

PROCESAREA INFORMAȚIEI

1. Date despre unitatea de curs/modul

Facultatea	Calculatoare, Informatică și Microelectronică				
Catedra/departamentul	Informatica și Ingineria Sistemelor				
Ciclul de studii	Studii superioare de licență, ciclul I				
Programul de studiu	0612.2 Managementul Informației				
Anul de studiu	Semestrul	Tip de evaluare	Categoria formativă	Categoria de opționalitate	Credite ECTS
II (învățământ cu frecvență);	3	E	F – unitate de curs fundamentală	O - unitate de curs obligatorie	3

2. Timpul total estimat

Total ore în planul de învățământ	Din care				
	Ore auditoriale		Lucrul individual		
	Curs	Laborator	Proiect de an	Studiul materialului teoretic	Pregătire aplicații
90	30	15		20	25

3. Precondiții de acces la unitatea de curs/modul

Conform planului de învățământ	Limbaje Formale, Programarea Calculatoarelor, Programarea Interactivă
Conform competențelor	Tehnici specifice de procesare a informației

4. Condiții de desfășurare a procesului educațional pentru

Curs	Pentru prezentarea materialului teoretic în sala de curs este nevoie de proiector și calculator. Nu vor fi tolerate întârzierile studenților, precum și convorbirile telefonice în timpul cursului.
Laborator	Studenții vor perfecta rapoarte conform condițiilor impuse de indicațiile metodice. Termenul de predare a lucrării de laborator – o săptămână după finalizarea acesteia. Pentru predarea cu întârziere a lucrării aceasta se depuncea cu 1pct./săptămână de întârziere.

5. Competențe specifice acumulate

Competențe profesionale	<p>CP2. Utilizarea sistemelor informatice de prelucrare și gestiune a datelor</p> <ul style="list-style-type: none"> ✓ Descrierea structurii și a modului de funcționare a sistemelor informatice în general. ✓ Explicarea rolului, funcționalității și utilității sistemelor informatice în general și a sistemelor de prelucrare și gestiune a datelor în domeniul specializării. ✓ Utilizarea componentelor software ale sistemelor informatice, folosind algoritmi, protocoale, limbaje, structuri de date. ✓ Utilizarea adecvată de criterii și metode standard de apreciere a caracteristicilor și a calităților sistemelor informațional. ✓ Prelucrarea și gestionarea datelor utilizând sisteme informatice dedicate
Competențe profesionale	<p>CP4. Rezolvarea problemelor economice și inginerești folosind metode matematice, metode statistice și tehnici informatice</p> <ul style="list-style-type: none"> ✓ Identificarea problemelor economice și inginerești care se pretează la modelare matematică și statistică precum și a metodelor utilizabile din informatica aplicată .

	<ul style="list-style-type: none"> ✓ Însușirea cunoștințelor teoretice și a deprinderilor de actualizare a acestora, necesare în procesul de interpretare a rezultatelor experimentale ✓ Abilitatea de a utiliza și adapta șabloane de soluții specifice matematicii și informaticii aplicate pentru rezolvarea problemelor economice și inginerești. ✓ Utilizarea criteriilor și metodelor de evaluare pentru adoptarea procedeeor, tehnicilor și metodelor de bază, necesare creșterii performanțelor. ✓ Analiza cerințelor pieței și tendințelor contemporane privind dezvoltarea unor soluții la problemele economice și inginerești folosind principii și metode ale matematicii, statisticii și informaticii aplicate
Competențe transversale	<p>CT1. Realizarea proiectului de an cu utilizarea corectă a surselor bibliografice și metodelor specifice, în condiții de autonomie restrânsă și asistență calificată, precum și susținerea acestora cu demonstrarea capacității de evaluare calitativă și cantitativă a unor soluții tehnice din domeniu.</p> <p>CT3. Identificarea nevoii de formare profesională, cu analiza critică a propriei activități de formare și a nivelului de dezvoltare profesională și utilizarea eficientă a resurselor de comunicare și formare profesională (Internet, e-mail, baze de date, cursuri on-line etc.), inclusiv folosind limbi străine.</p>

6. Obiectivele unității de curs/modulului

Obiectivul general	Familiarizarea, inițierea și instruirea studenților cum să utilizeze în practică conceptul modern de „Procesarea informației nestructurate”(PIN), și să utilizeze procesarea structurilor lingvistice a textului în dezvoltarea sistemelor moderne existente.
Obiectivele specifice	<p>Să înțeleagă conceptul de analiză sintactică a textului și proiectarea dicționarelor (uni - și bilingve).</p> <p>Să proiecteze reguli sintactico-semantice și gramatici formale pentru sistemele de înțelegere a textului limbajului natural.</p> <p>Să înțeleagă conceptul de analiză probabilistico-statistică a textului și proiectarea modelelor stohastice pentru rezolvarea problemelor PDN.</p> <p>Să programeze în limbajul Python în cadrul elaborării softului în baza algoritmilor probabilistico-statistice: Top-Down, Bottom-Up, Early, Word Sense Disambiguation și Viterbi.</p>

7. Conținutul unității de curs/modulului

Tematica activităților didactice	Numărul de ore	
	învățământ cu frecvență	învățământ cu frecvență redusă
Tematica prelegerilor		
T1. Introducere. Lingvistica și Procesarea datelor nestructurate. Metode empirice de construire a dicționarului electronic cu codarea (taging) pentru analiza sintactică. Metode empirice pentru parsing morfologic- etapa inițială a analizei sintactice a propoziției.	4	
T2. Tehnici de parsare. Top-Down Parser. Bottom-Up Parser. Top-Down Chart parsing. Transition Network Grammars. Algoritmul Early Algorithm.	4	
T3. Gramatici Probabilistice libere de context. Parsare cu PCFG. Algoritmul Viterbi. Algoritmul CYK	4	
T4. Rezolvarea ambiguității. Metode statistice. Dezambiguizarea sensului cuvintelor. Dezambiguizarea supervizată.	4	
T5. Regăsirea informației. Tehnici specifice. Metode.	4	

T6. Extragerea informației. Tehnici specifice de extragere a informației.	4	
T7. Metode de Clasificare și Clusterizare a informației.	6	
Total prelegeri:	30	

Tematica activităților didactice	Numărul de ore	
	învățământ cu frecvență	învățământ cu frecvență redusă
Tematica lucrărilor de laborator/seminarelor		
LL1. Metode empirice de construire a dicționarului electronic cu codificarea (taging) pentru analiza sintactică.	4	
LL2. Aplicarea tehnicilor Top-Down și Bottom-up chart parser pentru elaborarea unei aplicații.	4	
LL3. Implementarea algoritmului Early. Testarea modulelor softului creat	4	
LL4. Aplicarea tehnicilor studiate pentru căutarea și extragerea informațiilor din texte date	3	
Total lucrări de laborator/seminare:	15	

8. Referințe bibliografice

Principale	<ol style="list-style-type: none"> 1. Natural Language Processing With Python, Steven Bird, Ewan Klein, Edward Loper, Ed. O'Reilly. 2009, ISBN: 978-0-596-51649-9 2. Clive Matthews. An Introduction to Natural Language Processing Through Prolog. Logman 1998. 3. Christopher D.Manning and Hinrich Schutze.Foundation of Statistical Natural Language Processing. Cambridge 2002 4. Roman Feldman, James Sauger. The Text Mining Handbook. Advanced Approaches in Analyzing Unstructured Data. Cambridge 2008. 5. James Allen. Natural Language Understanding. 1995. 6. Mastering Natural Language Processing with Python Mastering, Deepti Chopra, Nisheeth Joshi, Iti Mathur, 2016, ISBN 139781783989041
Suplimentare	<ol style="list-style-type: none"> 1. http://www.nltk.org/book/ 2. https://www.kaggle.com/c/word2vec-nlp-tutorial/details/part-1-for-beginners-bag-of-words 3. https://blog.pusher.com/introduction-to-natural-language-processing-with-python/ 4. https://github.com/bonzanini/nlp-tutorial 5. https://www.dataquest.io/blog/natural-language-processing-with-python/

9. Evaluare

Forma de învățământ	Periodică		Curentă	Lucrul individual	Examen final
	Atestarea 1	Atestarea 2			
Cu frecvență	15%	15%	15%	15%	40%
Standard minim de performanță					
Prezența la lecții; activitatea și calitatea pregătirii la / pentru prelegeri și lucrări de laborator; Obținerea notei minime de „5” la fiecare dintre atestări și lucrări de laborator; Demonstrarea în lucrarea de examinare finală a cunoștințelor minime.					