



# CIS 600

## Social Media & Data Mining

Tuesday & Thursday, 14:00 - 15:20

Eggers Hall 010

Instructor: Martin Harrison  
Instructor Email: mharriso@syr.edu  
Office Location: CST 242-A  
Office Hours: 15:30 - 17:00, TR  
TA: Chunxu Tang  
TA Email: ctang02@syr.edu  
TA Office Location: CST 3<sup>rd</sup> Floor  
TA Office Hours: 15:00 - 16:30, MF  
Website: dataminingsu.slack.com

**Text:** There is no official text for this course. Papers and books will be provided digitally through BB and Slack.

### Course Description

Social media mining is a branch of data mining, with its focus on social media. Data mining refers to the process of discovering patterns, or extracting knowledge, from data. Accordingly, social media mining refers to the process of discovering patterns, or extracting knowledge, from social media. In this course, we will introduce popular methods for discovering patterns and extracting knowledge from social media. In introducing these methods, we will balance theory and practice. We will explain the key concepts and theoretical foundations of these methods and we will get hands-on experience by developing applications with publicly available software.

### Course Objectives

Students should understand the technical challenges as well as the algorithms (and theoretical foundations thereof) involved in extracting and analyzing social media data. Students will develop social media mining applications.

### Course Outcomes

At the completion of this course, students will be able to:

- Understand six types of social media.
- Use social network APIs together with the Python languages to pull social media data.
- Understand and describe key concepts and principles of social network analysis, including centrality, influence and community detection.
- Analyze social media structure as well as user attributes and behaviors.
- Perform sentiment analysis using natural language processing methods.
- Develop social media mining applications.

## Topics Covered

Below are topics to be covered, not precisely in order of appearance.

- Types of Social Media
- Python
- Twitter API
- Social Networks - graphs, centrality, influence, community
- NLP & Sentiment Analysis
- Predictive Modeling
- Recommendation Systems
- Facebook API
- Google+ API
- Slack API

## Term Project

The term project deliverables are:

- A slide presentation
- A report describing your approach, data, experimental results and analysis. Single-spaced, font size at most 12.
- Source code (well-commented!!), data files, URLs, etc.
- A README file describing how to install, run and use your system. If your code is available online, give a URL in the README.

## Grade Breakdown

Your final course grade will be calculated using the following weights.

Participation	10% (Attendance is mandatory)
Homework	10%
Midterm	40%
Term Project	40%

## Academic Honesty Policy Summary: Do not plagiarize or commit fraud.

*Fraud* includes faking answers as if they are correct. *Plagiarism* includes using the work of another without proper credit.

Any student involved in fraud or plagiarism will fail the course and will be reported to the dean and director.

*If you have any doubts, then don't do it!*

Please refer to the SU Academic Integrity policy: [http://supolicies.syr.edu/ethics/acad\\_integrity.htm](http://supolicies.syr.edu/ethics/acad_integrity.htm)

## Accommodations for Disability

If you need accommodations because of a disability, the contact both the instructor and the Office of Disability Services, 804 University Avenue, Suite 303, 315-443-4498 (voice) or 315-443-5019 (TDD).