

FIȘA DISCIPLINEI

Introducere în prelucrarea limbajului natural

Anul universitar 2025-2026

1. Date despre program

1.1. Instituția de învățământ superior	Universitatea Babeș-Bolyai Cluj-Napoca
1.2. Facultatea	Facultatea de Matematică și Informatică
1.3. Departamentul	Departamentul de Informatică
1.4. Domeniul de studii	Informatică
1.5. Ciclul de studii	Licență
1.6. Programul de studii / Calificarea	Informatică în limba engleză
1.7. Forma de învățământ	Cu frecvență

2. Date despre disciplină

2.1. Denumirea disciplinei	Introducere în prelucrarea limbajului natural			Codul disciplinei	MLE8151		
2.2. Titularul activităților de curs	Lect. dr. Lupea Mihaiela						
2.3. Titularul activităților de seminar	Lect. Dr. Lupea Mihaiela						
2.4. Anul de studiu	3	2.5. Semestrul	6	2.6. Tipul de evaluare	C	2.7. Regimul disciplinei	Optional

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	1lab+ 2pr
3.4. Total ore din planul de învățământ	60	din care: 3.5. curs	24	3.6 seminar/laborator/proiect	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)					15
Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate și pe teren					14
Pregătire seminare/ laboratoare/ proiecte, teme, referate, portofolii și eseuri					14
Tutoriat (consiliere profesională)					5
Examinări					7
Proiect practic					10
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	Limbaje formale, Structuri de date și algoritmi
4.2. de competențe	Cunoașterea unui limbaj de programare de nivel înalt.

5. Condiții (acolo unde este cazul)

5.1. de desfășurare a cursului	-
5.2. de desfășurare a seminarului/ laboratorului	Laborator cu calculatoare, medii de programare (.NET , Java)

6.1. Competențele specifice acumulate¹

Competențe profesionale/esențiale	<ul style="list-style-type: none">• Utilizarea instrumentelor informatice în context interdisciplinar• Folosirea conceptelor și tehnicilor inteligenței artificiale la rezolvarea unor probleme din lumea reală
Competențe transversale	<ul style="list-style-type: none">• Aplicarea regulilor de muncă organizată și eficientă, a unor atitudini responsabile față de domeniul didactic-științific, pentru valorificarea creativă a propriului potențial, cu respectarea principiilor și a normelor de etică profesională• Utilizarea unor metode și tehnici eficiente de învățare, informare, cercetare și dezvoltare a capacităților de valorificare a cunoștințelor, de adaptare la cerințele unei societăți dinamice și de comunicare în limba română și într-o limbă de circulație internațională

6.2. Rezultatele învățării

Cunoștințe	<ul style="list-style-type: none">• Absolventul are cunoștințe necesare pentru utilizarea calculatoarelor, dezvoltarea programelor și aplicațiilor software, procesarea informațiilor.• Absolventul are cunoștințe necesare pentru revizuirea literaturii de specialitate.
Aptitudini	<ul style="list-style-type: none">• Absolventul este capabil să identifice probleme complexe și să examineze probleme conexe pentru a dezvolta opțiuni de rezolvare și implementa soluții.• Absolventul are capacitatea de a alege și folosi paradigme de programare (procedural, orientat obiect, funcțional) pentru realizarea de aplicații software adecvate specificului domeniului aplicației dezvoltate.• Absolventul are abilitatea de a alege și folosi module și medii existente pentru dezvoltarea aplicațiilor• Absolventul este capabil de a redacta un raport științific.
Responsabilități și autonomie	<ul style="list-style-type: none">• Absolventul este capabil de a defini/identifica/înțelege probleme de cercetare în domeniul informaticii.• Absolventul are abilitatea de a înțelege și comunica eficient informațiile.• Absolventul are deprinderile necesare pentru utilizarea instrumentelor de sprijinire a cercetării.

7. Obiectivele disciplinei (reieșind din grila competențelor acumulate)

7.1 Obiectivul general al disciplinei	<ul style="list-style-type: none">• Introducerea principiilor de bază și problemelor specifice domeniului prelucrării limbajului natural (PLN).• Cunoașterea stadiului actual al cercetării în domeniul prelucrării limbajului natural, pentru a implementa sisteme care rezolva task-uri specifice PLN.
----------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

¹ Se poate opta pentru competențe sau pentru rezultatele învățării, respectiv pentru ambele. În cazul în care se alege o singură variantă, se va șterge tabelul aferent celeilalte opțiuni, iar opțiunea păstrată va fi numerotată cu 6.

7.2 Obiectivele specifice	<ul style="list-style-type: none"> • Aplicarea și utilizarea de modele formale (logici, gramatici formale, grafe), modele statistice (HMM- Hidden Markov Model), tehnici și algoritmi proprii inteligenței artificiale pentru rezolvarea unor probleme specifice domeniului PLN (limbile română și engleză): detectarea părților de vorbire, analiza sintactică, dezambiguizare, sumarizare de documente, rezolvarea anaferei, analiza sentimentelor și emoțiilor în texte. • Utilizarea de modele mari de limbaje (LLMs) pentru a rezolva task-uri PLN.
----------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

8. Conținuturi

8.1 Curs	Metode de predare	Observații
1. Prelucrarea limbajului natural (PLN): principii de bază, probleme specifice, abordări și aplicații.	Prelegere teoretică, explicații, exemple, studiu de caz, dezbateri, dialog	
2. Detectarea părților de vorbire și a entităților lexicale în limbile engleză și română Baze de date lexicale pentru limbile engleză și română - WordNet și RoWordNet - structura, relații semantice și relații lexicale		
3. Reprezentare textelor și modele de limbaje naturale. Clasificarea textelor.		
4. Analiza sintactică a textelor reguli gramaticale pentru construirea frazelor în limba engleză; gramatici independente de context algoritmul Cocke-Kasami-Yonger de analiză sintactică		
5. HMM (Hidden Markov Model) și aplicarea lui la detectarea părților de vorbire în texte. Algoritmii "forward" și Viterbi.		
6. Extragerea cuvintelor cheie din texte -algoritmii TextRank și RAKE		
7. Sumarizarea documentelor folosind abordări bazate pe clustering și grafe		
8. Prezentări rapoarte de cercetare de către studenți	Evaluare	
9. Detectarea opiniilor, sentimentelor și emoțiilor din text	Prelegere teoretică, explicații, exemple, studiu de caz, dezbateri, dialog	
10. Rezolvarea anaferei în texte		
11. Dezambiguizarea cuvintelor în texte –metode bazate de dicționare și grafe		
12. Prezentare proiecte de către studenți	Evaluare	
<p>Bibliografie</p> <ol style="list-style-type: none"> 1. J.ALLEN : Natural language understanding, Benjamin/Cummings Publisher, 2nd ed., 1995. 2. E. CHARNIAK: Statistical language learning, MIT press, 1996. 3. L. DENG, Y. LIU: Deep learning in Natural Language Processing, Springer Verlag, Singapore, 2018 4. D.FEHRER et al: Description logics for natural language processing. In Proc. of the 1994 Description Logic Workshop (DL'94), 1994. 5. H. HELBIG: Knowledge Representation and the Semantics of Natural Language, Springer, 2006. 6. D.JURAFSKY, J.MARTIN: Speech and language processing, Prentice Hall, 2000. 7. C.MANNING, H.SCHUTZE: Foundation of statistical natural language processing, MIT, 1999. 8. R. MITKOV(ed): The Oxford Handbook of Computational Linguistics, Oxford University Press, 2003. 9. D. TATAR: Inteligența artificială. Aplicații în prelucrarea limbajului natural, Editura Albastra, Microinformatică, 2003, ISBN 973-650-100-01. 10. S. VAJJALA, B. MAJUMDER, A. GUPTA, H. SURANA: Practical Natural Language Processing. A Comprehensive Guide to Building Real-World NLP Systems, O'REILLY. 2020. 		
8.2 Seminar / laborator	Metode de predare	Observații
1. Studiul de platforme și biblioteci în diferite limbaje de programare care oferă funcții de	Documentare pe platforme electronice, studii de caz, explicații	Laboratorul este

preprocesarea textelor in limbile romana si engleza.		structurat sub forma a 2 ore din 2 în 2 săptămâni.
2. Studiul modelelor mari de limbaje (LLM) pentru limba romana si engleza. Studiul bazelor de date lexicale WordNet (limba engleză), RoWordNet (limba română) Utilizarea de tool-uri existente pentru diferite task-uri: extragere cuvinte cheie, sumarizare, rezolvarea anaferei, analiza sentimentelor.		
3. Identificarea de task-uri PLN pentru limba romana. Alegerea task-ului care va fi rezolvat, studiul diferitelor abordari de rezolvare si alegerea metodei care va fi implementata. Cautare de date de intrare pentru task-ul ales.		
4. Prezentarea de catre studenti a referatului teoretic despre un task sau tool PLN.	Evaluare	
5. Discutii despre implementarea unui PLN tool. Dezvoltarea de resurse lingvistice pentru limba romana.	Explicatii, dialog, studiu de caz	
6. Prezentarea proiectelor de către studenți	Evaluare	
Bibliografie 1. Rada Mihalcea: www.cs.unt.edu/~rada/downloads.html 2. Resurse lingvistice in limba romana: www.racai.ro		

9. Coroborarea conținuturilor disciplinei cu așteptările reprezentanților comunității epistemice, asociațiilor profesionale și angajatori reprezentativi din domeniul aferent programului

- Cursul respectă Recomandările IEEE și ACM Curricula pentru studii de informatică;
- Cursul există în programul de studii al tuturor universităților majore din România și din străinătate;
- Optimizarea căutării pe Web, interfețele în limbaj natural și aspectele recente ale procesării textelor necesită o bună înțelegere a prelucrării limbajului natural.

10. Evaluare

Tip activitate	10.1 Criterii de evaluare	10.2 Metode de evaluare	10.3 Pondere din nota finală
10.4 Curs	Redactarea unui referat teoretic	Raport teoretic - prezentarea unei probleme specifice domeniului și a unui instrument existent care o rezolvă	35%
10.5 Seminar/ laborator	Aplicarea unor concepte teoretice in rezolvarea unor probleme practice din domeniu PLN	Dezvoltarea unor resurse lingvistice specifice limbii romane	20%
	Implementarea unor algoritmi prezentati la curs	Proiect practic – rezolvarea unei probleme specifice domeniului prelucrării limbajului natural.	35%
10.6 Activitate	Activitatea la curs și laborator	Prezența la activități, realizarea la timp a activităților solicitate	10%
10.7 Standard minim de performanță			
<ul style="list-style-type: none"> • Cel puțin nota 5 (pe o scara de 1 la 10) la lucrările scrise și activitatea de la seminarii. 			

11. Etichete ODD (Obiective de Dezvoltare Durabilă / Sustainable Development Goals)²

Nu se aplică.

² Păstrați doar etichetele care, în conformitate cu *Procedura de aplicare a etichetelor ODD în procesul academic*, se potrivesc disciplinei și ștergeți-le pe celelalte, inclusiv eticheta generală pentru *Dezvoltare durabilă* - dacă nu se aplică. Dacă nicio etichetă nu descrie disciplina, ștergeți-le pe toate și scrieți "Nu se aplică".

Data completării:
1.04.2025

Semnătura titularului de curs

Lect. dr. Mihaiela LUPEA

Semnătura titularului de seminar

Lect. dr.. Mihaiela LUPEA

Data avizării în departament:

...

Semnătura directorului de departament

Conf.dr. Adrian STERCA