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

LARGE LANGUAGE MODELS

1. Course/Module information

Faculty	Computers, Informatics, and Microelectronics				
Chair/department	Informatics and Systems Engineering				
Study cycle	Cycle II, Master's Studies				
Study program	**Data Science**				
Year of study	**Semester**	**Evaluation type**	**Formative category**	**Optionality category**	**ECTS credits**
I (full-time education)	2	E	F – fundamental course unit	O - mandatory course unit	5

2. Estimated total time

Total hours in the curriculum	Including				
	Auditory hours		Individual work		
	Lecture	Laboratory/ seminar	Term paper	Study of theoretical material	Application development
150	20	20		60	50

3. Prerequisites for access to the course/module

According to the curriculum plan	Linear Algebra, Advanced Mathematics, Probability and Information Theory, Mathematical Statistics, Mathematical Models and Optimization, Exploratory Data Analysis and Data Modeling, Data Visualization, Machine Learning, and Data Mining.
According to competencies	Skills in Data Manipulation and Preprocessing: Knowledge of techniques for data cleaning, normalization, feature scaling, and handling missing data. Data Visualization: Ability to create and interpret visualizations to represent data effectively. Critical Thinking and Problem-Solving: Ability to approach machine learning problems efficiently. Self-Learning Skills: Ability and willingness to engage in self-directed learning and research to deepen understanding and keep up with the rapidly evolving field of artificial intelligence. Understanding of Business Context: Awareness of how AI integrates into business and real-world applications.

4. Conditions for conducting the educational process

Lecture	Theoretical material will be presented in the classroom using a projector and a computer. Teaching materials will be made available to students on the course page of the department's pedagogical server.
Laboratory/seminar	Students will complete reports according to the conditions outlined in the methodological guidelines. The deadline for submitting the laboratory work is two

	weeks after its completion. Late submission of the work will result in a penalty of 0.25 points per day of delay.

5. Specific competencies acquired

Professional competencies	**CPM 1. Development and design of architecture** 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 company internal standards K5 Emerging technologies (e.g., distributed systems, virtualization models, datasets, mobile systems) **CPM 3. Application development. Component integration. Systems engineering** K1 Appropriate programs/modules, DBMS, and suitable programming languages. Cutting-edge technologies K3 Impact of system integration on the organization or the existing system K4 Interface techniques between modules, systems, and components K5 Integration testing techniques K6 Best design practices K7 Hardware components, tools, and hardware architectures K8 Functional and technical design K9 Fundamentals of information security K10 Prototyping
Transversal competencies	***CT1.*** Honorable, responsible, and ethical behavior in accordance with the law to ensure the fulfillment of professional tasks. ***CT2.*** Demonstrating the ability to work in a team, identifying individual and shared roles and responsibilities, making decisions, and assigning tasks, applying relationship and effective teamwork techniques. ***CT3.*** Demonstrating initiative and action for professional and personal development through continuous training using sources of documentation in Romanian and international languages.

6. Course/Module objectives

General objective	Development of theoretical and practical knowledge for building and using large linguistic models in natural language processing, providing innovative solutions to current challenges.
Specific objectives	As a result of studying the course, the student will know: • Understanding the architecture of transformers and their application to different NLP tasks. • Implementing and fine-tuning pre-trained models for specific analyses. • Ethical evaluation of the use of artificial intelligence in natural language processing.

7. Course/Module content

Syllabus of teaching activities	Number of hours Full-time education
Course topics	
T1. Introduction to NLP and Text Preprocessing: Tokenization, Stemming, Lemmatization.	4
T2. Word Embedding: Word2Vec, GloVe, FastText. Vector visualization.	4
T3. Neural Networks for NLP: RNN, LSTM, GRU.	4
T4. Transformers: BERT, GPT. Architecture and attention mechanisms.	6
T5. Ethics and responsible use of NLP models.	2
Total lectures:	**20**
Topics for laboratory/seminar work	
LL1. Building a classification model using RNN.	6
LL2. Fine-tuning BERT for sentiment analysis.	6
LL3. Text generation with GPT.	4
LL4. Evaluating and optimizing an NLP model.	4
Total laboratory/seminar work:	**20**

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. Vaswani, A., et al. "Attention Is All You Need." NeurIPS, 2017. 2. Jurafsky, D., & Martin, J. H. "Speech and Language Processing." Pearson, 2020. 3. Lewis, T., et al. "Natural Language Processing with Transformers." O'Reilly, 2021.
Supplementary	1. Goldberg, Y. "Neural Network Methods for Natural Language Processing." Morgan & Claypool, 2017. 2. Hugging Face Documentation: https://huggingface.co/docs

10. Evaluation

Form of education	Periodic		Current	Individual work	Exam final
	Mid-term 1	Mid-term 2			
Full-time	15%	15%	15%	15%	40%
Minimum performance standard					
Attendance and participation in lectures and laboratory work Achieving a minimum grade of "5" in each evaluation and laboratory work					

11. Evaluation criteria

Activity	Evaluation components	Evaluation method, evaluation criteria	Weight in final grade for the Activity	Weight in course evaluation
Full-time education				
Mid term I	Theoretical content, topics 1-3	Test	100%	**15%**
Mid term II	Theoretical content, topics 4-5	Activities during practical work/seminar	100%	**15%**
Current evaluation	Practical activity	Attendance and participation in classes	50%	**15%**
Individual study	Classification of research by activity type	Presentation/Discussion on the topic	100%	**15%**
Final examination	Theoretical and practical content	Oral exam. Grading according to grading scale	100%	**40%**