

FI A DISCIPLINEI

1. Date despre program

1.1 Instituția de învățământ superior	Universitatea Babeș-Bolyai Cluj-Napoca
1.2 Facultatea	Facultatea de Matematica și Informatică
1.3 Departamentul	Departamentul de Informatică
1.4 Domeniul de studii	Informatică
1.5 Ciclul de studii	Master
1.6 Programul de studiu / Calificarea	Sisteme distribuite în Internet - română

2. Date despre disciplină

2.1 Denumirea disciplinei	Data Mining						
2.2 Titularul activităților de curs	Lect. Dr. Anca Andreica						
2.3 Titularul activităților de seminar	Lect. Dr. Anca Andreica						
2.4 Anul de studiu	2	2.5 Semestrul	3	2.6. Tipul de evaluare	E	2.7 Regimul disciplinei	Optional

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	1 sem + 1 proiect
3.4 Total ore din planul de învățământ	56	Din care: 3.5 curs	28	3.6 seminar/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					40
Pregătire seminarii/laboratoare, teme, referate, portofolii și eseuri					40
Tutoriat					10
Examinări					14
Alte activități:					
3.7 Total ore studiu individual			144		
3.8 Total ore pe semestru			200		
3.9 Numărul de credite			8		

4. Precondiții (acolo unde este cazul)

4.1 de curriculum	•
4.2 de competențe	• Cunoștințe medii de programare într-un limbaj de programare de nivel înalt

5. Condiții (acolo unde este cazul)

5.1 De desfășurare a cursului	•
5.2 De desfășurare a seminarului/laboratorului	• Laborator cu calculatoare

6. Competențele specifice acumulate

Competențe profesionale	<ul style="list-style-type: none"> • Abilități practice pentru modelarea și rezolvarea problemelor reale • Abilitatea de a lucra în echipă, asumarea diferitelor roluri de execuție și conducere de proiecte, asumarea activităților profesionale cu autonomie și responsabilitate
Competențe transversale	<ul style="list-style-type: none"> • Studenții vor învăța diferite tehnici de analiză a datelor și vor aplica aceste tehnici pentru rezolvarea diferitelor probleme reale de data mining folosind instrumente speciale

7. Obiectivele disciplinei (reie ind din grila competen elor acumulate)

7.1 Obiectivul general al disciplinei	<ul style="list-style-type: none"> • Însușirea conceptelor, metodelor și tehnicilor data mining, din perspectiva bazelor de date
7.2 Obiectivele specifice	<ul style="list-style-type: none"> • Studenții vor deprinde diverse tehnici de analiză a datelor, și vor aplica aceste tehnici pentru rezolvarea unor probleme data mining folosind sisteme software speciale • Se va forma o percepție asupra data mining ca fiind un puternic domeniu aplicativ, dar și un domeniu de cercetare semnificativ în bazele de date.

8. Conținuturi

8.1 Curs	Metode de predare	Observații
1. Introducere; Tipuri de date explorate în data mining; Funcționalități (task-uri) data mining; Algoritmii; Data mining din perspectiva bazelor de date; Aplicații data mining	Expunerea, conversația, dezbateră, problematizarea, descoperirea	
2-3. Descrierea conceptelor; Definiții; Generalizarea datelor și caracterizarea bazată pe rezumare; Caracterizarea analitică: analiza relevanței atributelor; Compararea claselor: diferențierea claselor; Măsurile statistice descriptive în baze de date mari	Expunerea, conversația, dezbateră, problematizarea, descoperirea	
4. Preprocesarea datelor; Curățarea datelor; Transformarea și integrarea datelor; Reducerea datelor	Expunerea, conversația, dezbateră, problematizarea, descoperirea	

5-6. Descoperirea regulilor de asociere; Definitia problemei; Algoritmi pentru extragerea regulilor de asociere unidimensionale, booleene; Algoritmi pentru extragerea regulilor de asociere multinivel, multidimensionale, cu constrângeri; Analiza corelațiilor	Expunerea, conversația, dezbateră, problematizarea, descoperirea	
7-9. Clasificare și predicție; Definitia problemei; Clasificare folosind metoda k-nearest neighbor; Clasificare folosind arbori de decizie; Clasificare și estimare folosind rețele neuronale; Clasificare Bayesiană	Expunerea, conversația, dezbateră, problematizarea, descoperirea	
10-11. Clusterizare (analiza clusterilor); Definitia problemei; Tipuri de date în analiza clusterilor; Metode de clusterizare: de partiționare, ierarhice, bazate pe densitate, bazate pe griduri, bazate pe modele; Detectarea excepțiilor	Expunerea, conversația, dezbateră, problematizarea, descoperirea	
12. Standarde și software data mining	Expunerea, conversația, dezbateră, problematizarea, descoperirea	
13-14. Prezentările rapoartelor teoretice		
Bibliografie:		
<ol style="list-style-type: none"> 1. S. Chakrabarti et al, Data Mining. Know It All, Morgan Kaufmann, 2009. 2. K. Cios, W. Pedrycz, R. Swiniarski, L. Kurgan, Data Mining. A Knowledge Discovery Approach, Springer, 2007. 3. J. Han, M. Kamber, Data Mining: Concepts and Techniques, 2nd Edition, Morgan Kaufmann, 2006. 4. P. Tan, M. Steinbach, V. Kumar, Introduction to Data Mining, Addison Wesley, 2006. 5. D. Larose, Discovering Knowledge in Data. An Introduction to Data Mining, John Wiley & Sons, 2005. 6. Han, J., Kamber, M., Data Mining: Concepts and Techniques, 1st Edition, Morgan Kaufmann, 2000. 7. 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). 		
8.2 Seminar / laborator	Metode de predare	Observații
1. Preprocesarea datelor	Proiecte practice	
2. Data Mining software tools		
3. WEKA software		
4-6. Aplicații Data Mining		
7. Prezentările proiectelor		
Bibliografie:		

1. S. Chakrabarti et al, Data Mining. Know It All, Morgan Kaufmann, 2009.
2. K. Cios, W. Pedrycz, R. Swiniarski, L. Kurgan, Data Mining. A Knowledge Discovery Approach, Springer, 2007.
3. J. Han, M. Kamber, Data Mining: Concepts and Techniques, 2nd Edition, Morgan Kaufmann, 2006.
4. P. Tan, M. Steinbach, V. Kumar, Introduction to Data Mining, Addison Wesley, 2006.
5. D. Larose, Discovering Knowledge in Data. An Introduction to Data Mining, John Wiley & Sons, 2005.
6. Han, J., Kamber, M., Data Mining: Concepts and Techniques, 1st Edition, Morgan Kaufmann, 2000.
7. 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).

9. Coroborarea con inuturilor disciplinei cu a tept rile reprezentan ilor comunit ii epistemice, asocia iilor profesionale i angajatori reprezentativi din domeniul aferent programului

- Acest curs exista in programul de studiu al tuturor universitatilor importante din Romania si strainatate
- Acest curs asigura cunostintele de baza pe care orice programator trebuie sa la aiba

10. Evaluare

Tip activitate	10.1 Criterii de evaluare	10.2 metode de evaluare	10.3 Pondere din nota final
10.4 Curs	- cunoasterea principiilor de baza ale domeniului	Examen scris	50%
10.5 Seminar/laborator	- rezolvarea de probleme de data mining	Teme laborator	50%
10.6 Standard minim de performan			
<ul style="list-style-type: none"> • Pentru promovare este necesara obtinerea notei minim 5 la examenul scris si la temele de laborator 			

Data complet rii

24.04.2013

Titular de curs

Lect. Dr. Anca Andreica

Titular de seminar

Lect. Dr. Anca Andreica

Data aviz rii în departament

.....

Director de departament

.....