

# FIȘA DISCIPLINEI<sup>1)</sup>

## 1. Date despre program

1.1. Instituția de învățământ superior	Universitatea Petrol-Gaze din Ploiesti
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	Licență
1.6. Programul de studii universitare	Informatică

## 2. Date despre disciplină

2.1. Denumirea disciplinei	Introducere în Data Mining
2.2. Titularul activităților de curs	Conf.dr.mat. Marinoiu Cristian
2.3. Titularul activităților aplicative	Lector dr. Ioniță Irina
2.4. Anul de studiu	III
2.5. Semestrul *	6
2.6. Tipul de evaluare	Examen
2.7. Categoria formativă** / regimul*** disciplinei	F0/O

\* numărul semestrului este conform planului de învățământ;

\*\* fundamentală = F0; de domeniu = D1; de specialitate = S2; complementară = C3

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

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

3.1. Număr de ore pe săptămână	4	din care: 3.2. curs	2	3.3. Seminar/laborator	2
3.4. Total ore din planul de învățământ	40	din care: 3.5. curs	20	3.6. Seminar/laborator	20
3.7. Distribuția fondului de timp					ore
Studiu după manual, suport de curs, bibliografie și notițe					20
Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate și pe teren					20
Pregătire seminarii/laboratoare, teme, referate, portofolii și eseuri					28
Tutoriat					14
Examinări					3
Alte activități					-
3.7. Total ore studiu individual	85				
3.8. Total ore pe semestru	125				
3.9. Numărul de credite	5				

## 4. Precondiții (acolo unde este cazul)

4.1. de curriculum	Modelare și Simulare
4.2. de competențe	Abilități de modelare

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	<ul style="list-style-type: none"> <li>Sală cu dotare clasică</li> </ul>
5.2. de desfășurare a seminarului/laboratorului	<ul style="list-style-type: none"> <li>Sală de calculatoare cu soft adecvat</li> </ul>

## 6. Competențe specifice acumulate

<b>Competențe profesionale</b>	<p>C3.1 Descrierea de concepte, teorii și modele folosite în domeniul de aplicare</p> <p>C3.2 Identificarea și explicarea modelelor informatice de bază adecvate domeniului de aplicare</p> <p>C2.5 Realizarea unor proiecte informatice dedicate</p> <p>C3.3 Utilizarea modelelor și instrumentelor informatice și matematice pentru rezolvarea problemelor specifice domeniului de aplicare</p> <p>C3.4 Analiza datelor și a modelelor</p> <p>C4.2 Interpretarea de modele matematice și informatice</p> <p>C4.3 Identificarea modelelor și metodelor adecvate pentru rezolvarea unor probleme reale</p>
<b>Competențe transversale</b>	<p>CT1 Aplicarea regulilor de muncă organizată și eficientă, a unor atitudini responsabile față de domeniul didactic-științific, pentru valorificarea creativă a propriului potențial, cu respectarea principiilor și a normelor de etică profesională</p>

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

7.1. Obiectivul general al disciplinei	<p>Obiectivul principal al disciplinei constă în înțelegerea și asimilarea de către studenți a principiilor de bază ale data mining-ului prin intermediul unor tehnici clasice din acest domeniu</p>
7.2. Obiectivele specifice	<p>După parcurgerea disciplinei studenții vor putea să:</p> <ul style="list-style-type: none"> <li>facă diferența între o problemă de învățare supervizată și una de învățare nesupervizată</li> <li>propună și să rezolve modele liniare de regresie simplă</li> <li>valideze modelul liniar propus și să-l utilizeze pentru predicție</li> <li>înțeleagă și să utilizeze regresia logistică pentru predicție</li> <li>să utilizeze o tehnică de clusterizare pentru detectarea claselor</li> </ul>

## 8. Conținuturi

8.1. Curs	Nr.ore	Metode de predare	Observații
1. Introducere în Data Mining	1.2	Expunere, studii de caz,	
2. Modele de regresie liniară simplă	2.8	conversație, dezbateri	

(Covarianță, corelație, regresie. Rezolvarea modelului prin metoda celor mai mici pătrate. Estimarea coeficienților de regresie. Reziduri-proprietăți. Estimarea dispersiei erorilor. Teorema Gauss-Markov. Utilizarea modelului liniar de regresie simplă pentru predicție.) 3. Modelul liniar de regresie liniară simplă în ipoteza de normalitate a erorilor ( Intervale, elipse de încredere, bandă de încredere. Teste pentru verificarea ipotezelor statistice asupra parametrilor modelului. Validarea modelului liniar de regresie simplă. 4. Tehnici de clasificare supervizată și nesupervizată : regresia logistică binară și clusteri ierarhici	3.6 4.4		
Bibliografie			
1. Cristian Marinoiu, Modele de regresie liniară, Editura Universității Petrol-Gaze din Ploiești, Ploiești, 2015			
2. Cornillon, Pierre-André, Matzner-Løber , Régression Théorie et applications, Springer Paris, 2007			
3. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The elements of Statistical Learning, Springer-Verlag, New York, 2001			
4. Douglas Montgomery, Elizabeth A. Peck, G. Geoffrey Vining, <i>Introduction to Linear Regression Analysis</i> , John Wiley and Sons, Inc Publication, New Jersey, U.S.A, 2012			
5. Max Bramer, Principles of Data Mining, Springer Verlag, London Limited 2007			
<b>8.2. Seminar / laborator/proiect</b>	Nr. ore	Metode de predare	Observații
1. Organizarea datelor pentru minerit in date (data mining)	1.2 2.12	Expunere, studii de caz, conversație, dezbateri: aplicații	
2. Aplicații ale regresiei liniare simple	3.6	pe calculator	
3. Aplicații care utilizează regresia logistică binară și clusteri ierarhici			
Bibliografie			
1. Cornillon, Pierre-André, Matzner-Løber , Régression Théorie et applications, Springer Paris, 2007			
2. Trevor Hastie, Robert Tibshirani, Jerome Friedman, The elements of Statistical Learning, Springer-Verlag, New York, 2001			
3. Douglas Montgomery, Elizabeth A. Peck, G. Geoffrey Vining, <i>Introduction to Linear Regression Analysis</i> , John Wiley and Sons, Inc Publication, New Jersey, U.S.A, 2012			
4. Max Bramer, Principles of Data Mining, Springer Verlag, London Limited 2007			
5. Octave Tutorial <a href="https://en.wikibooks.org/wiki/Octave_Programming_Tutorial">https://en.wikibooks.org/wiki/Octave_Programming_Tutorial</a>			

## 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](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 cu bilete	50%
10.5. Seminar/laborator/proiect	Calitatea activității desfășurate în cadrul laboratorului	Intrebări privitoare la modul de rezolvare a problemelor propuse in cadrul laboratorului	25%
	Calitatea răspunsurilor la testele de control	Teste de control pe parcursul semestrului	25%
10.6. Standard minim de performanță			
Studentul trebuie să			
<ul style="list-style-type: none"> <li>• descrie corect modelele de regresie liniară, cel de regresie logistică binară și cel de clustering;</li> <li>• știe să utilizeze pachetul de programe prezentat la laborator pentru rezolvarea regresii liniare simple și a regresiei logistice binare și să interpreteze corect rezultatele.</li> </ul>			

Data completării  
15.09.2018

Semnătura titularului de curs  
Conf. dr. Marinou Cristian

Semnătura titularului de seminar/laborator  
Lector dr. Ioniță Irina

Data avizării în departament

Semnătura directorului de departament

Conf. dr. Gabriela Moise