CS 6890 Syllabus
Social Computing and Data Management
Utah State University, Fall 2013
Tuesday 5:15pm - 7:45pm, MAIN 117 with Broadcast

Instructor:
Kyumin Lee, MAIN 401E, (435) 797-8420, kyumin.lee@usu.edu
Office Hours: 10:00-11:00am Tu/Th, or by appointment

Course Summary:
This course introduces advanced research topics in social computing and data management, addressing the relevant theoretical foundations, methods, and tools from a wide spectrum, including (i) large-scale information management, (ii) data and text mining techniques and algorithms; and (iii) data privacy and security issues in large-scale social systems.

By the end of the semester you will be able to:
• Define and explain the key concepts, methods, and tools relevant to social computing and data management
• Design, implement and evaluate core algorithms and approaches including distributed text mining, large graph manipulation and analysis, etc.
• Identify the salient features and apply recent research results in social computing

Communication:
All course announcements will be posted to the course mailing list. If you have a specific question for me, please send me an email with 6890 in the subject line.

Prerequisites:
Graduate classification or approval of instructor.

Required Text:
No primary textbooks, but we will refer to the following textbooks:

In addition, we will read some research papers from top computer science conferences and journals.

Class Format:
During the first few weeks, I will present an overview of social computing and web data management. The remaining weeks will be devoted to intense paper discussion.

Grading:
The course grading policy is as follows:
15% In-class discussion and emailed questions
15%  Paper presentation  
20%  Midterm  
15%  Project proposal  
15%  Project presentation  
20%  Final project deliverable  


**Midterm:** the midterm exam is closed book and will be held in class. You may bring one standard 8.5" by 11" piece of paper with any notes you think appropriate or significant (front and back). No electronic devices allowed.

The most significant portion (50%) of your grade is based on the project, for which you may work in teams of up to two persons. The general goal of research is to contribute new knowledge. Thus, it is important to ask yourself what research question(s) you aim at answering and what challenges you aim at solving.

In general, a project should include both: (i) an empirical evaluation of an algorithm or model on an interesting dataset, in order to better understand these methods, and possibly further improve them; and (ii) a proposal for a new algorithm and model, including a comparison with a baseline.

When picking a topic, try to ask yourself the following questions:

- Is the topic addressing an important problem? Would any one care about it if you solve the problem?
- To what extent has previous research work addressed this problem? And what remains unknown?
- Do you have any idea at all about how to solve the problem? If not, can you reformulate the problem to make it easier?
- Would you be able to evaluate your solution? That is, how can you demonstrate your solution is good and solves the problem well?

**Proposal:** The emphasis of the course is on recent and current scholarly material. We will read a number of papers from the past few years and these papers should stimulate you to think about unexplored avenues of research. For the project proposal, you should pick two or three related papers as the basis for your proposal. These related papers will serve as the research grounding for your proposal. So first in your proposal, you must address these questions (in about 2 or 3 pages, single column, normal fonts and margins):

- What is the key technical content and interesting ideas behind the papers?
- How do the papers relate to the topics presented in the course?
- What are the strengths of the paper? What are its weaknesses? And how might you go about improving on the weaknesses?

In the remainder of the proposal (about 2 pages), I expect you to dig deeply into your proposed work (which should naturally build on the background materials you’ve already covered).

- What is your research question? Clearly define the research problem/question.
• Why is this an interesting question to ask and why would we care about the answer to this question or a solution to the problem?
• Has any existing research work tried to answer the same or a similar question, and if so, what is still unknown?
• How do you plan to work out the answer to the question. (At the proposal stage, you are only expected to have a sketch of your methods.)
• How would you evaluate your solution. That is, how do you plan to demonstrate that your solution/answer is good or is reasonable.
• A rough timeline to show when you expect to finish what. List a couple of milestones.

You should plan to write a proposal of around 5 pages. The project proposal is due on September 27 (Friday) by 11:59pm. Email me your project proposal in PDF with a pithy filename. One per team. Final Project Deliverable: At the end of the semester, you will deliver a final paper and participate in a presentation (maybe with a demo) session.

• The presentation session will be held on December 3, 2013. 15% of your project grade is based on the project presentation. Your team’s grade will be based on the clarity of the presentation itself, audience feedback, and your answers to my questions.
• The final paper is due on December 6, 2013 by 11:59pm. You should email me your final paper in PDF with a pithy filename. The paper should be in the standard ACM conference template http://www.acm.org/sigs/publications/proceedings-templates (double-column) and up to 8 pages maximum. The paper counts for 20% of your project grade. You should write your report as if you were writing a conference paper. You should address the same questions as those you have addressed in the proposal, only with more details, especially regarding some of the challenges that you need to solve and your experimental results. You should also include your conclusions from the study and point out how your work can be further extended (i.e., future work).

Code of Conduct
As a computer scientist, or someone taking a computer science class, you are expected to perform your work at all times in an ethical manner. This means that in addition to doing your own work and giving appropriate credit when the work of others is used, you are required to protect your work.

A student that protects their work will not allow another student access to that work whether it be allowing it to be copied, or treating its security in such a way as to give unintentional access, such as "accidental" loss. It is the policy of the department that when duplicate (essentially the same) work is turned in by two or more students, without acknowledgement of allowed cooperation, all involved students will be considered in violation of this department policy. Under such circumstances, each student will receive minus the points possible for the work. Thus, for a 15-point assignment, all would receive -15 points. If the infraction is deemed more egregious, then further action may be taken.

Plagiarism will not be tolerated in any form. This includes, but is not limited to, using solutions/answers found online, and purchasing solutions in any manner. Of course, you may use online resources to learn and study, but you MAY NOT submit the work as your own. This pertains to program code and to traditional written assignments/projects/etc. For additional information please visit: http://cs.usu.edu/htm/code-of-conduct/
ADA Compliance
Students with ADA-documented physical, sensory, emotional or medical impairments may be eligible for reasonable accommodations. Veterans may also be eligible for services. All accommodations are coordinated through the Disability Resource Center (DRC) in Room 101 of the University Inn, (435)797-2444 voice, (435)797-0740 TTY, or toll free at 1-800-259-2966. Please contact the DRC as early in the semester as possible. Alternate format materials (Braille, large print or digital) are available with advance notice.