FIŞA DISCIPLINEI

Introducere în prelucrarea limbajului natural

Anul universitar 2025-2026

1. Date despre program

1.1. Instituția de învăţământ superior	**Universitatea Babeş-Bolyai Cluj-Napoca**
1.2. Facultatea	**Facultatea de Matematică şi 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 romana**
1.7. Forma de învățământ	**Cu frecventa**

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 activităţilor 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) şi activități de autoinstruire (AI)					**ore**
Studiul după manual, suport de curs, bibliografie şi notițe (AI)					15
Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate şi pe teren					14
Pregătire seminare/ laboratoare/ proiecte, teme, referate, portofolii şi eseuri					14
Tutoriat (consiliere profesională)					5
Examinări					7
Proiect practic					10
3.7. Total ore studiu individual (SI) şi activităţi de autoinstruire (AI)				**65**	
3.8. Total ore pe semestru				**125**	
3.9. Numărul de credite				**5**	

4. Precondiţii (acolo unde este cazul)

4.1. de curriculum	Limbaje formale, Structuri de date şi algoritmi	
4.2. de competențe	Cunoasterea unui limbaj de programare de nivel înalt.	

5. Condiţii (acolo unde este cazul)

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

6.1. Competenţele specifice acumulate[1]

Competenţe profesionale/esenţiale	• Utilizarea instrumentelor informatice în context interdisciplinar • Folosirea conceptelor şi tehnicilor inteligenţei artificiale la rezolvarea unor probleme din lumea reală
Competenţe transversale	• Aplicarea regulilor de muncă organizată şi eficientă, a unor atitudini responsabile faţă de domeniul didactic-ştiinţific, pentru valorificarea creativă a propriului potenţial, cu respectarea principiilor şi a normelor de etică profesională • Utilizarea unor metode şi tehnici eficiente de învăţare, informare, cercetare şi dezvoltare a capacităţilor de valorificare a cunoştinţelor, de adaptare la cerinţele unei societăţi dinamice şi de comunicare în limba română şi într-o limbă de circulaţie internaţională

6.2. Rezultatele învăţării

Cunoştinţe	• Absolventul are cunoştinţe necesare pentru utilizarea calculatoarelor, dezvoltarea programelor şi aplicaţiilor software, procesarea informaţiilor. • Absolventul are cunoştinţe necesare pentru revizuirea literaturii de specialitate.
Aptitudini	• Absolventul este capabil să identifice probleme complexe şi să examineze probleme conexe pentru a dezvolta opţiuni de rezolvare şi implementa soluţii. • Absolventul are capacitatea de a alege si folosi paradigme de programare (procedural, orientat obiect, funcţional) pentru realizarea de aplicaţii software adecvate specificului domeniului aplicaţiei dezvoltate. • Absolventul are abilitatea de a alege si folosi module si medii existente pentru dezvoltarea aplicaţiilor • Absolventul este capabil de a redacta un raport ştiinţific.
Responsabilităţi şi autonomie	• Absolventul este capabil de a defini/identifica/înţelege probleme de cercetare în domeniul informaticii. • Absolventul are abilitatea de a înţelege şi comunica eficient informaţiile. • Absolventul are deprinderile necesare pentru utilizarea instrumentelor de sprijinire a cercetării.

7. Obiectivele disciplinei (reieşind din grila competenţelor acumulate)

7.1 Obiectivul general al disciplinei	• Introducerea principiilor de bază şi problemelor specifice domeniului prelucrării limbajului natural (PLN). • Cunoaşterea stadiului actual al cercetării în domeniul prelucrării limbajului natural, pentru a implementa sisteme care rezolva task-uri specifice PLN.

[1] Se poate opta pentru competenţe sau pentru rezultatele învăţării, respectiv pentru ambele. În cazul în care se alege o singură variantă, se va şterge tabelul aferent celeilalte opţiuni, iar opţiunea păstrată va fi numerotată cu 6.

7.2 Obiectivele specifice	• Aplicarea şi utilizarea de modele formale (logici, gramatici formale, grafe), modele statistice (HMM- Hidden Markov Model), tehnici şi algoritmi proprii inteligenţei artificiale pentru rezolvarea unor probleme specifice domeniului PLN (limbile română şi engleză): detectarea părţilor de vorbire, analiza sintactică, dezambiguizare, sumarizare de documente, rezolvarea anaforei, analiza sentimentelor şi emoţiilor în texte. • Utilizarea de modele mari de limbaje (LLMs) pentru a rezolva task-uri PLN.

8. Conţinuturi

8.1 Curs	Metode de predare	Observaţii
1. Prelucrarea limbajului natural (PLN): principii de bază, probleme specifice, abordări şi aplicaţii.	Prelegere teoretică, explicaţii, exemple, studiu de caz, dezbatere, dialog	
2. Detectarea părţilor de vorbire şi a entităţilor lexicale în limbile engleză şi română Baze de date lexicale pentru limbile engleză şi română - WordNet şi RoWordNet - structura, relaţii semantice şi relaţii lexicale		
3. Reprezentare textelor si modele de limbaje naturale. Clasificarea textelor.		
4. Analiza sintactica 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) si aplicarea lui la detectarea partilor de vorbire in texte. Algoritmii ”forward” si Viterbi.		
6 Extragerea cuvintelor cheie din texte -algoritmii TextRank si RAKE		
7. Sumarizarea documentelor folosind abordari bazate pe clustering si grafe		
8. Prezentari rapoarte de cercetare de catre studenti	Evaluare	
9. Detectarea opiniilor, sentimentelor si emotiilor din text	Prelegere teoretică, explicaţii, exemple, studiu de caz, dezbatere, dialog	
10. Rezolvarea anaforei in texte		
11. Dezambiguizarea cuvintelor in texte –metode bazate de dictionare si grafe		
12. Prezentare proiecte de catre studenţi	Evaluare	

Bibliografie

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: Inteligenta artificiala. Aplicatii in prelucrarea limbajului natural, Editura Albastra, Microinformatica, 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	Observaţii
1. Studiul de platforme si biblioteci in diferite limbaje de programare care ofera functii de	Documentare pe platforme electronice, studii de caz, explicatii	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 anaforei, 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 studenţi	Evaluare	
Bibliografie 1. Rada Mihalcea: www.cs.unt.edu/~rada/downloads.html 2. Resurse lingvistice in limba romana: www.racai.ro		

9. Coroborarea conţinuturilor disciplinei cu aşteptările reprezentanţilor comunităţii epistemice, asociaţiilor profesionale şi angajatori reprezentativi din domeniul aferent programului

- Cursul respectă Recomandările IEEE şi ACM Curricula pentru studii de informatică;
- Cursul există în programul de studii al tuturor universităţilor majore din România şi din străinătate;
- Optimizarea căutării pe Web, interfeţele în limbaj natural şi aspectele recente ale procesării textelor necesită o bună înţelegere 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 şi 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 prelucrarii limbajului natural.	35%
10.6 Activitate	Activitatea la curs şi laborator	Prezenţa la activităţi, realizarea la timp a activităţilor solicitate	10%
10.7 Standard minim de performanţă			
• Cel puţin nota 5 (pe o scara de 1 la 10) la lucrările scrise şi activitatea de la seminarii.			

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

Nu se aplică.

[2] Păstraţi doar etichetele care, în conformitate cu *Procedura de aplicare a etichetelor ODD în procesul academic*, se potrivesc disciplinei şi ştergeţi-le pe celelalte, inclusiv eticheta generală pentru *Dezvoltare durabilă* - dacă nu se aplică. Dacă nicio etichetă nu descrie disciplina, ştergeţi-le pe toate şi scrieţi "*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