Syllabus

COSC 7397: Text Mining with Applications to Biomedicine/Security

Instructor: Rakesh Verma, 532PGH, 3-3348

Office Hrs: Tu/Th 1.15-1.45pm (or by appt.)

Spring 2012

Due to the deluge of information on the world-wide web, recently there has been increasing attention and some notable successes on using computers to summarize and/or extract information from text documents. This course will examine state of the art techniques for extracting and analyzing information from text. Two application areas that will motivate the course are: (a) scientific articles especially in the area of biomedicine, and (b) computer security (e.g. automatic phishing detection). Topics will include word-sense disambiguation, summarization, question/answering, extracting relations from text, and knowledge discovery. This is a seminar course and emphasis will be on reading current literature and presenting/discussing it.

Grading (weights are approximate and subject to change): Reading quizzes 20%, Presentations, Discussion and Scribing 40% and Term Paper/Project 40%.