

## **SYLLABUS: CSE 591 - Semantic Web Mining**

### **1. Contact Information**

Instructor: Paulo Shakarian  
Office: BYENG 408  
Telephone: (480) 727 - 5290  
E-Mail: [shak@asu.edu](mailto:shak@asu.edu)

### **2. Office Hours**

Fridays, 12:00-1:30pm

*Students should make every effort to utilize the scheduled office hours. However, if you are unable to attend these office hours, please attempt to schedule an appointment at least 24 hours in advance.*

### **3. Course Objectives and Expected Learning Outcomes**

Data mining techniques for structuring and organizing unstructured sources such as text and Web data into meaningful machine-processable information; computational aspects of information extraction and data linkage; discovery and prediction tasks where text serves as data such as detecting events, measuring public opinion and making recommendations.

Prerequisites: This is an advanced course intended for Computer Science and Computer Engineering graduate students with background in databases, algorithms and theory of computation.

Course Outline (tentative):

- Week 1. Logic programming and rule learning
- Week 2. [Chapter 1] Regular expressions and crawling the web
- Week 3. Social network analysis and diffusion
- Week 4. [Chapter 4.5] *Collaborative Filtering*
- Week 5. [Chap. 4] *Similarity and co-Clustering for Text and Graphs*
- Week 6. [Chapter 7.5] *Topic Distillation*
- Week 7. [Chapter 9.1 9.2] *Information Extraction & NLP*
- Week 8. [Reading List] *Semi-structured Data: XML, XSLT, XPATH, HTML*
- Week 9. [Chapter. 5] *Supervised Learning for Text Classification*
- Week 10. [Reading List] *Sentiment Analysis*
- Week 11. [Reading List] *Streaming Data - Trend Detection*
- Week 12. [Chapter 8] *Resource Discovery*
- Week 13-14. *Final Project Presentations*
- Week 15. Review s

#### 4. Grading Policies

Project proposal	10%
Final Project Report	30%
Graded homework (3x)	30%
Final Exam	30%

## 5. Absence & Make-Up Policies

Accommodations will be made for religious observances provided that students notify the instructor at the beginning of the semester concerning those dates. Students who expect to miss class due to officially university-sanctioned activities should inform the instructor early in the semester. Alternative arrangements will generally be made for any examinations and other graded in-class work affected by such absences.

## 6. Readings, Assignments, Examinations, Special Materials, Required Activities

Required Text: *Mining the Web* by Soumen Chakrabarti. Morgan Kaufmann

Optional: *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*, Liu, Springer

Reading List: [http://www.public.asu.edu/~hdavulcu/CSE591\\_Semantic\\_Web\\_Mining.html](http://www.public.asu.edu/~hdavulcu/CSE591_Semantic_Web_Mining.html)

**Project Assignments:** Students will be grouped into project teams of 4 or 5 students. They will be asked to design and implement components of an end-to-end integrated semantic web mining application which use social media (such as Twitter, Facebook) streams and online interfaces for capturing the user preferences and to identify new relevant items matching a user's interests. The system will utilize the profiles of similar other users and relevant online sources to create a personalized recommendation service (with web-based or mobile device based user interfaces) that would facilitate the discovery of new items (such as news, products or services). Various components that will be developed by project teams include: RSS and social media feed readers; faceted search interfaces; named-entity detectors; trending topic detectors; classifiers; personalization and recommendation algorithms.

## 7. Classroom Behavior

Cell phones and pagers (must be/or state alternative rule) turned off during class to avoid causing distractions. The use of recording devices (is/is not) permitted during class. Any violent or threatening conduct by an ASU student in this class will be reported to the ASU Police Department and the Office of the Dean of Students.

## 8. Academic Integrity

<http://provost.asu.edu/academicintegrity>) and should acquaint themselves with its content and

requirements, including a strict prohibition against plagiarism. All violations will be reported

□ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □

9. Disability Accommodations.

Suitable accommodations will be made for students having disabilities and students should notify the instructor as early as possible if they will require same. Such students must be registered with the Disability Resource Center and provide documentation to that effect.

Notice: Any information in this syllabus (other than grading and absence policies) may be subject to change with reasonable advance notice.