

	CSC869: Data Mining
Class Meetings	Every Spring semester
Prerequisites	A grade C or better in CSC 510 and CSC675; or consent of instructor.
Course Websites	Primary: http://cs.sfsu.edu/~huiyang/869-????/ Secondary: http://ilearn.sfsu.edu (submissions, grades, and discussion forum)
Course Description	<ul style="list-style-type: none"> • Introduction to the knowledge discovery process (KDD) process and basic statistics. • Data preprocessing: missing data, noisy data, discretization, and data transformation. • Association pattern mining and applications. • Classification and prediction: decision trees, Naive Bayesian classifier, ensemble learners • Clustering analysis and anomaly detection: k-means, hierarchical, and density based clustering. • Data mining applications originated from different application domains, including text data, web data, stream data, graph-structured data, time series, spatial data, bioinformatic data, fraud detection, and so on.
Main Objectives	<ul style="list-style-type: none"> • Master the KDD process and basic concepts. • Master key data mining algorithms, including association rule mining, classification and prediction, clustering, and anomaly detection. • Familiarize with applying data mining algorithms to a variety of data types, such as text data, web data, spatial data, graphical data, and bioinformatics data. • Familiarize with scalable data mining algorithms, including parallel algorithms, database integration, and data locality issues. • Gain hands-on experience by applying data mining techniques to solve real problems.
Textbooks	<p><i>Required:</i></p> <p><i>Data Mining Concepts and Techniques (Second Edition)</i> by J. Han and M. Kamber, ISBN: 1-55860-901-6, Publisher: Morgan Kaufmann</p> <p><i>Recommended:</i></p> <p><i>Introduction to Data Mining</i>, by Pang-Ning Tan, Michael Steinbach and Vipin Kumar, ISBN-10: 0321321367, ISBN-13: 9780321321367, Publisher: Addison-Wesley</p> <p><i>Principles of Data Mining</i> by David J. Hand, Heikki Mannila and Padhraic Smyth, ISBN-10: 0-262-08290-X, ISBN-13: 978-0-262-08290-7, Publisher: MIT Press</p>
Primary Teaching & Learning Methods	<ul style="list-style-type: none"> • <i>Instructor:</i> Traditional in-class lectures, Q&A session during office hours or by appointment • <i>Students:</i> active class participation, homework assignments, take-home midterm exam, annotation/presentation/discussion of research papers (~8 papers in total), and term projects. • <i>Graduate seminar series:</i> 5:30pm--6:30pm, most Wednesdays. Submit a short summary after each seminar to earn 0.4 bonus points. • <i>Guest lectures:</i> TBA
Term Project	<p>Each student is required to decide on a term project no later than the 4th week. You can work on the project individually or with a teammate. Recommended projects are available on the website.</p> <p>Deliverables include:</p> <ul style="list-style-type: none"> • A one page project description by the 4th week

	<ul style="list-style-type: none"> • A brief progress report in the form of class presentation starting from the 6th week • A final project report of no longer than 10 pages. 																										
Homework Assignments	<p>Assignments will be in two main formats: a set of problems to examine your understanding of basic concepts, and mini-projects requiring you to implement an algorithm and then apply it to a toy (or real) dataset. Homework must be turned in at the beginning of the class on the due date. No late homework will be accepted.</p>																										
Mid-term Exam	<ul style="list-style-type: none"> • A hands-on project that asks you to (1) implement 1-2 algorithms, (2) apply these algorithms to analyze datasets drawn from different application domains. • Take-home • During the 7th or 8th week. 																										
Final Exam	<p>In lieu of a written exam at the end of the term, you are required to (1) demonstrate your term project to the instructor, (2) deliver a 10 minutes presentation of your term project, and (3) submit a final project report using the Times font, 12 point font size, with a page limit of 10 pages.</p>																										
Grading Policy	<ul style="list-style-type: none"> • 5% Class participation • 35% 3-4 mini-projects • 25% Annotation, presentation, and discussion of research papers • 35% Term project (by individuals or teams or 2 members) <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>Percent</th> <th>Grade</th> </tr> </thead> <tbody> <tr> <td>> 92%</td> <td>A</td> </tr> <tr> <td>89% - 92%</td> <td>A-</td> </tr> <tr> <td>86% - 88%</td> <td>B+</td> </tr> <tr> <td>82% - 85%</td> <td>B</td> </tr> <tr> <td>79% - 81%</td> <td>B-</td> </tr> <tr> <td>76% - 78%</td> <td>C+</td> </tr> <tr> <td>72% - 75%</td> <td>C</td> </tr> <tr> <td>69% - 71%</td> <td>C-</td> </tr> <tr> <td>66% - 69%</td> <td>D+</td> </tr> <tr> <td>62% - 65%</td> <td>D</td> </tr> <tr> <td>59% - 61%</td> <td>D-</td> </tr> <tr> <td>< 59%</td> <td>F</td> </tr> </tbody> </table>	Percent	Grade	> 92%	A	89% - 92%	A-	86% - 88%	B+	82% - 85%	B	79% - 81%	B-	76% - 78%	C+	72% - 75%	C	69% - 71%	C-	66% - 69%	D+	62% - 65%	D	59% - 61%	D-	< 59%	F
Percent	Grade																										
> 92%	A																										
89% - 92%	A-																										
86% - 88%	B+																										
82% - 85%	B																										
79% - 81%	B-																										
76% - 78%	C+																										
72% - 75%	C																										
69% - 71%	C-																										
66% - 69%	D+																										
62% - 65%	D																										
59% - 61%	D-																										
< 59%	F																										
Missed Exams	<p>In general, there will be no make-up exams. However, they can be given in case of uncontrolled events (e.g., medical emergency) or if the exam dates conflict with your religious observances. In the former case, written documents are required. In the latter case, it is your responsibility to notify the instructor within the first four weeks</p>																										

PLEASE RESPECT YOUR FELLOW STUDENTS: USE ELECTRONIC DEVICES WISELY IN THE CLASSROOM.