

Adatbányászati technikák (BMEVISZM185)

Dr. Buza Krisztián Antal

Számítástudományi és Információelméleti Tanszék

http://www.cs.bme.hu/~buza/edu/dm_techn/index.html

e-mail: buza@cs.bme.hu

Órák

- Előadás:
 - Hétfő, 8:30-10:00, QB104 – minden héten (nincs szünet)
 - Csütörtök, 10:15-12:00, R509 – páratlan oktatási heteken (14 naponta)
- Gyakorlat
 - Kéthetente
 - 2 csoport (páros ill. páratlan oktatási heteken)
 - Csütörtök, 8:30-10:00, R3U (nincs szünet)
 - Első héten nincs gyakorlat, márc. 15-én is elmarad
 - **KATALÓGUS**: 6 alkalomból 1-szer lehet hiányozni

Tematika

- Alapfogalmak, előfeldolgozás...
- Osztályozás, numerikus predikció (regresszió)
- Klaszterezés
- Asszociációs szabályok
- ...
- és ezeket implementáló WEKA szoftvercsomag kezelése (gyakorlat), WEKA moduljainak beépítése **saját programunka**
- Részletes beosztás: **GoogleNaptár**

Követelmények

- 2 ZH (2x35 pont)
3 kisHF (3x10 pont) → 100 pont
 - mind a két zh-n külön-külön el kell érni 40%-t (14 pont)
- első ZH: márc. 26., második ZH: ápr. 23.
(előadás idejében)
- pótZH: május 7. - az írhat pótZH-t, aki a 2 zh közül legalább egyet sikeresen teljesített
- pótpótZH: pótlási héten
- **JELLENLÉT A GYAKORLATOKON
(KATALÓGUS!)**

ZH-król

- 4-5 feladat, kb. 80 perc munkaidő
- Fókusz: elmélet
 - (gyakorlati ismeretek számonkérése: kisHF-k)
- 4-5 feladat közül 1-2 típuspélda, többi elmélet
- ZH-k előtt „Gyakorlás”
 - többféle típuspéldát gyakorlunk, a ZH 1-2 feladata hasonló lesz a bemutatott típuspéldák valamelyikéhez

kisHF-k

- Gyakorlatai ismeretek számonkérése
- 3x10 pont
- **Programozás nélkül** egy-egy feladatra **legfeljebb 6 pont** kapható
- Kiadás: márc. 8., április 12., április 26.
- Beadás: folyamatosan,
de legkésőbb május 10-i laboron

Könyv

- Pang-Ning Tan, Michael Steinbach, Vipin Kumar:
Introduction to Data Mining
- <http://www-users.cs.umn.edu/~kumar/dmbook/index.php>

Introduction to Data Mining - Mozilla Firefox

Fájl Szerkesztés Nézet Előzmények Könyvjelzők Eszközök Súgó

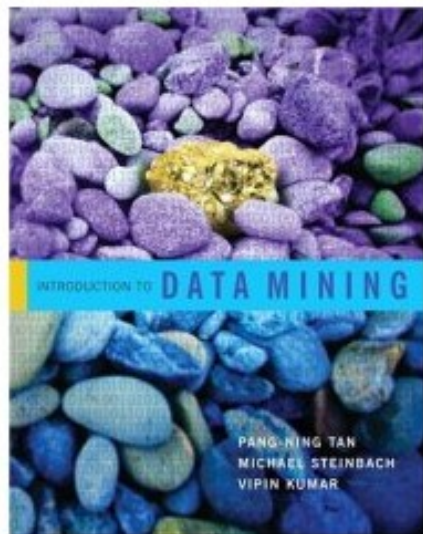
Introduction to Data Mining +

www-users.cs.umn.edu/~kumar/dmbook/index.php

Google

Legtöbbször látogatott Bevezetés Friss hírek Ingyenes Hotmail Javasolt helyek

Introduction to Data Mining



Pang-Ning Tan, Michigan State University,
Michael Steinbach, University of Minnesota
Vipin Kumar, University of Minnesota

- [Table of Contents](#)
- [Sample Chapters](#)
- [Resources for Instructors and Students](#)
- [Instructor Solution Manual](#)
- [Errata \(March 25, 2006\)](#)
- [Addison-Wesley Companion Book Site](#)

Contact info: dmbook@cs.umn.edu

Highlights:

- Provides both theoretical and practical coverage of all data mining topics.
- Includes extensive number of integrated examples and figures.
- Offers instructor resources including solutions for exercises and complete set of lecture slides.
- Assumes only a modest statistics or mathematics background, and no database knowledge is needed.

Sample Chapters:

These sample chapters are also available at the publisher's Web site.

- [Chapter 4. Classification: *Basic Concepts, Decision Trees, and Model Evaluation* \(444KB\)](#)
- [Chapter 6. Association Analysis: *Basic Concepts and Algorithms* \(612KB\)](#)
- [Chapter 8. Cluster Analysis: *Basic Concepts and Algorithms* \(1.3MB\)](#)

All files are in Adobe's PDF format and require [Acrobat Reader](#).

Resources for Instructors and Students:

[Link to PowerPoint Slides](#)

[Link to Figures as PowerPoint Slides](#)

[Links to Data Mining Software and Data Sets](#)

[Suggestions for Term Papers and Projects](#)

[Tutorials](#)

[Errata](#)

PowerPoint Slides:

1. **Introduction** (lecture slides: [\[PPT\]](#) [\[PDF\]](#))
2. **Data** (lecture slides: [\[PPT\]](#) [\[PDF\]](#))
3. **Exploring Data** (lecture slides: [\[PPT\]](#) [\[PDF\]](#))

Google Naptár (beosztás tervezete)

Google Naptár - Mozilla Firefox

Fájl Szerkesztés Nézet Előzmények Könyvjelzők Eszközök Súgó

Introduction to Data Mining Neptun.Net BME_W2K8O_D1 Adatbányászati Algoritmusok (fejlesztés ... Gmail - Inbox (6) - krisztian.buza@gmail... Google Naptár

google.com https://www.google.com/calendar/render?tab=mc&pli=1

Legtöbbször látogatott Bevezetés Friss hírek Ingyenes Hotmail Javasolt helyek

+Te Gmail Naptár Dokumentumok Fényképek Webhelyek Keresés Egyebek krisztian.buza@gmail.com

Google Keresés a Google Naptárban keresési lehetőségek megjelenítése

Naptár Ma < > 2012. február Ilap Hét Hónap 4 nap Ilapló Továbbiak

LÉTREHOZÁS

2012. február < >

h	k	sze	cs	p	szó	v	
30	31	febr. 1.	2	3	4	5	
6	7	8	9	10	11	12	
13	14	15	16	17	18	19	
20	21	22	23	24	25	26	
27	28	29	30	1	2	3	
4	5	6	7	8	9	10	11

Saját naptárak

- Krisztian Buza
- Adatbányászati techn...
- Teendők

Más naptárak

- Ismerős naptárának felvétel
- Ismerősök szület...
- Magyar névnapok
- Magyar ünnepnapok

08:15 Előadás - Bevezetés

10:15 Előadás - Adattípusok

08:15 Előadás - Aggregáció,

08:15 Labor1 - A csoport

08:15 Előadás - Feature Sele

08:15 Labor1 - B csoport

10:15 Előadás - Döntési fák

08:15 Előadás - Osztályozók I

08:15 Labor2 - A csoport