



FIȘA DISCIPLINEI

1. Date despre program

1.1 Instituția de învățământ superior	Universitatea Națională de Știință și Tehnologie Politehnica București
1.2 Facultatea	Electronică, Telecomunicații și Tehnologia Informației
1.3 Departamentul	Telecomunicații
1.4 Domeniul de studii	Inginerie Electronică, Telecomunicații și Tehnologii Informaționale
1.5 Ciclul de studii	Masterat
1.6 Specializarea	Ingineria Informației și a Sistemelor de Calcul

2. Date despre disciplină

2.1 Denumirea disciplinei (ro)		Metode de explorare a datelor heterogene					
(en)							
2.2 Titularul activităților de curs		Conf. Dr. Daniela FAUR					
2.3 Titularul activităților de seminar / laborator		Conf. Dr. Daniela FAUR					
2.4 Anul de studiu	2	2.5 Semestrul	I	2.6. Tipul de evaluare	E	2.7 Regimul disciplinei	Ob
2.8 Tipul disciplinei	DA	2.9 Codul disciplinei	UPB.04.M3.O.16-15	2.10 Tipul de notare	Nota		

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

3.1 Număr de ore pe săptămână	3	Din care: 3.2 curs	2.00	3.3 seminar/laborator	1
3.4 Total ore din planul de învățământ	42.00	Din care: 3.5 curs	28	3.6 seminar/laborator	14
Distribuția fondului de timp:					ore
Studiul după manual, suport de curs, bibliografie și notițe					48
Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate					
Pregătire seminarii/ laboratoare/proiecte, teme, referate, portofolii și eseuri					
Tutorat					4
Examinări					6
Alte activități (dacă există):					0
3.7 Total ore studiu individual	58.00				
3.8 Total ore pe semestru	100				
3.9 Numărul de credite	4				

4. Precondiții (acolo unde este cazul)

4.1 de curriculum	Nu este cazul.
4.2 de rezultate ale învățării	Acumularea de cunoștințe prin parcurgerea următoarelor discipline: Structuri de date și algoritmi Programare obiect orientată Prelucrarea digitală a semnalelor Prelucrarea și analiza imaginilor Recunoașterea formelor și inteligența artificială



5. Condiții necesare pentru desfășurarea optimă a activităților didactice (acolo unde este cazul)

5.1 Curs	Sală dotată cu videoproiector și computer.
5.2 Seminar/ Laborator/Proiect	Nu este cazul.

6. Obiectiv general (*Se referă la intențiile profesorilor pentru studenți, la ceea ce studenții vor fi învățați în timpul cursului. Oferă o orientare cu privire la locul cursului în cadrul domeniului științific abordat, precum și la rolul pe care acesta îl are în cadrul specializării studiate. Vor fi descrise de o manieră generală tematicile abordate, justificarea includerii cursului în planul de învățământ al specializării studiate etc.*)

Analiza exploratorie a seturilor mari de date presupune aprofundarea unor metode și algoritmi aflați la intersecția unor domenii ca statistica, machine learning și baze de date. Obiectivul acestui proces este extragerea informației din seturi mari de date și transformarea acesteia într-un format inteligibil, familiar pentru utilizator. Cursul își propune să trateze această analiză considerând 2 categorii de date eterogene: datele satelitare și limbajul natural scris. Conținutul acestui curs urmărește dezvoltarea abilităților studenților de a aplica cunoștințe fundamentale în ingineria informației pentru rezolvarea unor probleme de procesare a informației legate de analiza automată și semi automată a bazelor mari de date pentru a extrage modele anterior necunoscute, similare din punct de vedere al conținutului (clase), esanțioane neobisnuite (detectie de anomalii) sau corelații (reguli de asociere).

7. Competențe (*Capacitatea dovedită de a utiliza cunoștințe, aptitudini și abilități personale, sociale și/sau metodologice în situații de muncă sau de studiu și pentru dezvoltarea profesională și personală. Reflectă cerințele angajatorilor.*)

Specifice	Aplicarea cunoștințelor fundamentale în ingineria informației: metode statistice, prelucrare avansată a semnalelor, analiză și prelucrarea imaginilor, sisteme de operare Proiectarea și realizarea de aplicații în sisteme informatice și baze de date (data mining, teoria și proiectarea bazelor de date, inclusiv distribuite, e-learning și e-activities, tehnologii multimedia);
Transversale (generale)	- Executarea responsabilă a unor sarcini de lucru în echipă pluridisciplinară, cu asumarea de roluri pe diferite paliere ierarhice; - Identificarea nevoii de formare continuă și utilizarea eficientă a surselor informaționale și a resurselor de comunicare și formare profesională asistată (portaluri Internet, aplicații software de specialitate, baze de date, cursuri on-line, etc.) atât în limba română cât și într-o limbă de circulație internațională. - Comunicarea orală și scrisă în limba engleză, demonstrarea intelgenței termenilor specifici domeniului și prezentarea rezultatelor de cercetare sub forma unui articol științific.

8. Rezultatele învățării (*Sunt enunțuri sintetice referitoare la ceea ce un student va fi capabil să facă sau să demonstreze la finalizarea unui curs. Rezultatele învățării reflectă realizările studentului și mai puțin intențiile profesorului. Rezultatele învățării informează studenții despre ceea ce se așteaptă de la ei din punct de vedere al performanței, pentru a obține notele și creditele dorite. Sunt definite în termeni concreți, folosind verbe similare exemplurilor de mai jos și indică ceea ce se va urmări prin evaluare. Rezultatele învățării vor fi astfel redactate încât să fie evidențiată clar relația față de competențele definite la punctul 7.)*



Cunoștințe	<p>Rezultatul asimilării de informații prin învățare. Cunoștințele reprezintă ansamblul de fapte, principii, teorii și practici legate de un anumit domeniu de muncă sau de studiu. Pot fi teoretice și/sau faptice.</p> <p>Inteleg și descrie concepte importante care au marcat dezvoltarea domeniului.</p> <p>Definește noțiuni specifice domeniului.</p> <p>Descrie/clasifică noțiuni/procese/fenomene/structuri.</p>
Aptitudini	<p>Capacitatea de a aplica cunoștințe și de a utiliza know-how pentru a duce la îndeplinire sarcini și a rezolva probleme. Aptitudinile sunt descrise ca fiind cognitive (implicând utilizarea gândirii logice, intuitive și creative) sau practice (implicând dexteritate manuală și utilizarea de metode, materiale, unelte și instrumente).</p> <p>Selectează și grupează informații relevante – state of the art in domeniul temei de proiect.</p> <p>Identifica metode de implementare și optimizare pentru o temă dată.</p> <p>Verifica experimental soluțiile, analizează, compară și interpretează rezultatele.</p> <p>Elaborează un articol științific parcurgând etape impuse: motivația temei, analiza stadiului actual, descrierea setului de date, implementarea metodei propuse și interpretarea rezultatelor</p>
Responsabilitate și autonomie	<p>Capacitatea cursantului de a aplica în mod autonom și responsabil cunoștințele și aptitudinile sale.</p> <p>Selectează surse bibliografice din domeniul proiectului alocat, le analizează și face corelații.</p> <p>Respectă principiile de etică academică, citând corect sursele bibliografice utilizate.</p> <p>Demonstrează receptivitate pentru contexte noi de învățare</p> <p>Manifestă colaborare cu ceilalți colegi și cadre didactice în desfășurarea activităților didactice</p> <p>Demonstrează autonomie în organizarea situației/contextului de învățare sau a proiectului.</p>

9. Metode de predare (Se vor avea în vedere metode care să asigure predarea centrată pe student. Se va descrie modul în care se asigură participarea studenților la stabilirea propriului parcurs de învățare, cum se identifică eventualele rămăneri în urmă și ce măsuri remediale se adoptă în astfel de cazuri.)

Procesul de predare va îmbina expunerea și conversațiile interactive, problematizarea. Pentru prezentarea aspectelor teoretice se vor folosi mijloace multimedia – smart board. Notiunile teoretice vor fi însoțite de numeroase exemple/aplicații practice/ filmulete ce susțin și fundamentează conceptele prezentate. Indrumarea studenților pentru realizarea aplicațiilor practice ce reprezintă teme de proiect va încuraja gândirea critică și feedbackul permanent. Prezentarea proiectului impune exersarea abilităților de comunicare, de sumarizare precum și ascultarea activă a celorlalte teme de proiect și dezvoltarea mecanismelor de construcție a feedbackului.

10. Conținuturi

CURS		
Capitolul	Conținutul	Nr. ore
1	I. Introducere. Metode de explorare, date eterogene, analiza exploratorie a datelor. (Definții)	2
2	I. Introducere. Metode de explorare, date eterogene, analiza exploratorie a datelor. (Definții)	14



3	II. Regasirea si extragerea informatiei din date satelitare. - Imagini satelitare (notiuni fundamentale, metode de preprocesare, extragerea trasaturilor); - arbori MST, descompunerea liniara a unei imagini, algoritmi de clasificare si segmentare a datelor satelitare; - Determinarea automata a claselor de obiecte; - Extragerea vectorilor de contur pentru reconstrucia obiectelor. - Algoritmi de inteligenta artificiala (algoritmi genetici, retele neuronale si analiza fuzzy) pentru data mining in datele satelitare. - Metode de visual analytics pentru explorarea bazelor de date satelitare, reducerea dimensionalitatii si generarea proiectiilor 3D	12
4	III. Regasirea si extrageea informatiei din limbajul natural scris -semantic Web, extragerea automata a textului principal de pe o pagina web; recunoasterea limbii;	0
5	III. Regasirea si extrageea informatiei din limbajul natural scris -semantic Web, extragerea automata a textului principal de pe o pagina web; recunoasterea limbii;	0
Total:		

Bibliografie:

1. Catalin Cucu Dumitrescu - Elemente de data mining. Tehnici si aplicatii pentru imagini satelitare si analiza limbajelor natural, Editura Matrix Rom, ISBN: 978-973-755-756-8
2. Fundamentals of remote sensing - <http://www.nrcan.gc.ca/node/9309>
3. Roger D. Peng- Exploratory Data Analysis with R
4. Wang, Lidong. (2017). Heterogeneous Data and Big Data Analytics. Automatic Control and Information Sciences. 3. 8-15. 10.12691/acis-3-1-3.
5. Salama M, Kader HA, Abdelwahab A. An analytic framework for enhancing the performance of big heterogeneous data analysis. International Journal of Engineering Business Management. 2021;13. doi:10.1177/184797902199052
6. Allen B. Downey - Think stats, 2015 - Practical tools of exploratory data analysis.
7. Y. Guo, S. Guo, Z. Jin, S. Kaul, D. Gotz and N. Cao, "Survey on Visual Analysis of Event Sequence Data," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 12, pp. 5091-5112, 1 Dec. 2022, doi: 10.1109/TVCG.2021.3100413
8. J. Wang *et al.*, "Heterogeneous Data Feature Extraction Technology based on Long-Short Term Memory," *2020 IEEE International Conference on Artificial Intelligence and Information Systems (ICAIS)*, Dalian, China, 2020, pp. 141-144, doi: 10.1109/ICAIS49377.2020.9194865.
9. J. Liu *et al.*, "Urban Flow Pattern Mining Based on Multi-Source Heterogeneous Data Fusion and Knowledge Graph Embedding," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 2, pp. 2133-2146, 1 Feb. 2023, doi: 10.1109/TKDE.2021.3098612
10. Y. Yilmaz, M. Aktukmak and A. O. Hero, "Multimodal Data Fusion in High-Dimensional Heterogeneous Datasets Via Generative Models," in *IEEE Transactions on Signal Processing*, vol. 69, pp. 5175-5188, 2021, doi: 10.1109/TSP.2021.3109375.
11. S. Singh and M. Siwach, "Handling Heterogeneous Data in Knowledge Graphs: A Survey," in *Journal of Web Engineering*, vol. 21, no. 4, pp. 1145-1186, June 2022, doi: 10.13052/jwe1540-9589.2147

PROIECT

Nr. crt.	Conținutul	Nr. ore
----------	------------	---------



1	Implementarea unei metode de vizualizare a datelor eterogene prin parcurgerea unor etape impuse: identificarea descriptorilor/trasaturilor in corelatie cu natura datelor, extragerea trasaturilor, reducerea dimensionalitati spatiului multidimensional prin implementarea unei metode impuse, analiza si interpretarea rezultatelor	14
		Total:
Bibliografie:		

11. Evaluare

Tip activitate	11.1 Criterii de evaluare	11.2 Metode de evaluare	11.3 Pondere din nota finală
11.4 Curs	Cunoasterea si intelegerea notiunilor teoretice fundamentale;	Examen final, scris si oral, in timpul sesiunii, fixat de comun acord cu studentii.	60%
	Cunoasterea si intelegerea notiunilor teoretice fundamentale;	Examen final, scris si oral, in timpul sesiunii, fixat de comun acord cu studentii.	
11.5 Seminar/laborator/proiect	Cunoasterea, intelegerea si implementarea algoritmilor de extragere de trasaturi, clasificare si vizualizare pentru dezvoltarea unor aplicatii	Prezentare orala la sfarsitul semestrului; Verificarea aplicatie.	40%
	%		
11.6 Condiții de promovare			
- înțelegerea cunoștințelor teoretice fundamentale și aplicarea lor pentru rezolvarea unor aplicații practice; - proiectarea unor sisteme de extragere a informației bazată pe conținut din seturi mari de date (imagini, text) - 50% din punctajul total			

12. Coroborarea conținutului disciplinei cu așteptările reprezentanților angajatorilor și asociațiilor profesionale reprezentative din domeniul aferent programului, precum și cu stadiul actual al cunoașterii în domeniul științific abordat și practicile în instituții de învățământ superior din Spațiul European al Învățământului Superior (SEİS)

Detinerea informației relevante la momentul ideal este de importanță crucială pentru luarea deciziilor. Datorită progresului tehnologic intens cantitatea de informație de interes pentru luarea unei decizii crește foarte rapid în orice domeniu. Unul din motive este fluxul în continuă creștere al datelor ca urmare a automatizării activităților în domenii ca afaceri, inginerie, știință sau guvernare. Seturile mari de date reprezintă o importantă resursă nu întotdeauna ușor de valorificat din cauza dificultății de a extrage informația relevantă pentru un anumit context. Conținutul cursului „Metode de explorare a datelor eterogene” răspunde în mod concret și eficient cerințelor actuale de dezvoltare și evoluție, cu atât mai mult cu cât ne aflăm în plină eră „Big Data”. Orele de curs asigură însușirea cunoștințelor teoretice relevante



Universitatea Națională de Știință și Tehnologie Politehnica București

**Facultatea de Electronică, Telecomunicații și
Tehnologia Informației**



pentru extragerea informației din date eterogene iar orele dedicate proiectului urmaresc pregatirea studentilor pentru realizarea unor aplicatii de extragere a informației bazata de continut stimuland asumarea responsabilitatii in echipa si utilizarea eficienta a resurselor informationale.

Data completării

Titular de curs

Titular(i) de aplicații

Conf. Dr. Daniela FAUR

Conf. Dr. Daniela FAUR

Data avizării în departament

Director de departament

29.10.2024

Conf. Dr. Bogdan Cristian FLOREA

Data aprobării în Consiliul Facultății

Decan

29.10.2024

Prof. Dr. Mihnea Udrea