

MD-2045, CHIȘINĂU, STR. STUDENȚILOR, 9/7, TEL: 022 50-99-01 | FAX: 022 50-99-05,

www.utm.md

ANALIZA STATISTICĂ A DATELOR CU APLICAȚII

1. Date despre unitatea de curs/modul

Facultatea	Calculatoare, Informatică și Microelectronică				
Catedra/departamentul	Informatică și Ingineria Sistemelor				
Ciclul de studii	Studii de masterat, ciclul II				
Programul de studiu	Managementul Aplicațiilor Informaționale				
Anul de studiu	Semestrul	Tip de evaluare	Categoria formativă	Categoria de opționalitate	Credite ECTS
I (învățământ cu frecvență);	2	E	F – unitate de curs de fundamentală	O - unitate de curs obligatorie	5

2. Timpul total estimat

Total ore în planul de învățământ	Din care				
	Ore auditoriale		Lucrul individual		
	Curs	Laborator/seminar	Proiect de an	Studiul materialului teoretic	Pregătire aplicații
150	20	20	-	55	55

3. Precondiții de acces la unitatea de curs/modul

Conform planului de învățământ	Matematici speciale, Matematica computațională, Modele matematice și optimizări
Conform competențelor	Matematica, algoritmizarea problemelor, logica matematică, bazele statisticii.

4. Condiții de desfășurare a procesului educațional pentru

Curs	Pentru prezentarea materialului teoretic în sala de curs este nevoie de proiector și calculator. La fel este de dorit și posibilitatea de a conecta calculatorul la internet pe parcursul lecțiilor.
Laborator/seminar	Studentii vor perfecta lucrările în laborator conform condițiilor și variantelor impuse de indicațiile metodice și vor prezenta profesorului în format electronic. Termenul de predare a lucrării de laborator – două săptămâni după finalizarea acesteia. Pentru predarea cu întârziere a lucrării aceasta se depunțează cu 1pct./două săptămâni de întârziere.

5. Competențe specifice acumulate

Competențe profesionale	CP1. Operarea adecvată cu conceptele fundamentale ale științelor exacte, informaticii aplicate și științei calculatoarelor. CP1.1 Identificarea conceptelor, principiilor, paradigmelor, metodelor și tehnicilor de descriere, modelare, verificare și implementare a aplicațiilor software sistemelor de calcul și rețelelor de calculatoare. CP1.2 Utilizarea de teorii și instrumente specifice domeniului (algoritmi, metode, tehnici, protocoale, etc.) pentru explicarea funcționării aplicațiilor
-------------------------	---

	software inteligente, sistemelor de calcul și rețelelor de calculatoare.
Competențe profesionale	CP2. Conceperea, proiectarea, implementarea și gestiunea sistemelor informaționale complexe inteligente CP2.5 Proiectarea și implementarea compo-nentelor, aplicațiilor software, sistemelor de calcul de uz general și dedicate și rețelelor de calculatoare.
Competențe profesionale	CP3. Identificarea, formularea și soluționarea problemelor folosind instrumentele științei și ingineriei calculatoarelor CP3.1 Recunoașterea și descrierea unor tehnici și metode de rezolvare a sarcinilor de sinteză, modelare, simulare, verificare și implementare a aplicațiilor software, echipamentelor, sistemelor de calcul și obiectelor inteligente. CP3.2 Utilizarea adecvată a cunoștințelor interdisciplinare, a metodelor de soluționare și a mediilor de dezvoltare, efectuarea experimentelor și interpretarea rezultatelor.
Competențe profesionale	CP4. Managementul proceselor de elaborare a sistemelor informatice și asigurarea calității produselor și a serviciilor CP4.5 Dezvoltarea de aplicații pentru monitorizarea și controlul parametrilor de bază a sistemelor informatice pentru îmbunătățirea și optimizarea performanțelor.
Competențe profesionale	CP5. Utilizarea de metode avansate de modelare, simulare, identificare și analiză a sistemelor software, a fenomenelor și proceselor din domeniul sistemelor inteligente informatice CP5.1 Descrierea instrumentelor de modelare, simulare, verificare și evaluare a performanțelor sistemelor software inteligente, hardware și de comunicații. CP5.4 Testarea, validarea și evaluarea calitativă a caracteristicilor comportamentale și structurale ale aplicațiilor informaționale, obiectelor inteligente sistemelor de calcul și rețelelor de calculatoare.
Competențe profesionale	CP6. Cercetare științifică în domeniul științei tehnologiei informației și comunicațiilor; Managementul produselor și al serviciilor IT în concordanță cu cerințele pieței. CP6.1 Descrierea procedeeleor, tehnicilor și metodelor de bază necesare în domeniul științei calculatoarelor, tehnologiei informației și comunicațiilor.
Competențe transversale	CT1. Comportarea onorabilă, responsabilă, etică, în spiritul legii pentru a asigura îndeplinirea sarcinilor profesionale.
Competențe transversale	CT2. Demonstrarea capacității de lucru în echipă, identificarea rolurilor și responsabilităților individuale și comune, luarea deciziilor și atribuirea de sarcini, cu aplicarea de tehnici de relaționare și muncă eficientă în cadrul echipei.
	CT3. Demonstrarea spiritului de inițiativă și acțiune pentru dezvoltarea profesională și personală, prin formare continuă utilizând surse de documentare în limba română și în limbile de circulație internațională.

6. Obiectivele unității de curs/modulului

Obiectivul general	Inițierea în procesul de analiză statistică a datelor, data mining și învățare automată.
Obiectivele specifice	Inițierea în teoriile de analiză statistică a datelor, clasificarea, reprezentarea în forma grafică, modelarea și prognozarea trendurilor.

	Studierea modelelor probabilistice pentru datele statistice, data mining, observarea dependențelor între datele contigue. Studierea metodelor matematice din teoria probabilităților utilizate pentru luarea deciziilor în baza datelor statistice acumulate. Studierea metodelor de învățare automată, instrumentelor ce pot fi utilizate în scopul acesta, preprocesarea datelor și interpretarea rezultatelor obținute.
--	--

7. Conținutul unității de curs/modulului

Tematica activităților didactice	Numărul de ore	
	învățământ cu frecvență	
Tematica prelegerilor		
T1. Noțiunea de statistică, analiza statistică a datelor. Metode de analiză statistică a datelor. Vizualizarea grafică a datelor, compararea și prognozarea lor, estimarea legii de repartiție a datelor statistice, sistematizarea datelor după clase de valori; prelucrarea numerică a datelor.	4	
T2. Metoda ipotezelor statistici.	2	
T3. Modelarea secvențelor de evenimente stohastice, lanțul Markov. Google Page Rank în baza lanțului Markov.	4	
T4. Teoria lui Shannon, teoria informației, entropia, entropia condiționată. Entropia și redundanța textului. Estimarea probabilităților, legea Zipf, problema frecvențelor mici și frecvenței egale cu zero, metode de netezire.	2	
T5. Învățarea automată, metode și algoritmi de învățare automată.	4	
T6. Arbori de decizie, construirea arborelor optimi	2	
T7. Metoda k vecinilor celor mai apropiați	2	
Total prelegeri:	20	

Tematica activităților didactice	Numărul de ore	
	învățământ cu frecvență	
Tematica seminarelor		
S1. Analiză statistică a datelor; vizualizarea grafică a datelor, compararea și prognozarea lor. Prelucrarea numerică a datelor.	2	
S2. Testarea ipotezelor statistice	2	
S3. Lanțul Markov, calculul Google Page Rank în baza lanțului Markov.	2	
S4. WEKA – un soft ce conține algoritmi de învățare automată.	4	
S5. Pregătirea și preprocesarea datelor de antrenare și testare pentru metodele de învățare automată.	4	
S6. Procesarea datelor cu algoritmi WEKA.	6	
Total seminare:	20	

8. Referințe bibliografice

Principal e	1. Lucia Căbulea, Nicoleta Breaz. INTERPRETAREA STATISTICĂ A INFORMAȚIILOR. ELEMENTE DE DATA MINING ȘI PROGNOZĂ, Universitatea „1 Decembrie 1918” Alba Iulia.
----------------	---

	http://www.uab.ro/cursuri_perfectionare/program_perfectionare_profesionala/pagini/cursuri/Modul-7-Cabulea-Breaz/curs_modul7.doc 2. T. Andrei, Statistică și econometrie, Ed. Economică, 2003. 3. C. Anghelache, E. Bugudui, S. Gresoi, E. Niculescu, Statistică aplicată-indicatori, sinteze și studii de caz, Ed. Economică, 2006 4. E.N. Bâzdoacă, S. Matei, N.G. Bâzdoacă, Inițiere în Excel, Ed.Arves, Colecția Inițiere în calculator 2, 2002 5. Liviu Ciortuz , Alina Munteanu , Elena Badarau Exercitii de invatare automata ed UNIVERSITATII "ALEXANDRU IOAN CUZA" (2015) ISBN: 606-714-197-9 6. MARCUS, S., NICOLAU, E., STATI, S., 1966. Introducere în lingvistica matematică. București, Editura științifică, p.336. 7. An estimate of an upper bound for the entropy of English. Brown, Della Pietra, Mercer, Della Pietra, Lai. Computational Linguistics, 18(1), pp31-40, 1992. 8. VLAD, A., MITREA, A., MITREA, M., 2005. Limba română scrisă ca sursă de informație. Paideia, România, 286 p. 9. Roman Feldman, James Sauger. The Text Mining Handbook. Advanced Approaches in Analyzing Unstructured Data. Cambridge 2008.
Suplimentare	- MIHALCEA, R., 2002. Diacritics Restoration: Learning from Letters versus Learning from Words. in Proceedings of the Third International Conference on Intelligent Text Processing and Computational Linguistics, Mexico, p.339-348. - Christopher D.Manning, Hinrich Schutze Foundations of Statistical Natural Language Processing Mit Press, 1999. - Ronald Rosenfeld. Two decades of Statistical Language Modeling: Where Do We Go From Here? Proceedings of the IEEE, 88(8), 2000. - Lillian Lee. "I'm sorry Dave, I'm afraid I can't do that": Linguistics, Statistics, and Natural Language Processing. Computer Science: Reflections on the Field, Reflections from the Field, pp. 111--118, 2004.

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 și activitatea la prelegeri și lucrări de laborator					
Obținerea notei minime de „5” la fiecare dintre evaluări și lucrări de laborator					