San José State University
College of Science/Department of Computer Science
CS 274, Topics in XML and Web Intelligence, Section 1, Spring, 2018

Course and Contact Information

Instructor:	Dr. Teng Moh
Office Location:	MQH 411
Telephone:	(408) (924-5147)
Email:	MyFirstName <dot> MyLastName <at> SJSU <dot> EDU
Office Hours:	MW 20:45 to 21:45 *[See University [Policy S12-1](http://www.sjsu.edu/senate/docs/S12-1.pdf) at http://www.sjsu.edu/senate /docs/S12-1.pdf for guidelines]*
Class Days/Time:	MW 19:30 to 20:45
Classroom:	MQH 422
Prerequisites:	CS 160 or instructor consent. Familiarity with XML is assumed.

Course Description

XML: DTD, Schema, Namespace, XSLT, XPath, Xquery, Encryption, Signature in XML, applications in vertical industries; Semantic Web; RDF, RDFS, Ontology, Inferences; Web services, relevant tools for search, inference, data conversion in XML, Semantic web applications. This section focuses specifically on Semantic Web and Social Networks.

Course Learning Outcomes (CLO)

Upon successful completion of this course, students will be able to:

1. CLO 1 *understand Finding Similar Items*
2. CLO 2 *understand Mining Data Streams*
3. CLO 3 *understand Link Analysis*
4. CLO 4 *understand Frequent Itemsets*
5. CLO 5 *understand Advertising on the Web*
6. CLO 6 *understand Recommendation Systems*
7. CLO 7 *understand Mining Social-Network Graphs*
8. CLO 8 *implement new tools for web intelligence based on the above techniques*

Required Texts/Readings

Textbook

§ *Jure Leskovec, Anand Rajaraman, Jeff Ullman,* **Mining of Massive Datasets,** *Cambridge University Press, 2nd ed., 2014, ISBN: 978-1107077232*

Other Readings [Optional]

§ *Ian H. Witten, Eibe Frank, and Mark A. Hall,* **Data Mining***, Morgan Kaufmann, 3rd ed., 2011, ISBN: 978-0123748560*

Other equipment / material requirement

§ *Wireless laptop*

Course Requirements and Assignments

SJSU classes are designed such that in order to be successful, it is expected that students will spend a minimum of forty-five hours for each unit of credit (normally three hours per unit per week), including preparing for class, participating in course activities, completing assignments, and so on. More details about student workload can be found in University Policy S12-3 at http://www.sjsu.edu/senate /docs/S12-3.pdf.

Homework is due typewritten (include source code, but not executable files) by class starting time on the due date. Each assigned problem requires a solution and an explanation (or work) detailing how you arrived at your solution. Cite any outside sources used to solve a problem. When grading an assignment, I may ask for additional information. A subset of the assigned problems will typically be graded.

Refer the course website for latest information of homework assignments.

NOTE that University policy F69-24 at http://www.sjsu.edu/senate/docs/F69-24.pdf states that "Students should attend all meetings of their classes, not only because they are responsible for material discussed therein, but because active participation is frequently essential to insure maximum benefit for all members of the class. Attendance per se shall not be used as a criterion for grading."

Examinations

One mid-term exam scheduled approximately at the end of 9th week, and a final exam scheduled on Wednesday, May 16, 19:45-22:00.

Grading Policy

Grading information:
- *I will determine letter grades for the course, including +/- grades based on*

Percentage	Grade
92 and above	*A*

90 - 91   *A-*
88 - 89   *B+*
82 - 87   *B*
80 - 81   *B-*
78 - 79   *C+*
72 - 77   *C*
70 - 71   *C-*
60 - 69   *D*
59 and below   *F*

- *List of the percentage weight [or point value] assigned to various class assignments*
  - *Homework: 20%*
  - *Midterm (Wednesday, Mar. 21): 20%*
  - *Project: 30%*
  - *Final (Wednesday, May 16, 19:45 – 22:00): 30%*
- ***No*** *make-up exams will be given and **no** late homework will be accepted.*

Classroom Protocol

§ ***Always start your email subject with "CS274" to get my attention.***

§ ***Wireless laptop is required****. Your laptop must remain closed (preferably in your backpack and, in any case, not on your desk) until I inform you that it is needed for a particular activity.*

§ ***Cheating*** *will not be tolerated, but working together is encouraged*

§ *Student must be respectful of the instructor and other students. For example, but not limited*

     o *Turn off cell phones*

§ *To encourage participation from students, **no** recording is allowed.*

University Policies

Per University Policy S16-9, university-wide policy information relevant to all courses, such as academic integrity, accommodations, etc. will be available on Office of Graduate and Undergraduate Programs' [Syllabus Information web page](http://www.sjsu.edu/gup/syllabusinfo/) at http://www.sjsu.edu/gup/syllabusinfo/"

CS 274, Spring 2018, Course Schedule

The schedule is subject to change with fair notice and the notice will be made available in class.

Course Schedule

Week	Topics, Readings, Assignments, Deadlines
1	Shingling & Min-Hashing
2	Locality-Sensitive Hashing & Sampling a Fixed-Size Sample
3	Queries Over a Sliding Window & DGIM Method
4	Web Search & PageRank
5	Block-Based Update Algorithm & Frequent Itemset Mining
6	Association Rules & A-Priori Algorithm
7	Multistage Algorithm & Performance-Based Advertising
8	BALANCE Algorithm & Content-based Recommender Systems
9	Review & Midterm Exam (Wed., 3/21)
10	Collaborative Filtering
11	Community Detection & Spectral Clustering
12	Dimensionality Reduction
13	Applications of Mining Data Streams & Link Analysis
14	Applications of Frequent Itemsets & Advertising on the Web
15	Applications of Recommendation Systems & Mining Social-Network Graphs
16	Review
Final Exam	Wednesday, May 16, 19:45-22:00