

ANALIZA ȘI VIZUALIZAREA DATELOR

1. Date despre disciplină/modul

Facultatea	Calculatoare, Informatică și Microelectronică				
Catedra/departamentul	Informatică și Ingineria Sistemelor				
Ciclul de studii	Studii superioare de licență, ciclul II				
Programul de studiu	Știința Datelor				
Anul de studiu	Semestrul	Tip de evaluare	Categoria formativă	Categoria de opționalitate	Credite ECTS
II (învățământ cu frecvență)	3	E	F – unitate de curs fundamental	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	-	80	30

3. Precondiții de acces la disciplină/modul

Conform planului de învățământ	Analiza Matematică, Probabilitate și statistică aplicată, Structuri de date și algoritmi, Matematica discretă, Programarea calculatoarelor, Programare Interactivă.
Conform competențelor	Utilizarea de teorii și instrumente specifice domeniului (algoritmi, metode, tehnici, etc.) pentru analiza algoritmilor de bază din statistică, analiză matematică și ilustrarea datelor grafic.

4. Condiții de desfășurare a procesului educațional pentru

Curs	Pentru prezentarea materialului teoretic în sala de curs este nevoie de tablă, proiector și calculator.
Laborator/seminar	Studentii vor perfecta rapoarte conform condițiilor impuse de indicațiile metodice. Termenul de predare a lucrării de laborator – 2 săptămâni după finalizarea acesteia. Pentru predarea cu întârziere a lucrării aceasta se depunează cu 1 pct./săptămână de întârziere. Pentru desfășurarea seminarelor în auditoriu de studii este nevoie de tablă, proiector și calculator.

5. Competențe specifice acumulate

Competențe profesionale	CPM1 Elaborarea și proiectarea arhitecturii sistemului CPM2. Monitorizarea tendințelor tehnologice. Inovație. Dezvoltarea durabilă. CPM3 Dezvoltarea aplicațiilor. Integrarea componentelor. Ingineria sistemelor. CPM5. Îmbunătățirea proceselor.
-------------------------	---

6. Obiectivele disciplinei/modulului

Obiectivul general	Să descrie conceptele de bază, Să determine rolul vizualizării datelor, Să determine instrumentele și tehnicile utilizate
Obiectivele specifice	Să cunoască măsuri de tendință centrală: medie, mediană, mod precum și măsuri de dispersie: varianță, deviație standard, interval interquartil Să definească noțiunea de regresie, Să poată evalua modelele de regresie

	Să poată utiliza principii de design pentru vizualizarea datelor, să folosească instrumente de vizualizare: Matplotlib, Seaborn, Plotly și să creeze grafice interactive
--	--

7. Conținutul disciplinei/modulului

Tematica activităților didactice	Numărul de ore învățământ cu frecvență
Tematica prelegerilor	
Tema 1. Introducere în Analiza Statistică și Vizualizarea Datelor.	2
Tema 2. Colectarea și Preprocesarea Datelor.	2
Tema 3. Statistici Descriptive	2
Tema 4. Probabilitate și Distribuții de Probabilitate	2
Tema 5. Analiza Regresiei	2
Tema 6. Analiza Factorială.	2
Tema 7. Analiza de Grupare (Clustering)	2
Tema 8. Tehnici de Vizualizare a Datelor.	2
Tema 9. Vizualizarea Datelor Geospațiale.	2
Tema 10. Proiecte de Vizualizare a Datelor.	2
Total prelegeri:	20

Tematica activităților didactice	Numărul de ore învățământ cu frecvență
Tematica lucrărilor de laborator/seminarelor	
Lucrarea practică nr 1. Preprocesarea și Explorarea Datelor	4
Lucrarea practică nr 2. Vizualizări Statistice de Bază	4
Lucrarea practică nr 3. Vizualizarea și Analiza Distribuțiilor de Probabilitate	4
Lucrarea practică nr 4. Proiect de Analiză de Regresie	4
Lucrarea practică nr 5. Proiect de Vizualizare Interactivă	4
Total lucrări de laborator/practice:	20

8. Referințe bibliografice

Principale	- 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 T. Andrei, Statistică și econometrie, Ed. Economică, 2003. C. Anghelache, E. Bugudui, S. Gresoi, E. Niculescu, Statistică aplicată-indicatori, sinteze și studii de caz, Ed. Economică, 2006 - E.N. Bâzdoacă, S. Matei, N.G. Bâzdoacă, Inițiere în Excel, Ed.Arves, Colecția Inițiere în calculator 2, 2002 - Liviu Ciortuz , Alina Munteanu , Elena Badarau Exercitii de invatare automata ed UNIVERSITATII "ALEXANDRU IOAN CUZA" (2015) ISBN: 606-714-197-9 - MARCUS, S., NICOLAU, E., STATI, S., 1966. Introducere în lingvistica matematică. București, Editura științifică, p.336. - 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. - VLAD, A., MITREA, A., MITREA, M., 2005. Limba română scrisă ca sursă de informație. Paideia, România, 286 p.
-------------------	---

	10. Roman Feldman, James Sauger. The Text Mining Handbook. Advanced Approaches in Analyzing Unstructured Data. Cambridge 2008.
Suplimentare	1. 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. 2. Christopher D.Manning, Hinrich Schutze Foundations of Statistical Natural Language Processing Mit Press, 1999. 3. Ronald Rosenfeld. Two decades of Statistical Language Modeling: Where Do We Go From Here? Proceedings of the IEEE, 88(8), 2000. 4. 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, seminare și lucrări de laborator; Obținerea notei minime de „5” la fiecare dintre atestări și lucrări de laborator; Demonstrarea în lucrarea de examinare finală a cunoașterii algoritmilor de bază din teoria limbajelor formale, automatelor finite și automatelor pushdown.					