



SIMON
BUSINESS SCHOOL
UNIVERSITY of ROCHESTER

CIS 434: Social Media and Text Analytics
Course Syllabus (last modified 08/25/2021)
Fall A 2021

Course Logistics:

Day and Time:

Tuesday/Thursday	10:20 am – 11:50 am	lectures and in-class labs
Thursday	09:00 am – 10:00 am	labs

Location: S103

Course Blackboard Link: <https://learn.rochester.edu>

Zoom link: <https://rochester.zoom.us/my/cis434.11a>

Instructor:

Weiguang Wang 585.275.0804
Assistant Professor wwang90@simon.rochester.edu

Lab Instructor:

Junyuan Ke junyuan.ke@simon.rochester.edu

Teaching Assistants:

Fanruo Wang fanruo.wang@simon.rochester.edu

Office Hours:

Weiguang Wang	Friday	09:00 am – 10:00 am	https://rochester.zoom.us/my/cis434.11a
Junyuan Ke	Wednesday	02:00 pm – 03:00 pm	https://rochester.zoom.us/my/cis434.11a
Fanruo Wang	Monday	02:00 pm – 03:00 pm	https://rochester.zoom.us/my/cis434.11a

Course Information

Course Description and Learning Objectives

Overview: The rise of social media has fundamentally changed many aspects of business by empowering customers and the general public in an unprecedented way. They are well connected with each other through platforms like Facebook and Twitter. While text and multimedia channel substantially enable rich interactions, information can be distributed freely and publicly to large audiences in real time. This media revolution not only forces companies to actively manage their presence and engage with customers on social media platforms but also offers them a golden opportunity to extract intelligence from the vast amount of unstructured data, including text, image, audio, video, etc. Technology and strategies are increasingly intertwined in this new frontier of innovation and competition.

This course draws on a unique blend of social media data and the rapidly expanding supporting technologies with an emphasis on text analytics. While we mainly focus on topics related to the monitoring & analyzing of social media, we will also discuss applications of text analytics in other business settings where extracting intelligence from textual data has become increasingly important. Indeed, according to Wikipedia, text data is the primary form of unstructured data which is the source of 80 percent of business-relevant information.

The learning objectives of this course include: (1) getting comfortable with the management of text data; (2) being able to understand and apply commonly used methods to analyze text data; and (3) becoming familiarized with concepts and applications of the state-of-the-art technologies in text analytics.

Relation with other Courses at Simon Business School

This course is designed for students with entry level skills based on the curriculum of MS of Business Analytics and MS of Marketing Analytics. If you have less experience of programming, please be aware that the permission from instructor is needed and extra efforts might be required. If you are experienced in programming and machine learning, please make sure this course meets your expectation, as too complex models won't be covered for the best learning outcome of most students.

- **Programming Languages:** We will use Python in this course. Prior knowledge of Python is required.
- **Modeling:** Entry level or zero experience of NLP and Machine Learning are expected. Complex models will be discussed but not integrated in the labs. However, any modeling skills will be useful for the homework and final project.
- **Business theories:** This course will not discuss business strategies. However, the sense of business concepts, practice, etc. will be useful in homework and the final project as we are working on real social media data.
- **Programming Environment:** All students should install Python3 (preferably 3.6 or above) and your preferred IDE or any other coding interfaces such as Jupyter notebook, Sublime, or just text editor. Google Colab will be used by the teaching team for class demonstrations. Colab is also recommended to students with entry level of experience.

This is a hands-on course with many labs. Please have your laptop ready. Since we start the lab from Module 1, you are responsible to ensure Python runs smoothly on your machine before the first class.

Evaluation and Grading

The course grade is based on the following:

Homework assignments	50%
Team project	40%
Professionalism* and Class Participation	10%
Total	100%

Homework Assignments

The deadline of each assignment will be announced in class and on Blackboard. After the deadline, you will not be able to submit on Blackboard and no submission will be accepted afterwards. All homework assignments are individual based, and therefore must be the results of your own intellectual efforts.

Team Project

Teams can be self-formed. Team size is 5, and exceptions can only be made for the last team to accommodate all students. Team members will receive the same grade for the final team project unless peer evaluations warrant otherwise.

Professionalism and Class Participation

Apparently, your attendance is of utmost importance to your learning experience for such a hands-on course. Students attending class should arrive on time and not disrupt class by talking loud, eating, or using cell phones. Please bring your name tag in each class.

Accommodations and Accessibility

Both Simon and the University of Rochester respect and welcome students of all backgrounds and abilities. In the event you encounter any barrier(s) to full participation in this course due to the impact of a disability, please contact both your instructor and the Access Coordinator for Simon (in OSE), during the first two weeks of the course or sooner.

Additionally, you can contact the University's Office of Disability Resources. The access coordinators in the Office of Disability Resources can meet with you to discuss the barriers you are experiencing and explain the eligibility process for establishing academic accommodations. You can reach the Office of Disability Resources at disability@rochester.edu.

Credit-Hour Policy Adherence

This course follows the Simon credit-hour policy for 2.5-credit courses. The course meets twice weekly for 3 hours per week. In addition to these 3 hours of synchronous class sessions per week (in-person or via Zoom), students are required to complete approximately one hour per week on average of asynchronous learning activities (e.g., videos and simulations). Please see the course schedule for the details related to asynchronous activities and any related assignments or assessments.

Students are also expected to supplement each hour of class or asynchronous learning activities with two hours of supplemental work (e.g., class readings, assignments, and project work).

Academic Integrity

Simon's Code of Academic Integrity (see the section Academic Integrity Policy in the [Simon School Student Handbook](#)) states: "*Every Simon student is expected to be completely honest in all academic*

matters. Simon students will not in any way misrepresent their academic work or attempt to advance their academic position through fraudulent or unauthorized means. No Simon student will be involved knowingly, or unknowingly yet passively within a team, with another student's violation of this standard of honest behavior."

In addition to refraining from obvious forms of cheating and plagiarism:

- On assignments, do not copy or paraphrase work from each other, from students who have taken the class previously, from materials of mine distributed in a previous class, or from outside sources. Any written work should be entirely your own (or your team's, as applicable).
- Do not obtain advice, notes, solutions, or other material from students who took the class previously in ways that would give you an unfair advantage or would undermine the learning experience for you and the class (such as, notes from past case discussions). Similarly, do not use others' case analyses posted on-line.
- Use quotation marks when quoting any text directly. Changing a few words of a sentence or longer section does not make the work your own. Independently written texts rarely have even five consecutive words in common.

Most forms of disallowed shortcuts are easy to detect and will be referred to the school's Academic Integrity Committee. To help prevent other students from violating academic integrity, do not pass on notes or give advice on assignments to any students who are taking the course in a later term or are taking it at the same time in a different section. Please refer to the Student Handbook for any questions regarding the Code of Academic Integrity.

If a situation in your professional or personal life prevents you from finishing assigned work in a timely manner, please contact me before the deadline to discuss how to proceed. You should also contact your OSE advisors, if appropriate. Do not violate the Academic Integrity Code in an attempt to manage a difficult situation.

Professionalism

Please review the Professional Standards Policy in the Simon School Student Handbook. Students are expected to demonstrate the same professional behavior in class as they would in a business setting.

Unprofessional behavior has a negative impact on your participation grade. Specifically,

- Make every effort to attend each class.
- Display your printed name tent in every class.
- You are expected to be in your seat and ready for class at the beginning of each class. Should extenuating circumstances require you to leave early, please inform me before the start of class and sit near the back of the classroom so you can leave with as little disruption as possible.

Use of laptops and cell phones (allowed)

You are allowed to use laptops and tablets during lecture, but they should only be used for coursework related activities and not for email, social media, or other activities not directly related to the course. Cell phones must be turned off or silenced during class. No photography or recording of any kind is allowed, without express consent from me.

Students may not audio or video record class lectures or other classroom or laboratory activities without the instructor's permission. The sharing of course materials on an individual level for educational purposes

(e.g., working with groups or with a tutor) is permitted, provided that it has not been prohibited by the instructor. Students may not publish, distribute, or sell—electronically or otherwise—any course materials that the instructor has developed in any course of instruction in the University (e.g., presentation slides, lecture aids, video or audio recordings of lectures, and exams) without the explicit permission of the instructor. The sharing or distribution of course materials for purposes of giving or gaining unfair advantage in a course is prohibited. Students must further respect the requirements of copyright protection for materials that are made available for instructional purposes. Students may not provide or share Zoom Links or classroom materials to students not enrolled in the appropriate courses.

Course Outline and Schedule

Schedule (subject to change)

Session		Topics	In-class labs	Additional labs	Assignment
1	8/26	Introduction;	Basics of text and social media		
2	8/31	Lexical analysis	analysis		
3	9/2	Semantic analysis	Text info extraction and representation Sentiment analysis	Python basics Overfitting	
4	9/7				HW1
5	9/9	Topic modeling	Topic modeling	Sentiment analysis	TP
6	9/14				HW2
--	9/16	Career Fair Break - no classes			
7	9/21	Contextual information representation	Context information	General Inquirer Tagging	
8	9/23		Word2vec		HW3
9	9/28	ML in text analytics	ML in text analytics	Similarity	HW2 due
10	9/30				
11	10/5	Deep neural networks and AI	Neural networks	Movie review DL 1	HW3 due
12	10/7				HW1 due
13	10/12	Advanced topics;	No labs	Movie review DL 2	TP due
14	10/14	Team presentations			

*HW for homework, TP for team project