

Wright State University

## CORE Scholar

---

Computer Science & Engineering Syllabi

College of Engineering & Computer Science

---

Fall 2007

### CS 705: Introduction to Data Mining

Guozhu Dong

*Wright State University - Main Campus*, [guozhu.dong@wright.edu](mailto:guozhu.dong@wright.edu)

Follow this and additional works at: [https://corescholar.libraries.wright.edu/cecs\\_syllabi](https://corescholar.libraries.wright.edu/cecs_syllabi)



Part of the [Computer Engineering Commons](#), and the [Computer Sciences Commons](#)

---

#### Repository Citation

Dong, G. (2007). CS 705: Introduction to Data Mining. .  
[https://corescholar.libraries.wright.edu/cecs\\_syllabi/234](https://corescholar.libraries.wright.edu/cecs_syllabi/234)

This Syllabus is brought to you for free and open access by the College of Engineering & Computer Science at CORE Scholar. It has been accepted for inclusion in Computer Science & Engineering Syllabi by an authorized administrator of CORE Scholar. For more information, please contact [library-corescholar@wright.edu](mailto:library-corescholar@wright.edu).

# CS 705 Introduction to Data Mining

## Fall Quarter, 2007

**Description:** Data mining is concerned with the extraction of novel knowledge from large amounts of data. This course introduces and studies the concepts, issues, tasks and techniques of data mining. Topics include data preparation and feature selection, association rules, classification, clustering, evaluation and validation, scalability, spatial and sequence mining, privacy, and data mining applications. 3 hours lecture, 2 hours lab.

**Prerequisite:** CS 605 (Introduction to Database Systems), or CS 609 (Introduction to AI), or equivalent, or with consent of the instructor. Implicitly, CS 600 (Data Structures) is also required.

**Instructor:** Dr. Guozhu Dong. 383 Joshi.

**Phone & Email:** (937)-775-5066, guozhu.dong@wright.edu

**Class details:** 6:05 – 7:20, TTh, Millet 399

**Office hours:** 2:40 - 3:30, TTh. Use e-mail for short questions.

**Text Book:** Data Mining: Concepts and Techniques, 2nd edition. J. Han and M. Kamber. Morgan Kaufmann.

**Reference texts:** Data Mining: Practical Machine Learning Tools and Techniques, I.H. Witten and E. Frank, Morgan Kaufmann, 2nd Edition

Introduction to Data Mining, Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Addison Wesley, 1st Edition

Principles of Data Mining. D. Hand, H. Mannila, and P. Smith. MIT Press.

**Resources:** Class directory: Uptodate slides will be provided at [www.cs.wright.edu/~gdong/cs705](http://www.cs.wright.edu/~gdong/cs705)

Students may find this webpage useful: <http://www.kdnuggets.com/>, especially its pointers to datasets.

There are many Java programs for data mining at [www.cs.waikato.ac.nz/ml/weka](http://www.cs.waikato.ac.nz/ml/weka).

**Grading:** Homeworks: 10%, Midterm: 25%; Final: 35%; Projects 30%.

Final grade: A=[90,100], B=[80,90), C=[70,80), D=[60,70), F=[0,60).

The projects require extensive programming. Submissions that do not compile or that do not address project requirements will receive zero or very low marks.

**Handouts:** Handouts, and other course material will be distributed in class. It is the students' responsibility to collect them. The instructor plans to make these, plus slides masters, available in the course directory.

### Important dates:

6:05 – 7:20, Tuesday, 10/2, in class midterm.

8:00 – 10:00, Thursday, 11/15: Final.

There is no class on October 30, as the instructor will attend a conference.