



# CALIFORNIA STATE UNIVERSITY, CHICO

## CSCI 682: Topics in A.I. (Data Mining: Knowledge Discovery in Databases) Abbreviated Syllabus for Spring Semester 2008

Visit <http://www.ecst.csuchico.edu/~juliano/csci682/DataMining> for additional detail.

### Prerequisites

- Graduate standing; or
- Permission of instructor

### Description

3 units. This course introduces the student to basic concepts, tasks, methods, and techniques in data mining; in particular, the course focuses on practical machine learning tools and techniques used in data mining. Students will develop an understanding of the data mining process and issues, learn various techniques for data mining, and apply the techniques in solving data mining problems using data mining tools and systems.

**“Data mining, also known as knowledge-discovery in databases (KDD), is the practice of automatically searching large stores of data for patterns. To do this, data mining uses computational techniques from statistics and pattern recognition.”**

From [http://en.wikipedia.org/wiki/Data\\_mining](http://en.wikipedia.org/wiki/Data_mining)

**Note:** Students from departments such as Statistics, Biology, Mathematics, and Electrical & Computer Engineering who are working in interdisciplinary research (e.g., bioinformatics, modeling, data analysis) are especially encouraged to take this course.

Class #	Section	Act	Days	Time	Room	Instructors
6331	CSCI 682-01	DIS	TR	1230-0145pm	OCNL 239	Dr. J Juliano@csuChico.edu

### Instructor Information

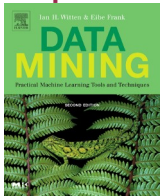
Dr. Juliano (a.k.a. Dr. J)  
<http://www.ecst.csuchico.edu/~juliano>

Office Hours: OCNL 222, T.B.A.  
Tel 530 898-4619 / 6442 (dept)  
Fax 530 898-5995  
*Appointments / walk-ins welcome.*

source data mining software application written in Java and developed at the University of Waikato in New Zealand.

3. Students are expected to familiarize themselves with Dr. J's general policies and expectations detailed at [/~juliano/Teaching/Policies.html](http://~juliano/Teaching/Policies.html)

### Required Textbook



*Data Mining: Practical Machine Learning Tools & Techniques, 2/e*  
Ian Witten and Eibe Frank, 2005.  
Elsevier Inc., Burlington, MA.  
ISBN 0-12-088407-0

### Grade Evaluation

60%	Written homework, assignments, and Laboratory (WEKA) projects
35%	Research paper (components and other deliverables; presentation: oral & written)
5%	Class participation

*Also see the on-line syllabus for details of final grade calculation.*

### Additional Requirements

1. Students are expected to open and maintain a *Chico State Connection* (CSC) Portal (see <http://portal.csuchico.edu>) account in order to access up-to-date *WebCT* tools that include an on-line calendar of events, current scores, discussion board, etc.
2. Students will also be using **WEKA**, the **W**aikato **E**nvironment for **K**nowledge **A**nalysis, an open

### Selected Topics for Research Papers:

Web search and Web mining, data mining for fraud detection, mining text and sequential data, data mining in bioinformatics (genomics, proteomics, etc).

### Additional Information

<http://www.ecst.csuchico.edu/~juliano/csci682/DataMining>  
<http://www.ecst.csuchico.edu/~juliano/DM/>  
<http://vista.csuchico.edu/>