

FIȘA DISCIPLINEI

Anul universitar 2023-2024

Anul de studiu I / Semestrul II

1. Date despre program

1.1. Instituția de învățământ superior	Universitatea „1 Decembrie 1918” din Alba Iulia
1.2. Facultatea	Facultatea de Informatică și Inginerie
1.3. Departamentul	Departamentul de Informatică, Matematică și Electronică
1.4. Domeniul de studii	Informatică
1.5. Ciclul de studii	Master
1.6. Programul de studii/calificarea	Programare avansată și baze de date/COR 251201, 251204, 251203

2. Date despre disciplină

2.1. Denumirea disciplinei	Data mining			2.2. Cod disciplină	MI 104		
2.3. Titularul activității de curs	Muntean Maria-Viorela						
2.4. Titularul activității de laborator	Muntean Maria-Viorela						
2.5. Anul de studiu	I	2.6. Semestrul	II	2.7. Tipul de evaluare (E/C/VP)	E	2.8. Regimul disciplinei (O – obligatorie, Op – opțională, F – facultativă)	

3. Timpul total estimat

3.1. Numar ore pe saptamana	4	din care: 3.2. curs	2	3.3. laborator	2
3.4. Total ore din planul de învățământ	56	din care: 3.5. curs	28	3.6. laborator	28
Distribuția fondului de timp					ore
Studiul după manual, suport de curs, bibliografie și notițe					40
Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate și pe teren					37
Pregătire laboratoare, teme, referate, portofolii și eseuri					40
Tutoriat					-
Examinări					2
Alte activități					-

3.7 Total ore studiu individual	119
3.9 Total ore pe semestru	175
3.10 Numărul de credite	7

4. Precondiții (acolo unde este cazul)

4.1. de curriculum	Baze de date Inteligența artificială
4.2. de competențe	

5. Condiții (acolo unde este cazul)

5.1. de desfășurare a cursului	Sală de curs dotată cu videoproiector/tablă / Platforma Microsoft Teams
5.2. de desfășurarea a laboratorului	Sală de laborator dotată cu: calculatoare (licență MS Windows), conexiune Internet și cu videoproiector / Platforma Microsoft Teams

6. Competențe specifice acumulate

Competențe profesionale	Capacitatea de a: 1) implementa tehnici de data mining, folosind diverse medii; 2) realiza sisteme de data mining. Capacitatea de a cunoaște și înțelege: 1) conceptele de bază ale procesului de data mining; 2) principalele tipuri și tehnici de învățare automată; 3) principalele abordări ale extragerii cunoștințelor din baze mari de date.
Competențe transversale	Nu este cazul

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

7.1 Obiectivul general al disciplinei	Cursul realizează o introducere în domeniul Data mining, tratând aspecte teoretice și practice ale acesteia.
7.2 Obiectivele specifice	7.2.1 Se pun în evidență principalele domenii și direcțiile fundamentale de evoluție ale Data mining. 7.2.2 Studenții se vor familiariza cu principalele abordări ale Data mining și cu o serie de aplicații de Data mining.

8. Conținuturi

8.1 Curs	Metode de predare	Observații
<p>1. INTRODUCERE</p> <p>1.1 Data mining – definirea domeniului, specificarea factorilor care au favorizat dezvoltarea domeniului și definirea procesului de extragere a cunoștințelor din baze de date (Knowledge Discovery in Databases)</p> <p>1.2. Exemplu de model de Data mining (Ob.sp. 7.2.1)</p>	<ul style="list-style-type: none"> • Slide-uri Power-Point • Explicații pe tablă • Intrebări/răspunsuri, discuții • Exerciții și probleme 	<p>Studentii au acces la suportul de curs în format electronic. Prezentare folosind laptop și videoproiector</p> <p>2 ore</p>
<p>2. CULEGEREA DATELOR ȘI SISTEMATIZAREA ACESTORA ÎN BAZE DE DATE</p> <p>2.1. Operații de depozitare date (acumulare, conversie, curățire, integrare și transformare, reducere și discretizare)</p> <p>2.2. Reconstituirea datelor lipsă (Ob.sp. 7.2.2)</p>	<ul style="list-style-type: none"> • Slide-uri Power-Point • Explicații pe tablă • Intrebări/răspunsuri, discuții • Exerciții și probleme 	<p>Studentii au acces la suportul de curs în format electronic. Prezentare folosind laptop și videoproiector.</p> <p>2 ore</p>
<p>3. CLASIFICARE ȘI PREDICȚIE</p> <p>3.1. Definirea problemei</p> <p>3.2. Arbori de clasificare/decizie</p> <p>3.3. Clasificatori bazați pe reguli de decizie</p> <p>3.4. Mașini cu suport vectorial</p> <p>3.5. Metaclassificatori</p> <p>3.6. Clasificatori de tip k-nearest neighbour</p> <p>3.7. Clasificatori Bayesieni</p> <p>3.8. Rețele neuronale artificiale. Deep learning</p> <p>3.9. Algoritmi genetici în clasificare</p> <p>3.10. Sisteme de învățare multiple (Ob.sp. 7.2.2)</p>	<ul style="list-style-type: none"> • Slide-uri Power-Point • Explicații pe tablă • Intrebări/răspunsuri, discuții • Exerciții și probleme 	<p>Studentii au acces la suportul de curs în format electronic. Prezentare folosind laptop și videoproiector.</p> <p>10 ore</p>
<p>4. METODE NUMERICE PENTRU PREVIZIUNE</p> <p>4.1. Regresia liniară</p> <p>4.2. Perceptronul</p> <p>4.3. Mașini cu suport vectorial</p> <p>4.4. Arbori de regresie (Ob.sp. 7.2.2)</p>	<ul style="list-style-type: none"> • Slide-uri Power-Point • Explicații pe tablă • Intrebări/răspunsuri, discuții • Exerciții și probleme 	<p>Studentii au acces la suportul de curs în format electronic. Prezentare folosind laptop și videoproiector.</p> <p>2 ore</p>
<p>5. CLUSTERIZARE (ANALIZA CLUSTERILOR)</p> <p>5.1. Definirea problemei</p> <p>5.2. Tipuri de date în analiza clusterilor</p> <p>5.3. Clasificarea metodelor de clusterizare</p> <p>5.4. Clase de metode de clusterizare: de partiționare, ierarhice, bazate pe densitate, bazate pe griduri, bazate pe modele</p> <p>5.5. Detectia excepțiilor (outliers) (Ob.sp. 7.2.2)</p>	<ul style="list-style-type: none"> • Slide-uri Power-Point • Explicații pe tablă • Intrebări/răspunsuri, discuții • Exerciții și probleme 	<p>Studentii au acces la suportul de curs în format electronic. Prezentare folosind laptop și videoproiector.</p> <p>2 ore</p>
<p>6. DESCOPERIREA REGULILOR DE ASOCIERE (ANALIZA DEPENDENȚELOR DINTRE LEGĂTURI)</p> <p>6.1. Definirea problemei</p> <p>6.2. Descoperirea asocierilor</p> <p>6.2.1. Algoritmul Apriori</p> <p>6.2.2. Dezvoltarea modelului frecvent</p> <p>6.3. Descoperirea metodelor secvențiale</p> <p>6.4. Descoperirea secvențelor de timp similare (Ob.sp. 7.2.2)</p>	<ul style="list-style-type: none"> • Slide-uri Power-Point • Explicații pe tablă • Intrebări/răspunsuri, discuții • Exerciții și probleme 	<p>Studentii au acces la suportul de curs în format electronic. Prezentare folosind laptop și videoproiector.</p> <p>2 ore</p>
<p>7. IDENTIFICAREA ANOMALIILOR ȘI DEVIATIILOR</p> <p>7.1. Metode grafice (box-plot, diagrame de dispersie, acoperire convexă);</p> <p>7.2. Metode statistice (testul Grubbs, testul verosimilității);</p> <p>7.3. Metode bazate pe măsurarea distanței (metoda k-nearest neighbour, metoda clustering-ului);</p> <p>7.4. Metode bazate pe modele (clasificarea, mașini instruibile, modele autoregresive). (Ob.sp. 7.2.2)</p>	<ul style="list-style-type: none"> • Slide-uri Power-Point • Explicații pe tablă • Intrebări/răspunsuri, discuții • Exerciții și probleme 	<p>Studentii au acces la suportul de curs în format electronic. Prezentare folosind laptop și videoproiector.</p> <p>2 ore</p>
<p>8. EVALUAREA MODELELOR</p> <p>8.1. Holdout</p> <p>8.2. K-folds cross validation</p> <p>8.3. Bootstrap</p> <p>8.4. Costul și acuratețea clasificării</p> <p>8.5. K-statistics</p> <p>8.6. Lift-charts</p> <p>8.7. Curbe ROC (Receiver Operating Characteristic)</p> <p>8.8. Coeficientul de corelație (Ob.sp. 7.2.2)</p>	<ul style="list-style-type: none"> • Slide-uri Power-Point • Explicații pe tablă /Intrebări/răspunsuri, discuții • Exerciții și probleme 	<p>Studentii au acces la suportul de curs în format electronic. Prezentare folosind laptop și videoproiector.</p> <p>2 ore</p>

<p>9. APLICAȚII ȘI TENDINȚE IN DATA MINING</p> <p>9.1. Aplicații: telecomunicații, analiza datelor financiare, analiza datelor biologice, etc.</p> <p>9.2. Data mining în baze de date statistice, audio, video</p> <p>9.3. Data mining, securitatea și secretul (privacy) datelor (Ob.sp. 7.2.2)</p>	<ul style="list-style-type: none"> • Slide-uri Power-Point • Explicații pe tablă • Întrebări/răspunsuri, discuții • Exerciții și probleme 	<p>Studentii au acces la suportul de curs în format electronic. Prezentare folosind laptop și videoproiector.</p> <p>4 ore</p>
<p>Bibliografie</p> <ol style="list-style-type: none"> 1. Muntean, M., <i>Data mining. Teorie și aplicații</i>, Ed. Aeternitas, 2011, Alba Iulia, ISBN: 978-606-613-015-8. 2. Ileană, I., Rotar, C., Muntean, M., <i>Inteligență artificială</i>, Ed. Aeternitas, Alba Iulia, 2009, ISBN: 978-973-1890-49-4. 3. Radu G. Crestulescu; Daniel I. Morariu, <i>Text Mining. Tehnici de clasificare și clustering al documentelor</i>, Ed. Albastră, Cluj-Napoca, 2012 4. Gorunescu, F., <i>Data mining. Concepte, modele și tehnici</i>, Ed. Albastră, Cluj-Napoca, 2006, ISBN: 973-650-169-8. 5. Mark Last, Piotr S. Szczepaniak, Zeev Volkovich, Abraham Kandel (Eds.): <i>Advances in Web Intelligence and Data Mining. Studies in Computational Intelligence</i> Vol. 23 Springer 2006. 6. Daniel Barbará, Sushil Jajodia, <i>Applications of data mining in computer security</i>, Kluwer Academic Publisher, 2002. 7. Hillol Kargupta, Jiawei Han, Philip S. Yu, Rajeev Motwani, Vipin Kumar, <i>Next Generation of Data Mining</i>, Taylor and Francis Group, Chapman & Hall, 2009. 		
<p>8.2. Laborator</p>		
<p>1. Instalarea softului necesar dezvoltării și implementării aplicațiilor. Introducere</p>	<ul style="list-style-type: none"> • Explicații pe studii de caz utilizând videoproiector • Întrebări/răspunsuri, discuții • Lucrul individual pe calculator pentru instalarea soft-ului necesar 	<p>2 ore</p>
<p>2. Metode de preprocesare a datelor. Aplicații</p>	<ul style="list-style-type: none"> • Explicații pe studii de caz utilizând videoproiector/ tablă • Întrebări/răspunsuri, discuții • Lucrul individual sau în echipe, pe calculator • Aplicații propuse • Activități de testare 	<p>2 ore</p>
<p>3. Algoritmi bazați pe arbori decizionali. Aplicații</p>	<ul style="list-style-type: none"> • Explicații pe studii de caz utilizând videoproiector/ tablă • Întrebări/răspunsuri, discuții • Lucrul individual sau în echipe, pe calculator • Aplicații propuse • Activități de testare 	<p>2 ore</p>
<p>4. Algoritmi bazați pe reguli. Aplicații</p>	<ul style="list-style-type: none"> • Explicații pe studii de caz utilizând videoproiector/ tablă • Întrebări/răspunsuri, discuții • Lucrul individual sau în echipe, pe calculator • Aplicații propuse • Activități de testare 	<p>2 ore</p>
<p>5. Clasificatori Bayesiani. Aplicații</p>	<ul style="list-style-type: none"> • Explicații pe studii de caz utilizând videoproiector/ tablă • Întrebări/răspunsuri, discuții • Lucrul individual sau în echipe, pe calculator • Aplicații propuse • Activități de testare 	<p>2 ore</p>
<p>6. Rețele neuronale artificiale. Aplicații</p>	<ul style="list-style-type: none"> • Explicații pe studii de caz utilizând videoproiector/ tablă • Întrebări/răspunsuri, discuții • Lucrul individual sau în echipe, pe calculator • Aplicații propuse • Activități de testare 	<p>2 ore</p>
<p>7. Clasificatori de tip k-nearest neighbour. Aplicații</p>	<ul style="list-style-type: none"> • Explicații pe studii de caz utilizând videoproiector/ tablă • Întrebări/răspunsuri, discuții • Lucrul individual sau în echipe, pe calculator • Aplicații propuse • Activități de testare 	<p>2 ore</p>
<p>8. Mașini cu suport vectorial. Aplicații</p>	<ul style="list-style-type: none"> • Explicații pe studii de caz utilizând videoproiector/ tablă • Întrebări/răspunsuri, discuții • Lucrul individual sau în echipe, pe calculator • Aplicații propuse • Activități de testare 	<p>2 ore</p>
<p>9. Optimizarea sistemelor de data mining cu algoritmi genetici. Aplicații</p>	<ul style="list-style-type: none"> • Explicații pe studii de caz utilizând videoproiector/ tablă • Întrebări/răspunsuri, discuții 	<p>2 ore</p>

	<ul style="list-style-type: none"> • Lucrul individual sau în echipe, pe calculator • Aplicații propuse • Activități de testare 	
10. Clase de metode de clusterizare: de partiționare, ierarhice, bazate pe densitate, bazate pe griduri, bazate pe modele. Aplicații	<ul style="list-style-type: none"> • Explicații pe studii de caz utilizând vidoproiector/ tablă • Întrebări/răspunsuri, discuții • Lucrul individual sau în echipe, pe calculator • Aplicații propuse • Activități de testare 	2 ore
11. Algoritmi pentru descoperirea regulilor de asociere. Algoritmul Apriori. Aplicații	<ul style="list-style-type: none"> • Explicații pe studii de caz utilizând vidoproiector/ tablă • Întrebări/răspunsuri, discuții • Lucrul individual sau în echipe, pe calculator • Aplicații propuse • Activități de testare 	2 ore
12. Identificarea anomaliilor și deviațiilor. Aplicații	<ul style="list-style-type: none"> • Explicații pe studii de caz utilizând vidoproiector/ tablă • Întrebări/răspunsuri, discuții • Lucrul individual sau în echipe, pe calculator • Aplicații propuse • Activități de testare 	2 ore
13. Curbe de învățare. Costul și acuratețea clasificării. Curbe ROC. Studii de caz	<ul style="list-style-type: none"> • Explicații pe studii de caz utilizând vidoproiector/ tablă • Întrebări/răspunsuri, discuții • Lucrul individual sau în echipe, pe calculator • Aplicații propuse • Activități de testare 	2 ore
14. Verificarea practică a cunoștințelor	<ul style="list-style-type: none"> • Lucrul individual sau în echipe, pe calculator • Activități de testare 	2 ore

Bibliografie

1. Muntean, M., *Data mining. Teorie și aplicații*, Ed. Aeternitas, 2011, Alba Iulia, ISBN: 978-606-613-015-8.
2. Ileană, I., Rotar, C., Muntean, M., *Inteligență artificială*, Ed. Aeternitas, Alba Iulia, 2009, ISBN: 978-973-1890-49-4.
3. Belciug, Smaranda, Gorunescu, Mariana, *DATA MINING: Modele predictive și de clasificare. Implementare în MATLAB și JAVA*, Editura Albastră, Cluj Napoca, 2012.
4. Nicoleta Petcu, *Tehnici de Data Mining rezolvate în SPSS Clementine*, Editura Albastră, Cluj Napoca, 2010
5. Weka system and documentation (<http://www.cs.waikato.ac.nz/ml/weka/>). Weka is a suite of machine learning / data mining software. It contains Java implementation for various mining algorithms, data preprocessing filters, and experimentation capabilities. Weka is free open-source software under the GNU General Public License (GPL).
6. <https://www.cs.waikato.ac.nz/ml/weka/>
7. <https://www.java.com/en/>

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

Aptitudini și competențe solicitate de reprezentanților comunității epistemice, asociațiilor profesionale și angajatori și dobândite de studenți la finalul cursului:

- abilitatea de a dezvolta și de a lucra cu sisteme de data mining;
- abilitatea de a integra sisteme de data mining în aplicații reale.

10. Evaluare

Tip activitate	10.1 Criterii de evaluare	10.2 Metode de evaluare	10.3 Pondere din nota finală
10.4 Curs	<p><i>Completitudinea și corectitudinea cunoștințelor</i></p> <p><i>Capacitatea de a opera cu cunoștințele asimilate în activități intelectuale complexe</i></p> <p><i>Capacitatea de analiză, de interpretare personală, originalitatea, creativitatea</i></p>	<i>Verificarea scrisă</i>	50%
10.5 Laborator	<p><i>Capacitatea de aplicare în practică, în contexte diferite, a cunoștințelor învățate</i></p> <p><i>Gradul de asimilare a limbajului de specialitate și capacitatea de comunicare</i></p>	<i>Verificarea practică</i>	50%
10.6 Standard minim de performanță: Cel puțin nota 5 (pe o scară de la 1 la 10) la verificarea scrisă. Cel puțin nota 5 (pe o scară de la 1 la 10) la verificarea practică.			

Data completării

Semnătura titularului de curs
Lect. univ. dr. Muntean Maria – Viorela

Semnătura titularului de laborator
Lect. univ. dr. Muntean Maria - Viorela

Data avizării în departament

Semnătura directorului de departament
Lect univ. dr. Aldea Mihaela

Data aprobării în Consiliul Facultății

Semnătura Decanul Facultății
Conf univ. dr. Rotar Corina

Anexă la Fișa disciplinei (facultativă)

ANEXĂ LA FIȘA DISCIPLINEI

b. Evaluare – mărirea de notă

Tip activitate	10.1 Criterii de evaluare	10.2 Metode de evaluare	10.3 Pondere din nota finală
10.4 Curs	<i>Completitudinea și corectitudinea cunoștințelor</i> <i>Capacitatea de a opera cu cunoștințele asimilate în activități intelectuale complexe</i> <i>Capacitatea de analiză, de interpretare personală, originalitatea, creativitatea</i>	Verificarea scrisă	50%
10.5 Laborator	<i>Capacitatea de aplicare în practică, în contexte diferite, a cunoștințelor învățate</i> <i>Gradul de asimilare a limbajului de specialitate și capacitatea de comunicare</i>	Verificarea practică	50%
10.6 Standard minim de performanță			
Cel puțin nota 5 (pe o scară de la 1 la 10) la verificarea scrisă. Cel puțin nota 5 (pe o scară de la 1 la 10) la verificarea practică.			
Data completării	Semnătura titularului de curs	Semnătura titularului de laborator	

c. Evaluare – restanță

Tip activitate	10.1 Criterii de evaluare	10.2 Metode de evaluare	10.3 Pondere din nota finală
10.4 Curs	<i>Completitudinea și corectitudinea cunoștințelor</i> <i>Capacitatea de a opera cu cunoștințele asimilate în activități intelectuale complexe</i> <i>Capacitatea de analiză, de interpretare personală, originalitatea, creativitatea</i>	Verificarea scrisă	50%
10.5 Laborator	<i>Capacitatea de aplicare în practică, în contexte diferite, a cunoștințelor învățate</i> <i>Gradul de asimilare a limbajului de specialitate și capacitatea de comunicare</i>	Verificarea practică	50%
10.6 Standard minim de performanță			
Cel puțin nota 5 (pe o scară de la 1 la 10) la verificarea scrisă. Cel puțin nota 5 (pe o scară de la 1 la 10) la verificarea practică.			
Data completării	Semnătura titularului de curs	Semnătura titularului de laborator	