

FIȘA DISCIPLINEI¹⁾

1. Date despre program

1.1. Instituția de învățământ superior	Universitatea Petrol-Gaze din Ploiești
1.2. Facultatea	Litere si Stiințe
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	Lector dr. Tudorică Daniela
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	➤ 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
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: ➤ 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). Forma standard de prezentare a datelor în Data Mining	2	Expunere, studii de caz, conversație, dezbateri	
2. Tipuri de învățare automată. Scale de măsurători. Pregătirea datelor pentru aplicarea tehnicilor de Data Mining	2	Expunere, studii de caz, conversație, dezbateri	
3.Funcție de cost, risc, risc empiric.	2	Expunere, studii de caz,	

Modele de regresie și de clasificare, Prezentarea modelului de regresie liniară		conversație, dezbateri	
4 Rezolvarea modelului de regresie liniară, ,	2	Expunere, studii de caz, conversație, dezbateri	
5.Considerații geometrice asupra modelului de regresie liniară. Overfitting, underfitting, compromisul deplasare- varianță	2	Expunere, studii de caz, conversație, dezbateri	
6.Regresia regularizată (Ridge, LASSO, ElasticNet).	2	Expunere, studii de caz, conversație, dezbateri	
7. Algoritmi de învățare supervizată pentru clasificare (Bayes, Naive Bayes, Analiza discriminantă lineară-LDA)	2	Expunere, studii de caz, conversație, dezbateri	
8. Algoritmi de învățare supervizată pentru clasificare (Regresia logistică, KNN)	2	Expunere, studii de caz, conversație, dezbateri	
9. Arbori de decizie. Algoritmul TDIDT	2	Expunere, studii de caz, conversație, dezbateri	
10. Utilizarea entropiei pentru selecția atributului pentru divizare. Algoritmul TDIDT pentru date continue și neadevrate	2	Expunere, studii de caz, conversație, dezbateri	
11. Clusteri neierarhici	2	Expunere, studii de caz, conversație, dezbateri	
12. Clusteri ierarhici	2	Expunere, studii de caz, conversație, dezbateri	
13. Metode și indicatori pentru evaluarea performanțelor modelelor de regresie	2	Expunere, studii de caz, conversație, dezbateri	
14. Metode și indicatori pentru evaluarea performanțelor modelelor de clasificare	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			
8.2. Seminar / laborator	Nr. ore	Metode de predare	Observații
Elemente de bază privind operarea într-un pachet de programe specializat	2	Expunere, studii de caz, conversație, dezbateri: aplicații pe calculator	
Aplicații ale modelelor de regresie liniară simplă și multiplă	2	Expunere, studii de caz, conversație, dezbateri: aplicații pe calculator	
Aplicații ale modelelor de regresie regularizate ((Ridge, LASSO, ElasticNet)	2	Expunere, studii de caz, conversație, dezbateri: aplicații pe calculator	

Utilizarea clasificatorilor LDA (Analiza discriminantă lineară), Regresia logistică, Naive Bayes și K Nearest Neighbours (KNN) în aplicații	4	Expunere, studii de caz, conversație, dezbateri: aplicații pe calculator	
Aplicații ale arborilor de decizie	2	Expunere, studii de caz, conversație, dezbateri: aplicații pe calculator	
Aplicații ale metodelor de clusterizare	2	Expunere, studii de caz, conversație, dezbateri: 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			

3. 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	Intrebari privitoare la modul de	10%

	desfășurate în cadrul laboratorului	rezolvare a problemelor propuse în cadrul laboratorului	
	Calitatea răspunsurilor la testele de control	Teste de control pe parcursul semestrului	20%
	Calitatea proiectului elaborat în cadrul laboratorului	Prezentarea proiectului I pe calculator	20%

10.6. Standard minim de performanță

Studentul trebuie să

- descrie corect metodele regresie și de clasificare prezentate la curs
- știe să utilizeze pachetul de programe prezentat la laborator pentru rezolvarea problemelor care pot fi soluționate utilizând modelele anterioare
- descrie corect cel puțin un indicator sau cel puțin una dintre metode pentru evaluarea performanțelor modelelor de regresie sau de clasificare prezentate la curs

Data completării	Semnătura titularului de curs	Semnătura titularului de seminar/laborator	Semnătura titularului de proiect
------------------	-------------------------------	--	----------------------------------

9.09.2020

Conf. dr. Marinoiu Cristian

Lector dr.

Tudorică Daniela

-

Data avizării în departament

Semnătura directorului de departament

Conf. dr. Gabriela Moise