



FIȘA DISCIPLINEI

1. Date despre program

1.1 Instituția de învățământ superior	Universitatea “Alexandru Ioan Cuza” din Iași
1.2 Facultatea	Facultatea de Informatică
1.3 Departamentul	Informatică
1.4 Domeniul de studii	Informatică
1.5 Ciclul de studii	Licență
1.6 Programul de studii / Calificarea	Informatică

2. Date despre disciplină

2.1 Denumirea disciplinei	Tehnici de Ingineria Limbajului Natural						
2.2 Titularul activităților de curs	CS II Dr. Daniela GÎFU						
2.3 Titularul activităților de seminar	CS II Dr. Daniela GÎFU, Conf. Dr. Diana TRANDABĂȚ, Conf. Dr. Ionuț PISTOL, dr. Mihaela Onofrei						
2.4 An de studiu	I	2.5 Semestru	2	2.6 Tip de evaluare	E	2.7 Regimul disciplinei*	OB

* OB – Obligatoriu / OP – Opțional

3. Timpul total estimat (ore pe semestru și activități didactice)

3.1 Număr de ore pe săptămână	4	din care: 3.2 curs	2	3.3 seminar/laborator	2
3.4 Total ore din planul de învățământ	56	din care: 3.5 curs	28	3.6 seminar/laborator	28
Distribuția fondului de timp					ore
Studiu după manual, suport de curs, bibliografie și altele					28
Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate și pe teren					28
Pregătire seminarii/laboratoare, teme, referate, portofolii și eseuri					48
Tutoriat					-
Examinări					4
Alte activități					-
3.7 Total ore studiu individual					104
3.8 Total ore pe semestru					164
3.9 Număr de credite					8

4. Precondiții (dacă este cazul)

4.1 De curriculum	Cursul de Inteligență artificială
4.2 De competențe	Cunoștințe de algoritmică și programare

5. Condiții (dacă este cazul)

5.1 De desfășurare a cursului	<ul style="list-style-type: none">- Studenții se vor prezenta la prelegerile ținute în clasă, seminarii/laboratoare, cu telefoanele mobile închise.- Studenții trebuie să fie pregătiți ca o parte din ore să fie ținute în regim online (pe perioada stării de urgență și alertă).
5.2 De desfășurare a seminarului/laboratorului	<ul style="list-style-type: none">- Termenul predării temelor de seminar este stabilit de titularul de laborator, de comun acord cu studenții.- Nu se acceptă cererile de amânare decât pe motive obiectiv întemeiate.



- Pentru predarea cu întârziere a temelor de seminar, acestea vor fi depunctate cu 0,5 pct./zi de întârziere.

6. Competențe specifice acumulate

Competențe profesionale	C1. Cunoașterea teoriilor semnificative și a metodelor de lucru utilizate în ingineria limbajului natural. C2. Dezvoltarea de deprinderi de utilizare a unor aplicații de prelucrarea limbajului natural. C3. Abilitatea de a adapta și implementa modele existente la specificul limbii române. C4. Capacitatea de a utiliza limbajul științific al domeniului tehnologiilor limbajului în interacțiuni cu specialiști informaticieni și lingviști.
Competențe transversale	CT1. Identificarea rolurilor și responsabilităților într-o echipă pluridisciplinară și folosirea de tehnici de comunicare, relaționare și muncă eficientă în cadrul unei echipe. CT2. Aplicarea normelor și valorilor de etică profesională CT3. Participarea la proiecte cu caracter științific în condiții de autonomie și independență profesională și demonstrarea capacității de a identifica oportunități pentru propria formare profesională în viitor. CT4. Utilizarea eficientă a surselor informaționale și a resurselor de comunicare și formare profesională asistată, atât în limba română, cât și într-o limbă de circulație internațională.

7. Obiectivele disciplinei (din grila competențelor specifice acumulate)

7.1 Obiectivul general	Inițierea în domeniul teoretic și aplicativ al ingineriei limbajului natural (ILN) / prelucrarea limbajului natural (PLN).
7.2 Obiectivele specifice	La finalizarea cu succes a acestei discipline, studenții vor fi capabili să: § explice caracteristicile modelelor folosite în ILN/PLN ; § deprindă utilizarea metodelor simbolice, statistice și neurale aplicate în ILN/PLN și adecvarea lor în diverse aplicații; § descrie conceptele și metodele utilizate în dezvoltarea aplicațiilor de ILN/PLN; § utilizeze corect termenii specifici modelelor de natură simbolică, statistică, neurală; § adapteze și să implementeze modele simbolice, statistice și neurale pe baza unor specificații date pentru diverse probleme reale din domeniul ILN/PLN; § analizeze, modeleze și evalueze soluția unei aplicații informatice simple din domeniul ILN/PLN.

8. Conținut

	Curs	Metode de predare	Observații (ore și referințe bibliografice)
1	Niveluri de prelucrare ale limbajului natural (PLN); prelucrări de bază ale documentelor în vederea implementării de modele	Expunere, prelegere participativă	2
2	Modele de proiecte PLN bazate pe machine learning	Expunere, prelegere participativă, exemplificare, aplicații	2
3	Modele de proiecte PLN bazate pe deep learning	Expunere, prelegere participativă, exemplificare, aplicații	2



4	Prelucrările nivelului sintactic	Expunere, prelegere participativă, exemplificare, aplicații	2
5	Prelucrări la nivelul semantic	Expunere, prelegere participativă, exemplificare, aplicații	2
6	Prelucrări la nivelul pragmatic	Expunere, prelegere participativă, exemplificare, aplicații	2
7	Relații semantice. Anafora, fenomene anaforice și rezoluția ei	Expunere, prelegere participativă, exemplificare, aplicații	2
8	Nivelul discursului: Teoria Structurilor Retorice, Teoria Nervurilor (coeziune)	Expunere, prelegere participativă, exemplificare, aplicații	2
9	Nivelul discursului: Teoria Nervurilor (coerență) și rezumare automată	Expunere, prelegere participativă, exemplificare, aplicații	2
10	Timpul în texte: logica temporală, modelul TIMEML, modelul Time Yards	Expunere, prelegere participativă, exemplificare, aplicații	2
11	Sisteme de întrebare răspuns; evaluarea sistemelor de PLN	Expunere, prelegere participativă, exemplificare, aplicații	2
12	Nivelul pragmatic: cogniție condiționată de percepție, conceptualizări, formarea limbajului	Expunere, prelegere participativă, exemplificare, aplicații	2
13	Abordări statistice, vectorale și neurale	Expunere, prelegere participativă, exemplificare, aplicații	2
14	Sesiune de întrebări recapitulative	Dialog cu studenții pe marginea materiei predate	2

Bibliografie

- R. Mitkov (2005). The Oxford Handbook of Computational Linguistics.
- R. Simionescu (2016). Statistical and Inference Based Approaches in Natural Language Processing – Applications to Morphology and Syntax, PhD thesis, UAIC.
- Chris Manning and Hinrich Schütze (1999). Foundations of Statistical Natural Language Processing, MIT Press. Cambridge, MA, <https://nlp.stanford.edu/fsnlp/>
- Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean (2014). Distributed Representations of Words and Phrases and their Compositionality: <https://arxiv.org/pdf/1310.4546.pdf>
- Chris Nicholson and Adam Gibson (2014). SkyMind: A Beginner's Guide to Word2Vec and Neural Word Embeddings: <https://skymind.ai/wiki/word2vec>
- TensorFlow: An open-source machine learning framework for everyone: <https://www.tensorflow.org/>
- Natural Language Processing Fundamentals in Python, Natural Language ToolKit, de ex. <http://nltk.org/book>.
- Site-ul NLP-Group@UAIC-FII, la <http://nlptools.info.uaic.ro/>

Seminar

Seminarele/laboratoarele urmează rezolvarea unui set de proiecte. Studenților li se indică pachete software și aplicații web cu care să construiască miniaplicații. Tehnologiile prezentate sunt exemplificate în aplicații de tipul: asistenți virtuali, regăsire de informații, analiza textelor din psihologie, prelucrarea vorbirii etc.



9. Coroborarea conținutului disciplinei cu așteptările reprezentanților comunității, asociațiilor profesionale și angajatorilor reprezentativi din domeniul aferent programului

Conținutul cursului este actualizat pe baza unor cursuri similare de la diferite universități de prestigiu din străinătate. Ingineria limbajului este considerat astăzi unul dintre domeniile cheie de dezvoltare în tehnologia informației. Corporații internaționale mari precum Google, Microsoft, IBM, Facebook și Nuance au investit substanțial în acest domeniu. Peste tot în lume, mii de start-up-uri se înființează pentru diverse aplicații și servicii de prelucrare a limbajului natural, esențiale pentru extragerea de informații din colecții imense de date structurate sau nestructurate. Tehnologia limbajului le permite oamenilor să colaboreze, să învețe, să facă afaceri și să împărtășească cunoașterea dincolo de granițele limbajului, indiferent de competențele de utilizare a calculatorului.

10. Evaluare

10.1 Tip activitate	10.2 Criterii de evaluare	10.3 Metode de evaluare	10.4 Pondere în nota finală (%)
Curs	- Corectitudinea și completitudinea cunoștințelor: - Capacitatea de a înțelege și explica modele simbolice, statistice și neurale aplicate limbajului natural. - Utilizarea corectă a terminologiei domeniului - Creativitate	Teste practice (evaluare parțială și evaluare finală online), proiecte	33%
Seminar/laborator	- Abilitatea de a construi aplicații de analiză a textelor folosind diferite pachete software.	Exerciții de seminar/laborator	33%
		Proiect de cercetare	33%
	Criterii ce vizează aspectele atitudinale: - Interesul pentru studiu individual - Interactivitatea, prezența activă, la curs și seminar	Participarea activă la cursuri și seminarii/laboratoare	bonus 10%

11. Standard minim de performanță

Pentru promovare trebuie îndeplinite simultan următoarele criterii:

- § Utilizarea corectă a terminologiei domeniului (testată în exprimările din teză);
- § Abilitatea de a explica un model simbolic, statistic ori neural aplicat PLN;
- § Capacitatea de a elabora un model simbolic, statistic ori neural pentru o problemă reală de PLN;
- § Minim 50% prezență la curs & seminar/laborator;
- § Minim 30% teme de la seminar/laborator;
- § Minim 20% proiect.

Notele se stabilesc conform cu criteriile ECTS.

Data completării
30 martie 2021

Titular de curs
CS II Dr. Daniela Gîfu

Titular de seminar
CS II Dr. Daniela Gîfu
Conf. Dr. Diana Trandabăț
Conf. Dr. Ionuț Pistol
Dr. Mihaela Onofrei

Data avizării în departament

Director de departament
Prof. Dr. Dorel LUCANU