2: Data Collection and Preprocessing	4
Topic 3: Exploratory Data Analysis (EDA)	4
Topic 4: Machine Learning	4
Topic 5: Model Performance Evaluation	4
Topic 6: Feature Selection	4
Topic 7: Model Validation and Generalization	4
Topic 8: Unsupervised Learning	4
Topic 9: Deep Learning	4
Topic 10: Model Optimization and Scaling	4
Total lectures:	**40**

8. Using generative AI

Permission to use	The use of generative AI in assignments and projects is permitted, provided that students adhere to the following rules: • Generative AI may be used to generate ideas, text structures, or code, but all generated materials must be reviewed and adjusted by the student to ensure that they meet academic requirements. • Any use of generative AI must be declared in the appendix section of each paper, using the phrase: "During the preparation of this paper, the author used [NAME OF TOOL / SERVICE] for the purpose of [REASON]. After using this tool / service, the author reviewed and edited the content as necessary and assumes full responsibility for the content of the paper."
Restrictions to use	Students ***MUSTN'T consider generative AI as a reliable source of information***, as it does not provide clear references or documented sources. • ***Direct citation of AI-generated content*** in academic papers as if it were a primary source ***isn't permitted.*** • Activities in which the use of **generative AI is prohibited** are specified by the teacher and are usually ***intermediate and final assessments*** or that don't involve professional competence development activities.

9. Bibliographic references

Main	1. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer. **A fundamental paper for understanding machine learning methods and their application in data analysis.** 2. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning with Applications in R*. Springer. **A practical guide to machine learning methods, including detailed examples using the R programming language.** 3. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. **An essential resource for deep learning, covering neural networks and advanced applications in detail.** 4. Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). *Scikit-learn: Machine Learning in Python*. Journal of Machine Learning Research. **An important reference for using Scikit-learn in machine learning projects with Python.** 5. McKinney, W. (2017). *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. O'Reilly Media. **A practical guide to data manipulation and analysis using Pandas and NumPy.** 6. Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. MIT Press. **A detailed introduction to model-based probabilistic machine learning, providing a solid mathematical approach.**
Supplementary	1. **Provost, F., & Fawcett, T. (2013).** *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*. O'Reilly Media. **A useful resource for understanding the impact and applicability of data science in a business context.** 2. **Wickham, H., & Grolemund, G. (2016).** *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media. **A very useful resource for R users who want to learn modern data analysis techniques using the Tidyverse package.**

10. Evaluation

Form of study	Periodic		Current	Project/thesis	Final exam
	Mid-term 1	Mid-term 2			
Full-time	10%	10%	10%	30%	40%
	Minimum performance standard				
	Student evaluation in the **Data Science Project** is based on a balanced combination of continuous activity, attendance at consultations, and the final project. Defining a real and relevant problem in data science is a crucial first step in student evaluation. Students must demonstrate the ability to formulate concrete solutions using specific techniques in analysis, design, and software development. Attendance and participation in consultations related to the project are constantly monitored, providing continuous feedback to students or the working team, allowing them to improve progress in the project. Ongoing assessment is formative, ensuring a continuous learning process and adjustment of solutions. The final exam is a summative evaluation, conducted orally, based on the project presented publicly by the team. During this presentation, each student will be individually evaluated, both for their contribution to the project and through individual discussions in the presence of the team. The grades obtained during the final exam are individual and constitute 40% of the final grade, reflecting both technical competencies and students' communication and argumentation skills.				