

FIȘA DISCIPLINEI¹⁾

1. Date despre program

1.1. Instituția de învățământ superior	Universitatea Petrol-Gaze din Ploiești
1.2. Facultatea	Litere si Științe
1.3. Departamentul	Informatică, Tehnologia Informației, Matematică, Fizică
1.4. Domeniul de studii universitare	Informatică
1.5. Ciclul de studii universitare	Master
1.6. Programul de studii universitare	Tehnologii Avansate pentru Prelucrarea Informației

2. Date despre disciplină

2.1. Denumirea disciplinei	Tehnici avansate de data mining
2.2. Titularul activităților de curs	Conf.dr. mat. Marinoiu Cristian
2.3. Titularul activităților seminar/laborator	Șef lucrări dr. Ioniță Irina
2.4. Titularul activității proiect	-
2.5. Anul de studiu	II
2.6. Semestrul*	3
2.7. Tipul de evaluare	Examen
2.8. Categoria formativă** / regimul*** disciplinei	DS/O

*numărul semestrului este conform planului de învățământ;

**DF - Discipline fundamentale; DD - discipline de domeniu; DS - discipline de specialitate; DC - discipline complementare, DA - disciplina de aprofundare, DSI- disciplina de sinteza.

***obligatorie = O; opțională = A; facultativă = L

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

3.1. Număr de ore pe săptămână	3	din care: 3.2. curs	2	3.3. Seminar/laborator	1	3.4. Proiect	-
3.5. Total ore din planul de învățământ	42	din care: 3.6. curs	28	3.7. Seminar/laborator	14	3.8. Proiect	-
3.9. Distribuția fondului de timp							ore
Studiu după manual, suport de curs, bibliografie și notițe							30
Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate și pe teren							44
Pregătire seminarii/laboratoare, teme, referate, portofolii și eseuri							42
Tutoriat							14
Examinări							3
Alte activități							-
3.10 Total ore studiu individual	133						
3.11. Total ore pe semestru	42						
3.12. Numărul de credite	7						

4. Precondiții (acolo unde este cazul)

4.1. de curriculum	➤ Învățare Automată, Analiza datelor
4.2. de competențe	➤ Abilități de programare, de calculul probabilităților și de calcul statistic

¹⁾ Adaptare după Ordinul Ministrului educației, cercetării, tineretului și sportului nr. 5 703/2011 privind implementarea Codului național al calificărilor din învățământul superior, publicat în Monitorul Oficial al României, partea I, nr.880 bis / 13.XII.2011

5. Condiții (acolo unde este cazul)

5.1. de desfășurare a cursului	➤ Sală cu dotare clasică, proiector
5.2. de desfășurare a seminarului/laboratorului	➤ Sală cu calculatoare dotate cu software adecvat

6. Competențe specifice acumulate

Competențe profesionale	C1.2 Utilizarea de teorii, modele conceptuale și instrumente specifice pentru explicarea structurii și funcționalității diverselor tipuri de sisteme software și aplicații informatice complexe de procesare avansată a informației C1.4 Alegerea criteriilor, metodelor și tehnicilor de evaluare a calității, performanțelor și limitărilor diverselor tipuri de sisteme software și aplicații informatice complexe de procesare avansată a informației C6.4 Evaluarea comparativă și analiza critică a soluțiilor de abordare a unor probleme
Competențe transversale	➤ CT1. Aplicarea regulilor de muncă organizată și eficientă, a unor atitudini responsabile față de domeniile didactic, științific și profesional, în vederea valorificării creative a propriului potențial, cu respectarea principiilor și normelor de etică profesională

7. Obiectivele disciplinei (reieșind din grila competențelor specifice acumulate)

7.1. Obiectivul general al disciplinei	Obiectivul principal al disciplinei constă în însușirea de către studenți a principalelor tehnici de data mining
7.2. Obiectivele specifice	După parcurgerea disciplinei studenții vor putea să: <ul style="list-style-type: none">➤ descrie tehnicile de data mining predate➤ utilizeze un software care permite aplicarea tehnicilor de data mining➤ utilizeze tehnici de data mining și să interpreteze rezultatele obținute

8. Conținuturi

8.1. Curs	Nr.ore	Metode de predare	Observații
1.Modele KDP (Knowledge Discovery Process). Metode de reducere a dimensionalității modelelor de regresie liniară. Regresia PCR și regresia PLS	2	Expunere, studii de caz, conversație, dezbateri	
2..Hiperplane separatoare, perceptronul lui Rosenblatt	2	Expunere, studii de caz, conversație, dezbateri	
3.Hiperplane separatoare optimale	2	Expunere, studii de caz, conversație, dezbateri	

4. Clasificatorul SVM (Support Vector Machine) liniar	2	Expunere, studii de caz, conversație, dezbateri	
5. Clasificatorul SVM neliniar	2	Expunere, studii de caz, conversație, dezbateri	
6. Clasificatorul SVR (Support Vector Regression) liniar	2	Expunere, studii de caz, conversație, dezbateri	
7. Clasificatorul SVR (Support Vector Regression) neliniar	2	Expunere, studii de caz, conversație, dezbateri	
8. Tehnici de Data Mining ansambliste : Bagging si Random forests	2	Expunere, studii de caz, conversație, dezbateri	
9. Tehnici de Data Mining ansambliste : Boosting si AdaBoost	2	Expunere, studii de caz, conversație, dezbateri	
10. Metode speciale de clusterizare: DBSCAN (Density Based Spatial Clustering of Applications with Noise)	2	Expunere, studii de caz, conversație, dezbateri	
11. Aplicații Data Mining utilizând mulțimi fuzzy	2	Expunere, studii de caz, conversație, dezbateri	
12. Serii de timp. Modele clasice de predicție bazate pe serii de timp. Etape pentru realizarea predicțiilor cu ajutorul seriilor de timp	2	Expunere, studii de caz, conversație, dezbateri	
13. Predicții bazate pe serii de timp, utilizând rețele neuronale autoregresive (Neural Network Autoregression- NNAR) Indicatori pentru evaluarea calității predicțiilor în cazul seriilor de timp	2	Expunere, studii de caz, conversație, dezbateri	
14. Elemente de Text Mining	2	Expunere, studii de caz, conversație, dezbateri	

Bibliografie

1. Cristian Marinoiu, *Introducere în Învățarea Automată*, Editura Universității Petrol-Gaze din Ploiești, 2019
2. Trevor Hastie, Robert Tibshirani, Jerome Friedman, *The elements of Statistical Learning*, Springer-Verlag, New York, 2001
3. Sergios Theodoridis, Konstantinos Koutroumbas, *Pattern recognition*, Elsevier Inc., 2009
4. Max Bramer, *Principles of Data Mining*, Springer Verlag, London Limited 2007
5. Christopher Bishop, *Pattern recognition and Machine Learning*, Springer Science+Business Media, 2006
6. Eric Biernat, *Data science: Fondamentaux et etudes de cas- Machine learning avec Python and R*
7. Rob J Hyndman and George Athanasopoulos, *Forecasting: Principles and Practice*, disponibil la <https://otexts.com/fpp3/>
8. Cristian Marinoiu, *The classification of Romania's counties using the fuzzy clustering technique*, Economic Insights – Trends and Challenges, Vol. VI(LXIX), No. 2/2017, pp. 15-23, <http://www.upg-bulletin-se.ro/archive-2-2017.html>
9. Cristian Marinoiu, *Using DBSCAN Method to Classify European Countries based on Employments Indicators in Technology and Knowledge -Intensive Sectors*, Proceedings of the 34th International Business Information Management Conference (34th IBIMA) Madrid, Spain 13-14 November 2019, ISBN:

8.2. Seminar / laborator	Nr. ore	Metode de predare	Observații
Aplicații care utilizează tehnici fundamentale de Data Mining	4	Expunere, studii de caz, conversație, dezbatere: aplicații pe calculator	
Aplicații ale tehnicilor SVM și SVR	2	Expunere, studii de caz, conversație, dezbatere: aplicații pe calculator	
Aplicații Data Mining care utilizează metode ansambliste, mulțimi fuzzy, serii de timp, Text Mining	8	Expunere, studii de caz, conversație, dezbatere: aplicații pe calculator	

Bibliografie

1. Trevor Hastie, Robert Tibshirani, Jerome Friedman, *The elements of Statistical Learning*, Springer-Verlag, New York, 2001
2. Sergios Theodoridis, Konstantinos Koutroumbas, *Pattern recognition*, Elsevier Inc., 2009
3. Max Bramer, *Principles of Data Mining*, Springer Verlag, London Limited 2007
4. Christopher Bishop, *Pattern recognition and Machine Learning*, Springer Science+Business Media, 2006
5. Documentație R: <https://cran.r-project.org/manuals.html>
6. Graham Williams, Rattle: A Data Mining GUI for R, *The R Journal*, Dec. 2009
7. Graham Williams, *Data Mining with Rattle and R*, Togaware Series, 2008
8. Emmanuel Paradis, R pentru începători, 2013, <https://fddocuments.in/document/emmanuelparadis-cranr-c-2002-2005-emmanuel-paradis-august-10-2013-permission.html?page=2>
9. <https://www.r-project.org/>
10. Tutorial R, https://www.tutorialspoint.com/data_mining/index.htm
11. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at <https://www.R-project.org/>.

9. Coroborarea conținuturilor disciplinei cu așteptările reprezentanților comunității epistemice, asociațiilor profesionale și angajatori reprezentativi din domeniul aferent programului

Tehnicile de data mining reprezintă o direcție relativ nouă în Informatică. Sprijinite puternic pe rezultatele mai vechi sau mai noi din domeniile statisticii matematice și ale inteligenței artificiale aceste tehnici sunt utilizate din ce în ce mai mult în lucrări de cercetare interdisciplinară datorită potențialului inovativ pe care- l oferă: extragerea cunoștințelor din uriașul volum de date acumulat de-a lungul timpului nu poate fi imaginată astăzi în afara utilizării acestor instrumente. La ora actuală meseria de „miner în date” este practică de diverși specialiști (matematicieni, informaticieni, ingineri, economiști, medici, psihologi, geneticieni etc.) care au un minim de cunoștințe în domeniu și care sunt preocupați în valorificarea informațiilor acumulate în domeniul lor de activitate. Un statut bine conturat al specialistului în data mining este pe cale de a se definitiva (vezi, de exemplu <http://www.socdm.org/index.php/about-us/5-launch-announcement> sau http://www.allanalytics.com/author.asp?section_id=2587&doc_id=271335

10. Evaluare

Tip activitate	10.1. Criterii de evaluare	10.2. Metode de evaluare	10.3. Pondere din nota finală
10.4. Curs	Calitatea prezentării subiectului și a răspunsurilor la examinarea finală	Examinare orală	50%
10.5. Seminar/laborator/	Calitatea activității desfășurate în cadrul laboratorului	Intrebari privitoare la modul de rezolvare a problemelor propuse in cadrul laboratorului	10%
	Calitatea răspunsurilor la testele de control	Teste de control pe parcursul semestrului	20%
	Calitatea proiectului elaborat în cadrul laboratorului	Prezentarea proiectului pe calculator	20%
10.6 Proiect	-----		
10.7. Standard minim de performanță			
<p>Studentul trebuie să</p> <ul style="list-style-type: none"> ➤ descrie corect ideile care stau la baza metodelor prezentate la curs ➤ știe să utilizeze pachetul de programe prezentat la laborator pentru rezolvarea problemelor care pot fi soluționate utilizând modelele anterioare 			

Data
completării

Semnătura titularului de curs

Semnătura titularului de
seminar/laborator

Semnătura titularului de proiect

Conf. dr. Marinoiu Cristian _

Sef lucrari dr. Ioniță Irina

Data avizării în
departament

Director de departament
Conf. dr. Gabriela Moise

Decan
Prof. dr. Suditu Mihaela