FIŞA DISCIPLINEI

Sisteme bazate pe cunoştinţe şi tehnologia limbajului

Anul universitar 2025-2026

1. Date despre program

1.1. Instituţia 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	**Master**
1.6. Programul de studii / Calificarea	**Calcul de Înaltă Performanţă şi Analiza Volumelor Mari de Date**
1.7. Forma de învăţământ	**Cu frecventa**

2. Date despre disciplină

2.1. Denumirea disciplinei	**Sisteme bazate pe cunoştinţe şi tehnologia limbajului**					Codul disciplinei	**MME8044**
2.2. Titularul activităţilor de curs	**Lect.dr. Lupea Mihaiela**						
2.3. Titularul activităţilor de seminar	**Lect.dr. Lupea Mihaiela**						
2.4. Anul de studiu	1	2.5. Semestrul	2	2.6. Tipul de evaluare	E	2.7. Regimul disciplinei	Optional

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

3.1. Număr de ore pe săptămână	**4**	din care: 3.2. curs	**2**	3.3. seminar/ laborator/proiect	**1sem +1pr**
3.4. Total ore din planul de învăţământ	56	din care: 3.5. curs	28	3.6 seminar/laborator/proiect	**28**
Distribuţia fondului de timp pentru studiul individual (SI) şi activităţi de autoinstruire (AI)					**ore**
Studiul după manual, suport de curs, bibliografie şi notiţe (AI)					25
Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate şi pe teren					15
Pregătire seminare/ laboratoare/ proiecte, teme, referate, portofolii şi eseuri					25
Tutoriat (consiliere profesională)					10
Examinări					15
Alte activităţi					29
3.7. Total ore studiu individual (SI) şi activităţi de autoinstruire (AI)				**119**	
3.8. Total ore pe semestru				**175**	
3.9. Numărul de credite				**7**	

4. Precondiţii (acolo unde este cazul)

4.1. de curriculum	Limbaje formale, structuri de date, invatare automata
4.2. de competenţe	Abilitati bune de programare

5. Condiţii (acolo unde este cazul)

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

6.1. Competenţele specifice acumulate[1]

Competenţe profesionale/esenţiale	• Înţelegerea şi operarea cu conceptele de bază ale calculului de înaltă performanţă si a tehnicilor de analiză a datelor; • Modelare şi rezolvarea eficienta de probleme din lumea reală.
Competenţe transversale	• Abilităţi de comunicare profesională: descrierea clară, concisă, verbală şi în scris, a rezultatelor profesionale, abilităţi de negociere; • Comportarea onorabilă, etică, respectarea deontologiei profesionale.

6.2. Rezultatele învăţării

Cunoştinţe	• Studentul/absolventul cunoaşte principiile, tehnicile şi aplicaţiile de bază ale învăţării automate • Studentul/absolventul cunoaşte metodele avansate de analiză a datelor
Aptitudini	• Studentul/absolventul este în măsură să ofere consultanţă ştiinţifică de specialitate şi să elaboreze materiale de specialitate • Studentul/absolventul ştie să gestioneze cantităţi (extrem de) mari de date digitale în diverse formate (text, video, financiar, medical etc.) • Studentul/absolventul este capabil să utilizeze algoritmi noi, infrastructuri software şi metodologii în scopul prelucrării (stocării, achiziţiei/preluării, analizei) a unor cantităţi mari de date
Responsabilităţi şi autonomie	• Studentul/absolventul cunoaşte şi respectă norme şi reguli etice şi deontologice în cercetarea ştiinţifică • Studentul/absolventul îşi asumă responsabilitatea pentru produsul muncii sale, solicită feedback şi îl utilizează constructiv • Studentul/absolventul îşi dezvoltă capacitatea de a translata cunoştinţele academice în context profesional, economic, social şi etic.

7. Obiectivele disciplinei (reieşind din grila competenţelor acumulate)

7.1 Obiectivul general al disciplinei	• Introducerea principiilor, tehnicilor şi aplicaţiilor de bază ale Tehnologiei limbajului şi ale sistemelor bazate pe cunoştinţe. • Studierea impactului utilizarii limbajului natural in diferite domenii.
7.2 Obiectivele specifice	• Aplicare modele formale (logici, gramatici), modele statistice (HMM), algoritmi de inteligenţă artificială (clusterizare, invatare automata) şi tehnici (nespervizate, supervizate) pentru a rezolva diferite task-uri la nivel sintactic (detectare parte de vorbire, analiza sintactica) şi nivel semantic (rezumarea documentelor, dezambiguizarea sensului cuvântului, extragerea informaţiilor, rezoluţia anaforei, analiza sentimentelor) în domeniul Procesării Limbajului Natural (PLN). • Utilizarea de modele mari de limbaj natural (LLMs) pentru rezolvarea unor task-uri de PLN in limbile romana si engleza.

[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.

8. Conţinuturi

8.1 Curs	Metode de predare	Observaţii
1. Procesarea limbajului Natural (PLN): task-uri specifice, abordari, aplicatii.	• Expunerea interactivă • Prezentarea • Explicarea • Exemple practice • Discuţii pe studii de caz	
2. Detectarea partilor de vorbire in texte. Prezentare WordNet si RoWordNet.		
3. Modalitati de reprezentare a textelor. Clasificarea textelor.		
4. Analiza sintactica a textelor		
5. Modelul statistic HMM si aplicarea lui la detectarea partilor de vorbire in text		
6. Extragerea cuvintelor cheie din text		
7. Sumarizarea documentelor		
8. Detectarea opiniilor, sentimentelor si emotiilor din text		
9. Prezentari rapoarte de cercetare de catre studenti		
10. Rezolvarea anaforei in texte		
11. Dezambiguizarea cuvintelor in texte		
12. Implicarea textuala		
13. Extragerea informatiei din text		
14. Prezentare proiecte de catre studenţi		

Bibliografie

1. J. ALLEN : Natural language understanding, Benjamin/Cummings Publisher, 2nd ed., 1995.
2. L. DENG, Y. LIU: Deep learning in Natural Language Processing, Springer Verlag, Singapore, 2018.
3. H. HELBIG: Knowledge Representation and the Semantics of Natural Language, Springer, 2006.
4. D. JURAFSKY, J. MARTIN: Speech and language processing, Prentice Hall, 2000.
5. C. MANNING, H. SCHUTZE: Foundation of statistical natural language processing, MIT, 1999.
6. R. MITKOV(ed): The Oxford Handbook of Computational Linguistics, Oxford University Press, 2003.
7. G. PAAS, S. Giesselbach: Foundation Models for Natural Language Processing. Pre-trained Language Models Integrating Media, Springer, Berlin, 2022.
8. D. ROTHMAN: Transformers for Natural Language Processing, Build, train, and fine-tune deep neural network architecture for NLP with Python, PyTorch, TensorFlow, BERT, and GPT-3, Second edition, Packt Publishing, 2022.
9. D. TATAR: Inteligenta artificiala. Aplicatii in prelucrarea limbajului natural, Editura Albastra, Microinformatica, 2003.
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 preprocesarea textelor in limbile romana si engleza.	Documentare pe platforme electronice, explicaţie, dialog, studii de caz.	2 ore de seminar la 2 saptamani
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.	Documentare pe platforme electronice, explicaţie, dialog, studii de caz	
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.	Documentare pe platforme electronice, explicaţie, dialog, studii de caz.	
4. Prezentari referate teoretice de catre studenti	Dialog, dezbatere.	

5. Discutii individuale cu studentii despre temele alese pentru proiectele practice.	Explicatii, dialog, dezbatere.	
6. Dezvoltare de resurse lingvistice pentru limba romana	Documentare pe platforme electronice, studii de caz.	
7. Prezentari proiecte practice de catre studenti	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 pentru Curricula in domeniul Informaticii
- Cursul există în programul de studii al tuturor universităţilor importante din România şi din străinătate

10. Evaluare

Tip activitate	10.1 Criterii de evaluare	10.2 Metode de evaluare	10.3 Pondere din nota finală
10.4 Curs	- Cunoaşterea conceptelor teoretice ale domeniului; - Aplicarea metodelor si algoritmilor de la curs in rezolvarea problemelor de PLN - Sintetizarea şi compararea diferitelor abordări pentru rezolvarea task-urilor specifice PLN.	Examen scris *sau* raport de cercetare in domeniul PLN	40%
10.5 Seminar/laborator	Identificarea unui task practic de PLN si implementarea unui instrument care il rezolva.	Implementarea proiectului şi prezentare.	40%
	Aplicarea conceptelor teoretice in aplicatii practice	Dezvoltarea de resurse lingvistice pentru limba romana.	20%
10.6 Standard minim de performanţă			
• Nota finala sa fie cel putin 5 (pe o scara de la 1 la 10).			

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

Nu se aplică.

Data completării:
15.04.2025

Semnătura titularului de curs

Lect.dr. Lupea Mihaiela

Semnătura titularului de seminar

Lect.dr. Lupea Mihaiela

Data avizării în departament:
...

Semnătura directorului de departament

Conf.dr. Adrian STERCA

[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.*".