Graph Data Mining (GDM): Theory, Algorithms, and Applications

DESCRIPTION
The discovery and forecasting of insightful patterns from graph data are at the core of analytical intelligence in government, industry, and science. This course teaches both basic and more advanced techniques required for routine analytical intelligence operations on graph data. Students will be exposed to the underlying theory and learn to design effective and efficient algorithms and data structures for dealing with huge volumes of complex and noisy graph data, as well as real-world applications.

PREREQUISITES
Undergraduate level knowledge of statistics, linear algebra, and programming. Basic knowledge of data mining and machine learning concepts as well as python is a plus. Otherwise, consent of the instructor is required.

LEARNING OUTCOMES
By the end of the course, students will have gained knowledge of different methods and techniques for graph mining and analytics, will be able to critically analyze the pros and cons of applying these techniques in different contexts, and will be aware of the applications that require such graph analytic techniques. Finally, students will be able to conceptualize and design efficient and effective graph mining solutions for different real world problems.

INSTRUCTOR
Dr. Nagiza Samatova <samatova@csc.ncsu.edu>
Phone: 865-566-5586 (Cell)
We could meet/Skype/WebEx by appointment.

TEACHING ASSISTANTS
Yifan Zhao <yzhao48@ncsu.edu>
Office hours are by appointment.
The best medium for communication is Piazza, but meetings can be arranged if necessary.

COMMUNICATION
- All questions/issues/suggestions must be posted on Piazza only.
- No emails will be answered; private posts are ok.
- Except for personal issues or assignment clarifications, TAs will wait 24 hrs before helping with questions to give students sufficient time to interact.

LECTURE TIME AND PLACE
Lectures are pre-recorded and available through engineeringonline.ncsu.edu. Please see the schedule on Piazza Resources for the course for the appropriate viewing order.
ASSIGNMENTS

- Individual unless stated otherwise.
- There will be approximately five homeworks and five projects. In addition, you must choose to complete either a Final exam or a Final Capstone Project. If you choose to do both, then one of them will be used as the bonus credit. The details on the bonus credits will be posted on Piazza later in the course.
- Each assignment due date is on the Assignment Schedule available on the Moodle course page. If you did not click the "submit" button, but uploaded your files and saved them in what is called a "draft" in Moodle, which we will still be able to view and grade.
- You are required to submit your own work (unless explicitly stated as a team-based assignment). However, you may discuss your approach to any assignment (except for the exam) with your classmates. In this case, you must provide the names of the students you discussed your approaches with. If you fail to mention your collaborators and similarity is detected there will be an automatic 20 percent deduction.
- Submission lockers will be located on Moodle.

GRADING

To reduce stress over grades and promote collaboration, we will adopt a grading system inspired by a number of US medical schools. The homeworks and projects will be graded as a binary pass or fail (P/F). If you complete roughly 80% of the assignment or more, you will be given a pass, namely, you met the learning objectives for this topic; otherwise, a fail. Ultimately, at the end of the course we will have to submit a letter grade. This letter grade will be determined as follows:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Ways to Achieve the Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>A+</td>
<td>Fail only one assignment (project or homework) and pass the final exam/capstone.</td>
</tr>
<tr>
<td>A</td>
<td>Fail only two assignments and pass the final/capstone.</td>
</tr>
<tr>
<td>B+</td>
<td>Fail only three assignments and pass the final/capstone.</td>
</tr>
<tr>
<td>B</td>
<td>Fail only four assignments and pass the final/capstone.</td>
</tr>
<tr>
<td>F</td>
<td>Fail five or more assignments or fail the final/capstone</td>
</tr>
</tbody>
</table>

As this is an advanced graduate level course, feedback will be implicit. Namely, if you receive a Pass, we will not be providing feedback, and it will be your responsibility to look into the provided solution to figure out if/where you might have made any mistakes. However, if you receive a Fail, we will provide feedback as to why you did not meet the learning objectives for that assignment.

IMPORTANT: We will not entertain grade disputes until the end of the semester during the last week of classes. If you feel an assignment was graded incorrectly, please, make a note for your record and bring up the issue to the TA/grader at that time.
**LATE POLICY**
Late assignments after the penalty days are not accepted, except for reasons specified in the N.C. State Academic Policy on Attendance Regulations. The web site on attendance policy has detailed instructions on how to document various kinds of excused absences.

**TEAMS**
- Some assignments may allow you to work in teams. If so, it will be explicitly stated.
- Team will be self-made; you will be allowed to pick your own teams of 2 people.
- If you do not want to work on a team you do not have to.
- You may change your team at any time, you do not have to inform or ask us.
- Only one submission per team, but include your team members’ names.

**STUDENTS WITH DISABILITIES**
Reasonable accommodations will be made for students with verifiable disabilities. In order to take advantage of available accommodations, students must register with the Disability Services Office at Suite 1900, Student Health Center, Campus Box 7509, 515-7653.

**ACADEMIC INTEGRITY**
All the policies outlined in the Code of Student Conduct (Section 8) apply. Please (scroll down) read this carefully and ask me if you have any questions.
The essence of University policy is: **Don’t do anything that will give you an unfair advantage over other students.** The Code of Student Conduct has a list of behaviors that are unacceptable.