

FIȘA DISCIPLINEI

Înțelegerea și Implementarea de Modele Lingvistice Mari (LLMs)

Anul universitar 2026-2027

1. Date despre program

1.1. Instituția de învățământ superior	Universitatea Babeș-Bolyai
1.2. Facultatea	Facultatea de Matematica și Informatica
1.3. Departamentul	Informatica
1.4. Domeniul de studii	Informatica
1.5. Ciclul de studii	Licenta
1.6. Programul de studii / Calificarea	Informatica
1.7. Forma de învățământ	Cu frecvență

2. Date despre disciplină

2.1. Denumirea disciplinei	Înțelegerea și Implementarea de Modele Lingvistice Mari (LLMs)			Codul disciplinei	<i>MLE5247</i>
2.2. Titularul activităților de curs	Lect. Dr. Bogdan MURSA				
2.3. Titularul activităților de seminar	Lect. Dr. Bogdan MURSA				
2.4. Anul de studiu	3	2.5. Semestrul	5	2.6. Tipul de evaluare	Colocviu
2.7. Regimul disciplinei	Opțional	2.8. Tipul disciplinei		Disciplină de specializare (DS)	

3. Timpul total estimat (ore pe semestru al activităților didactice)

3.1. Număr de ore pe săptămână	5	din care: 3.2. curs	2	3.3. seminar/ laborator/ proiect	1 lab + 1 proj
3.4. Total ore din planul de învățământ	60	din care: 3.5. curs	24	3.6 seminar/laborator	36
Distribuția fondului de timp pentru studiul individual (SI) și activități de autoinstruire (AI)					ore
Studiul după manual, suport de curs, bibliografie și notițe (AI)					12
Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate și pe teren					16
Pregătire seminare/ laboratoare/ proiecte, teme, referate, portofolii și eseuri					25
Tutoriat (consiliere profesională)					6
Examinări					6
Alte activități					-
3.7. Total ore studiu individual (SI) și activități de autoinstruire (AI)				65	
3.8. Total ore pe semestru				125	
3.9. Numărul de credite				5	

4. Precondiții (acolo unde este cazul)

4.1. de curriculum	<ul style="list-style-type: none">• Programare în Python• Algebră liniară• Statistică• Structuri de date și algoritmi
4.2. de competențe	<ul style="list-style-type: none">• Competențe medii de programare într-un limbaj de programare de nivel înalt și cunoștințe foarte bune despre structuri de date și algoritmi.

5. Condiții (acolo unde este cazul)

5.1. de desfășurare a cursului	• Sală de curs cu videoproiector	
5.2. de desfășurare a seminarului/ laboratorului	• Laborator echipat cu calculatoare performante și Python instalat.	

6.1. Competențele dobândite în urma absolvirii programului de studii

Competențe profesionale	
Codul competenței	Competență
CP1	crează softuri
CP18	utilizează învățarea automatizată
CP6	dezvoltă prototipul pentru software
Competențe transversale	
Codul competenței	Competență
CT2	Soluționează probleme
CT3	Gândește analitic

6.2. Rezultatele învățării specifice programului de studii

Rezultatele învățării vizate prin disciplină		
Codul competenței	Cunoștințe și înțelegere (Knowledge and understanding)	Abilități academice specifice (Specific academic skills)
CP1/CP6	<p>Studentul/absolventul identifică, explică și argumentează concepte fundamentale de structuri de date, algoritmi și paradigme de programare, precum și a arhitecturii calculatoarelor.</p> <p>Absolventul cunoaște, înțelege și aplică conceptele și algoritmi fundamentali utilizați în inteligența artificială și este capabil să le evalueze pe baza unor metrici.</p> <p>Absolventul cunoaște și înțelege conceptele și tehnicile de reprezentare a cunoștințelor și le poate aplica în vederea rezolvării de probleme.</p> <p>Absolventul cunoaște și înțelege fundamentele matematice necesare dezvoltării algoritmilor inteligenți și este capabil să le utilizeze pentru implementarea acestor algoritmi.</p> <p>Absolventul cunoaște, înțelege și utilizează metode pentru reprezentarea, analiza și manipularea unor volume mari de date.</p> <p>The graduate knows, understands and ap</p>	<p>Studentul/absolventul elaborează, dezvoltă și demonstrează soluții software complexe utilizând algoritmi eficienți și paradigme diverse de programare.</p> <p>Absolventul este capabil să descrie formal problemele abordate din diferite domenii, modelându-le ca probleme care se pot aborda cu tehnici din sfera inteligenței artificiale.</p> <p>Absolventul este capabil să aplice algoritmi fundamentali de inteligență artificială pentru a rezolva probleme din lumea reală. Absolventul este capabil să evalueze, atât în mod cantitativ și cât și în mod calitativ, performanța sistemelor inteligente. Absolventul este capabil să proiecteze și să implementeze sisteme software care folosesc metode ale inteligenței artificiale și să evalueze performanța acestora.</p>
CP18	<p>Studentul/absolventul identifică, compară, recunoaște și descrie concepte și tehnici avansate din domeniul inteligenței artificiale, învățării automate și procesării limbajului natural.</p>	<p>Studentul/absolventul proiectează, implementează, experimentează modele predictive și dezvoltă aplicații bazate pe algoritmi de învățare automată.</p>

CT2/CT3	Studentul/absolventul are cunoștințele necesare pentru a înțelege și soluționa probleme complexe, pentru a planifica și organiza procese avansate în diverse domenii.	Absolventul este capabil să identifice probleme complexe și să examineze probleme conexe pentru a dezvolta opțiuni de rezolvare și implementa soluții. Absolventul are abilitatea de a aplica reguli generale unor probleme specifice și de a produce soluții relevante. Absolventul este capabil să combine informații diverse pentru a formula soluții și genera idei de dezvoltare pentru noi produse și aplicații.
---------	---	--

7. Rezultatele învățării specifice disciplinei

Cunoștințe și înțelegere (Knowledge and understanding)
1. Înțelegerea conceptelor fundamentale și avansate de Inteligență Artificială și Machine Learning aplicate în contextul sistemelor conectate.
2. Înțelegerea principalelor paradigme de învățare automată: învățare supervizată, nesupervizată și deep learning.
3. Înțelegerea etapelor de proiectare, antrenare, evaluare și optimizare a modelelor de Machine Learning.
4. Înțelegerea provocărilor legate de scalabilitate, procesarea datelor și integrarea soluțiilor AI în arhitecturi Big Data și sisteme distribuite.
Abilități academice specifice (Specific academic skills)
1. Dezvoltarea și evaluarea unor modele de Machine Learning utilizând seturi reale de date și framework-uri moderne AI.
2. Analiza, preprocesarea și vizualizarea datelor în vederea construirii unor soluții inteligente robuste și eficiente.
3. Proiectarea și implementarea unor aplicații bazate pe Inteligență Artificială pentru rezolvarea unor probleme reale din domeniul sistemelor conectate.

8. Conținuturi

8.1 Curs	Metode de predare	Observații
1. Introducere în LLM-uri și peisajul inteligenței artificiale generative. Prezentare generală a istoriei procesării limbajului natural, cu accent pe modelele lingvistice de mari dimensiuni (LLM) și importanța acestora în domeniul inteligenței artificiale generative. Examinarea diverselor aplicații și sarcini în care sunt utilizate LLM-urile, evidențiindu-se versatilitatea lor.	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Demonstrație didactică 	
2. Evoluția tehnologiilor de generare a textului. Urmărirea dezvoltării generării de text, de la modelele pre-transformer până la metodologiile actuale.	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Demonstrație didactică 	
3. Explorare aprofundată a arhitecturii transformer. Tehnici și	<ul style="list-style-type: none"> • Expunere interactivă 	

strategii pentru utilizarea transformerelor în sarcini de generare de text. Explorarea arhitecturii transformer, fundamentul LLM-urilor moderne.	<ul style="list-style-type: none"> • Explicație • Conversație • Demonstrație didactică 	
4. Principiul atenției în transformare. Înțelegerea conceptului „Attention is all you need” și impactul său revoluționar asupra LLM-urilor.	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Demonstrație didactică 	
5. Stăpânirea ingineriei prompturilor. Învățarea modului eficient de proiectare a prompturilor pentru a ghida LLM-urile în generarea rezultatelor dorite.	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Demonstrație didactică 	
6. Pre-antrenarea modelelor LLM și legile de scalare. Perspective asupra procesului de pre-antrenare, provocările computaționale și principiile legilor de scalare pentru LLM-uri.	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Demonstrație didactică 	
7. Fine-tuning-ul LLM-urilor pentru sarcini specifice. Strategii pentru fine-tuning bazat pe instrucțiuni, inclusiv adaptări monotask și multitask.	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Demonstrație didactică 	
8. Tehnici avansate de fine-tuning. Introducere în metodele eficiente din punct de vedere al parametrilor (PEFT), cum ar fi LoRA și Soft Prompts.	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Demonstrație didactică 	
9. Învățare prin întărire din feedback uman (RLHF). Fundamentele alinierii LLM-urilor cu valorile umane prin RLHF, inclusiv colectarea de feedback și modelele de recompensă.	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Demonstrație didactică 	
10. Îmbunătățirea ieșirilor LLM prin raționament și acțiune. Explorarea strategiilor avansate de fine-tuning și prompting prin metode precum Chain-	<ul style="list-style-type: none"> • Expunere interactivă • Explicație 	

of-Thought (CoT, doar raționament), doar acțiuni și ReAct, evidențiind traiectoriile de rezolvare a sarcinilor și avantajele distincte ale abordării ReAct.	<ul style="list-style-type: none"> • Conversație • Demonstrație didactică 	
11. Implementarea LLM-urilor în aplicații din lumea reală și introducere în LangChain. Combinarea explorării implementării LLM-urilor în aplicații reale cu o introducere în LangChain: încărcarea documentelor, vector stores, embeddings și fundamentele generării augmentate prin regăsire (RAG).	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Demonstrație didactică 	
12. Etica în AI. Descoperirea domeniului în continuă evoluție al inteligenței artificiale generative, subliniind necesitatea utilizării responsabile și a inovației continue în aplicațiile bazate pe LLM-uri.	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Demonstrație didactică 	
13. Prezentarea proiectelor.	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Dialog, dezbateri 	
14. Prezentarea proiectelor.	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Dialog, dezbateri 	
Bibliografie 1. Chung, H. W., Hou, L., Longpre, S., Zoph, B., Tay, Y., Fedus, W., Li, Y., Wang, X., Dehghani, M., Brahma, S., Webson, A., Gu, S. S., Dai, Z., Suzgun, M., Chen, X., Chowdhery, A., Pellat, M., Robinson, K., Valter, D., . . . Wei, J. (2022). <i>Scaling Instruction-Finetuned Language Models</i> . 2. Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2022). <i>ReAct: Synergizing Reasoning and Acting in Language Models</i> . 3. Wu, S., Irsoy, O., Lu, S., Dabrovolski, V., Dredze, M., Gehrmann, S., Kambadur, P., Rosenberg, D., & Mann, G. (2023). <i>BloombergGPT: A Large Language Model for Finance</i> . 4. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). <i>Attention Is All You Need</i> 5. Alammar, J, Grootendorst, M. (2024). <i>Hands-On Large Language Models</i> . 6. Auffarth, B. (2023). <i>Generative AI</i>		
8.2 Seminar/Laborator	Metode de predare	Observații
1. Introducere în LLM-uri și generarea de text. Activitate practică cu operațiuni de bază ale LLM-urilor, cu	<ul style="list-style-type: none"> • Expunere interactivă 	

<p>accent pe generarea de text folosind modele pre-antrenate.</p>	<ul style="list-style-type: none"> • Explicație • Conversație • Lucru individual și în echipă • Dialog, dezbateri 	
<p>2. Explorarea arhitecturilor de tip transformer. Analiza modelelor transformer, înțelegerea mecanismelor de atenție și implementarea acestora în sarcini de generare de text.</p>	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Lucru individual și în echipă • Dialog, dezbateri 	
<p>3. Generare avansată de text și ingineria prompturilor. Experimente cu tehnici avansate de generare de text și învățarea artei de a construi prompturi eficiente pentru a ghida ieșirile LLM-urilor.</p>	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Lucru individual și în echipă • Dialog, dezbateri 	
<p>4. Strategii de pre-antrenare și fine-tuning. Sesiune practică privind bazele pre-antrenării LLM-urilor și strategiile de fine-tuning pentru sarcini specifice.</p>	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Lucru individual și în echipă • Dialog, dezbateri 	
<p>5. Învățare prin întărire din feedback uman (RLHF). Implementarea tehnicilor RLHF, configurarea buclelor de feedback și înțelegerea modelelor de recompensă pentru alinierea ieșirilor LLM-urilor cu valorile umane.</p>	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Lucru individual și în echipă • Dialog, dezbateri 	
<p>6. Introducere în LangChain și generare augmentată prin regăsire (RAG). Activități practice cu</p>	<ul style="list-style-type: none"> • Expunere interactivă 	

<p>LangChain: încărcarea documentelor, vector stores și embeddings. Explorarea implementării RAG pentru îmbunătățirea aplicațiilor LLM.</p>	<ul style="list-style-type: none"> • Explicație • Conversație • Lucru individual și în echipă • Dialog, dezbateri 	
<p>7. Crearea unui chatbot. Aplicarea cunoștințelor din LangChain și RAG pentru dezvoltarea unui chatbot funcțional.</p> <p>PROIECT</p> <p>Faza 1 (Săptămânile 1 și 2): Introducere și alegerea temei</p> <p>Prezentarea unei liste de teme de proiect ce includ utilizarea LLM-urilor, axate pe cerințele clienților din lumea reală. Studenții își aleg sau propun propriile teme de proiect, lucrând în echipe. Discuție despre proiectele alese pentru a asigura fezabilitatea și relevanța acestora, folosind metodologia ciclului de viață al proiectelor de AI generativă.</p> <p>Analiză inițială a stadiului actual al tehnologiei, concentrată pe abordările similare existente folosind LLM-uri.</p> <p>Faza 2 (Săptămânile 3 și 4): Pregătire și planificare</p> <p>În funcție de tema aleasă, fiecare echipă identifică și definește o listă de aplicații NLP, apoi realizează o cercetare a literaturii pentru a determina cele mai performante modele pre-antrenate pentru cazurile de utilizare respective.</p>	<ul style="list-style-type: none"> • Expunere interactivă • Explicație • Conversație • Lucru individual și în echipă • Dialog, dezbateri • Expunere interactivă • Explicație • Conversație • Lucru individual și în echipă 	

Faza 3 (Săptămânile 5 și 6): Adaptarea și alinierea modelului I

Aplicarea tehnicilor de inginerie a prompturilor pentru rafinarea ieșirilor modelului fără reantrenare, urmată de o evaluare a performanței acestuia.

Faza 4 (Săptămânile 7 și 8): Adaptarea și alinierea modelului II

Implementarea metodelor de fine-tuning pentru reantrenarea modelelor, îmbunătățindu-le performanța pentru particularitățile temei alese, urmată de evaluarea acestora.

Faza 5 (Săptămânile 9 și 10): Adaptarea și alinierea modelului III

Incorporarea RLHF și a modelelor de recompensă pentru a ajusta ieșirile LLM-ului în concordanță cu valorile umane.

Faza 6 (Săptămânile 11 și 12): LangChain și generare augmentată prin regăsire (RAG)

Folosind LangChain și RAG, studenții integrează LLM-ul dezvoltat într-un flux de lucru concret. Această integrare trebuie să asigure coerența rezultatelor modelului cu cerințele tematice, prin utilizarea încărcării documentelor, vector stores și embeddings.

Faza 7 (Săptămânile 13 și 14): Prezentări orale Prezentarea finală a proiectelor.		
Bibliografie 1. Chung, H. W., Hou, L., Longpre, S., Zoph, B., Tay, Y., Fedus, W., Li, Y., Wang, X., Dehghani, M., Brahma, S., Webson, A., Gu, S. S., Dai, Z., Suzgun, M., Chen, X., Chowdhery, A., Pellat, M., Robinson, K., Valter, D., . . . Wei, J. (2022). <i>Scaling Instruction-Finetuned Language Models</i> . 2. Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2022). <i>ReAct: Synergizing Reasoning and Acting in Language Models</i> . 3. Wu, S., Irsoy, O., Lu, S., Dabrovolski, V., Dredze, M., Gehrmann, S., Kambadur, P., Rosenberg, D., & Mann, G. (2023). <i>BloombergGPT: A Large Language Model for Finance</i> . 4. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). <i>Attention Is All You Need</i> . 5. Alammari, J, Grootendorst, M. (2024). <i>Hands-On Large Language Models</i> . 6. Auffarth, B. (2023). <i>Generative AI with LangChain: Build large language model (LLM) apps with Python, ChatGPT and other LLMs</i>		


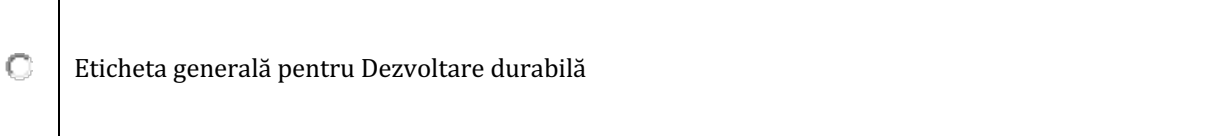









9. Evaluare

Tip activitate	9.1 Criterii de evaluare	9.2 Metode de evaluare	9.3 Pondere din nota finală
9.4 Curs	<ul style="list-style-type: none"> Capacitatea de a utiliza cunoștințele dobândite în cadrul cursului și aplicate în laboratoare pentru a aborda probleme practice și cerințe din lumea reală, cu aplicații în procesarea limbajului natural și inteligența artificială generativă. 	<ul style="list-style-type: none"> Examinare orală (proiect) 	60%
9.5 Seminar/laborator	<ul style="list-style-type: none"> Studentul deține o înțelegere aprofundată a conceptelor legate de modelele lingvistice de mari dimensiuni (LLM), inclusiv arhitecturile de tip transformer, ingineria prompturilor și aplicațiile LangChain. 	<ul style="list-style-type: none"> Examinare practică sub observație continuă (rezolvarea sarcinilor de laborator) 	40%
9.6 Standard minim de performanță			
<ul style="list-style-type: none"> Studentii trebuie să dovedească faptul că au dobândit un nivel acceptabil de cunoștințe și înțelegere a conceptelor de bază predate în cadrul cursului, că sunt capabili să utilizeze aceste cunoștințe într-o formă 			

coerentă, că au abilitatea de a stabili conexiuni și de a folosi informațiile pentru rezolvarea diverselor probleme de viziune computerizată.

• Nota finală (media ponderată între cele două metode de evaluare prezentate) trebuie să fie de cel puțin 5 (fără rotunjire, pe o scară de la 1 la 10).

10. Etichete ODD (Obiective de Dezvoltare Durabilă / Sustainable Development Goals)¹

								
								
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
								Nu se aplică nici o etichetă
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Data completării:

15.05.2026

Semnătura titularului de curs

Lect. Dr. Bogdan MURSA

Semnătura titularului de seminar

Lect. Dr. Bogdan MURSA

Data avizării în departament:

...

Semnătura directorului de departament

Conf.dr. Adrian STERCA

¹ Selectați o singură etichetă, cea care, în conformitate cu [Procedura de aplicare a etichetelor ODD în procesul academic](#), se potrivește cel mai bine disciplinei. Dacă disciplina tratează tema dezvoltării durabile la modul general (de ex. prin prezentarea/introducerea cadrului general al dezvoltării durabile etc.) atunci se poate alocă eticheta generală de Dezvoltare Durabilă. Dacă niciuna dintre etichete nu descrie disciplina, selectați ultima opțiune: „Nu se aplică nici o etichetă”.