Course Syllabus

Network Data Science

ENPM809G
Spring 2023

Course Overview

This course will introduce methods for analyzing and understanding the structure and function of networks, including social networks, web graphs, and sensor networks. The course will introduce students to the math and science of network analysis. Through real world examples, including analysis of their own networks, students will develop skills for describing and understanding the structure, patterns, and functionality of networks. Students will read classic and cutting-edge articles and books about these topics and discuss their applicability to various network types. The class will culminate with a capstone project in which students will apply the analysis methods they have learned to understanding a particular question about a network they choose.

Learning Outcomes

After successfully completing this course, you will be able to:

- Understand to the basic concepts of network analysis (Evaluation: problem sets)
- Collaborate with peers to apply these methods to a variety of networks (Evaluation: projects)
- Understand the link between qualitative and quantitative methods of network analysis (Evaluation: short analysis papers)

Dr. W. Lewis Collier

lcollier@umd.edu

Class Meets

Mondays

7:00pm – 9:40pm

Location: JMP TBD

Office Hours

TBD

and by appointment

Teaching Assistants

TBD

Prerequisites

Basic statistics and programming experience.

Course Communication

All communications outside of class time will be via email provided above. Please ensure that an email that is accessible by you from mobile devices, or on which you are receiving constant updates is provided in case alerts are needed to be sent.

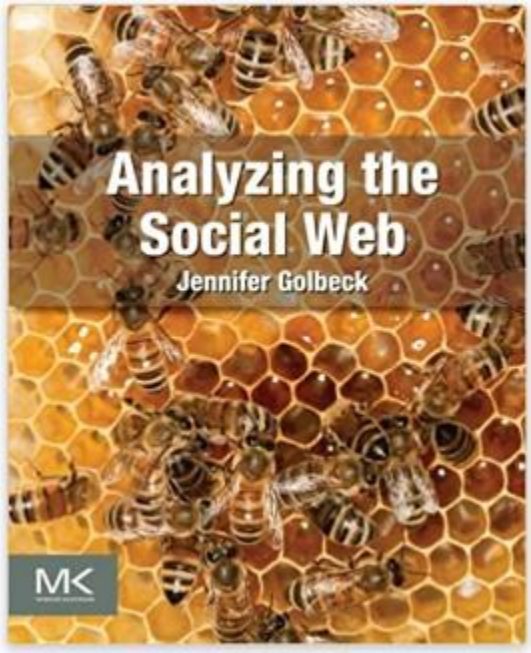
Course website
elms.umd.edu

Resources

Course website elms.umd.edu

Book

Analyzing the Social Web , by Jennifer Golbeck **ISBN-13:** 978-0124055315



Please also get the Errata here

Readings

Additional materials (slides) will be provided in web site.

Hardware/Software

This is TBD, but some version of Python is to be used.

Supplemental Resources

Readings:

Provided on course website.

Hardware/Software

Gephi, Anaconda.

Campus Policies

It is our shared responsibility to know and abide by the University of Maryland's policies that relate to all courses. Please visit https://academiccatalog.umd.edu/graduate/policies/academic-record/ for the Office of Graduate Studies' list of campus-wide policies.

Code of Academic Integrity

The University of Maryland, College Park has a nationally recognized Code of Academic Integrity, administered by the Student Honor Council. This Code sets standards for academic integrity at Maryland for all undergraduate and graduate students. As a student you are responsible for upholding these standards for this course. It is very important for you to be aware of the consequences of cheating, fabrication, facilitation, and plagiarism. For more information on the Code of Academic Integrity of the Student Honor Council, please visit https://tltc.umd.edu/integrity.

Activities, Learning Assessments, and Expectations for Students

The phenomenal growth of biometrics recognition and identification systems (both within organizations and publicly) combined with advances in technology (in cloud computing and machine learning) have allowed organizations to leverage biometrics processing to enhance their recognition and identification processes to achieve their missions. In addition, the availability of low-cost computing resources and sensors have accelerated this process.

This course focuses on traditional methods to extract usable recognition and identification information from biometrics sensors and communicate those technologies to stakeholders so they can be incorporated in operational work processes to achieve organizational goals. Students will gain a comprehension of and use relevant tools, technologies and approaches for biometrics recognition and identification processing. Architectural approaches for implementing biometrics-based recognition and identification will be covered.

Students will use analytics tools and scripting languages for the implementation of projects. Some example languages/tools introduced are using the Python scripting language. Prior knowledge of scripting language such as Python or R will be helpful – but not required. If not already capable, students will learn these basic scripting skills in order to complete assignments during the class.

All assignments include a written description that clearly, and concisely presents the results of the assignment. The goal of the reports is to build confidence in reporting of data analysis efforts. All assignments are to be presented in a “mini-report” format that follows the standard academic paper structure of abstract, introduction, methodology, results, and discussion. The final project is to be presented as a more in-depth report that follows the same structure as assignments, with an accompanying slide presentation.

While strongly recommended, but ultimately not required, the use of LaTeX for the reports if preferred. Instructions for utilizing LaTeX are readily available and instructions will be posted, including a sample class (style format) file, on the course website. TeXnic Center (a full featured IDE) with the MiKTeX distribution for MS-Windows is preferred, so that help can be found from multiple sources.

Performance in this class, and your associated grade, will be measured via three (3) assessment areas:

1. 50% Homework assignments (problem sets, papers, short analysis projects)
2. 25% Class Participation (includes in-class exercises)
3. 25% Final Project

You are expected to complete assignments on the target dates. The instructor may grant limited extensions of time for specific assignments for health or personal emergencies. That extension normally will be granted only if arranged with the instructor in advance. Otherwise, penalties will be assigned for late completion up to 15% per day.

To maximize your learning/success, class participation is important and expected. While there is a specific grading component for attendance and participation, these facets also contribute greatly to necessary

understanding for the other graded items. Thus, participation factors in much more than just in the participation grade component.

For each class, I have prepared items to be completed before you arrive at the class, what will be covered during the class, and what is expected to be completed after the class. The overall schedule of topics for the class are:

Date	Part 1 – Understanding Biometrics Terminology	
30 Jan 2023	**Week 1**	Network Structure I (Ch 1,2) [Homework 1 – 10 pts]
06 Feb 2023	**Week 2**	Network Structure II (Ch 3) [Homework 2 – 20 pts]
13 Feb 2023	**Week 3**	Visualization (Ch 4) [Homework 3 – 10 pts]
20 Feb 2023[1]	**Week 4**	Jupiter and NetworkX [Homework 4 – 10 pts]
27 Feb 2023	**Week 5**	Random Graphs in NetworkX [Homework 5 – 20 pts]
06 Mar 2023	**Week 6**	Random Networks, Power Laws, and Mixing [Homework 6 – 20 pts]
13 Mar 2023	**Week 7**	Network Building (Ch 8,9)
20 Mar 2023	**Week 8**	SPRING BREAK
27 Mar 2023	**Week 9**	Connecting Structure to Content [Homework 7 – 10 pts]
03 Apr 2023	**Week 10**	Network Propagation (Ch 10) [Final Project Assigned]
10 Apr 2023	**Week 11**	Network Propagation and Communities
17 Apr 2023	**Week 12**	Small Worlds, Milgram, and Data Science Ethics
24 Apr 2023	**Week 13**	Identifying Communities and Bipartite Graphs
01 May 2023	**Week 14**	Dynamic Networks
08 May 2023	**Week 15**	Final Presentations (12 May) Reading Day
13-19 May 2023	**Week 16**	Final Exam Week !!! Any missing assignments due 13 May 2023 !!!

Course-Specific Policies

I expect you to make the responsible and respectful decision to refrain from using your cellphone during class time. If you have critical communication to attend to, please excuse yourself and return when you are ready. For more information about the science behind the policy watch: http://youtu.be/WwPaw3Fx5Hk

Likewise, as I do not expect students to be buried in their computers, this will not be a PowerPoint presentation-centric course. I will use slides to highlight topics, and data sets, but I will not be reading them to you. I expect students to have read the material before the class so that in depth discussions can be had during class (see note above about grading for class participation). As stated above, the overarching goal of this class is to provide an introductory understanding of network data science algorithms. Hands on practice will lead to a better understanding when real data is processed, but discussion beforehand will afford the base understanding that enables the deeper comprehension when data is processed and analyzed.

The course is structured so that group efforts are used in the beginning of the course, followed by individual efforts. The intent is that the group members work together to get up to speed on the tools and methodologies. In addition, group study is encouraged for the mid-term, which focuses on the presented algorithms. The second half of the course focuses on ensuring that all students can utilize the base decision algorithm information and utilize the baseline understanding of the first half of the course to individually evaluate the data analysis tools for various applications. In this way, students can utilize the course methods for data that is pertinent to each

[1] This is Presidents Day but it appears that classes DO meet on this day.

individual's needs and interests. Group study for the individual projects and final exam are also encouraged, but ***ALL*** work products marked as individual efforts are expected to be created individually.

This is a graduate course so there is an expectation that all group members will contribute equally to the group work products. Some will provide more help in coding aspects, some with the underlying math, and some with the writing. The goal is for the groups to help all members become stronger in their less-effective areas so all students can thrive in data analysis for the second part of the course, and beyond.

For this course, some of your assignments will be collected via Turnitin on our course ELMS page. I have chosen to use this tool because it can help you improve your scholarly writing and help me verify the integrity of student work. For information about Turnitin, how it works, and the feedback reports you may have access to, visit Turnitin Originality Checker for Students

Accessibility and Reasonable Accommodations

The University of Maryland is committed to creating and maintaining a welcoming and inclusive educational, working, and living environment for people of all abilities. The University of Maryland is also committed to the principle that no qualified individual with a disability shall, on the basis of disability, be excluded from participation in or be denied the benefits of the services, programs, or activities of the University, or be subjected to discrimination. The University of Maryland provides reasonable accommodations to qualified individuals. Reasonable accommodations shall be made in a timely manner and on an individualized and flexible basis.

Discrimination against individuals on the grounds of disability is prohibited. The University also strictly prohibits retaliation against persons arising in connection with the assertion of rights under this Policy.

Accessibility & Disability Service (ADS) facilitates reasonable accommodations to qualified individuals. For assistance in obtaining an accommodation, contact Accessibility and Disability Service at 301.314.7682, or adsfrontdesk@umd.edu. More information is available from the Counseling Center.

Get Some Help!

You are expected to take personal responsibility for your own learning. This includes acknowledging when your performance does not match your goals and doing something about it. Everyone can benefit from some expert guidance on time management, note taking, and exam preparation, so I encourage you to consider visiting http://ter.ps/learn (there are specific resources for graduate students under handouts, but please explore to find what you need). Sharpen your communication skills (and improve your grade) by visiting https://gradschool.umd.edu/graduate-school-writing-center and schedule an appointment with the campus Graduate Writing Center. Finally, if you just need someone to talk to, visit http://www.counseling.umd.edu.

Everything is free because you have already paid for it, and **everyone needs help**… all you have to do is ask for it.

Names/Pronouns and Self Identifications

The University of Maryland recognizes the importance of a diverse student body, and we are committed to fostering equitable classroom environments. I invite you, if you wish, to tell us how you want to be referred to both in terms of your name and your pronouns (he/him, she/her, they/them, etc.). The pronouns someone indicates are not necessarily indicative of their gender identity. Visit trans.umd.edu to learn more.

Additionally, how you identify in terms of your gender, race, class, sexuality, religion, and dis/ability, among all aspects of your identity, is your choice whether to disclose (e.g., should it come up in classroom conversation about our experiences and perspectives) and should be self-identified, not presumed or imposed. I will do my best to address and refer to all students accordingly, and I ask you to do the same for all of your fellow Terps.

Detailed Course Schedule

The course will follow the schedule presented above. Each week consists of things to do before class, topics to be covered during the class, and things to do after the class.

Note: This is a tentative schedule, and subject to change as necessary – monitor the course ELMS page for current deadlines. In the unlikely event of a prolonged university closing, or an extended absence from the university, adjustments to the course schedule, deadlines, and assignments will be made based on the duration of the closing and the specific dates missed.

Grades

Grades are not given, but rather, they are earned. Your grade is determined by your performance on the learning assessments in the course and is assigned individually (not curved). If earning a particular grade is important to you, please speak with me at the beginning of the semester so that I can offer some helpful suggestions for achieving your goal. Grades for each will be weighted in the following manner:

As the assignments and reports are not expressly quantitative assessments, grades will be assigned as A+, A, A-, B+, and so forth) rather than strict numerical grades.

All assessment scores will be posted on the course ELMS page. If you would like to review any of your grades (including the exams), or have questions about how something was scored, please email me to schedule a time for us to meet before or after class, or virtually during office hours.

You are expected to complete assignments on the target dates. The instructor may grant limited extensions of time for specific assignments for health or personal emergencies. That extension normally will be granted only if arranged with the instructor in advance. Otherwise, penalties will be assigned for late completion of a fractional letter grade (A to A-, A- to B+, B+ to B, and so on) per day.

I am happy to discuss any of your grades with you, and if I have made a mistake, I will review it in very short order and correct it as applicable. Any formal grade disputes must be submitted in writing and within one week of receiving the grade.

Final letter grades are assigned based on the weighted percentage of total assessment points earned. The standard 4.0 scale will be applied to the letter grades for each assignment, exam, or final presentation, with a significant bonus for marks of A+. Final grades will be tabulated from the weighted averages, per the table below. Thus, an average of 3.700, or higher, would equate to a letter grade of A and an average of 3.400 to 3.699 would equate to a letter grade of A-. To be fair to everyone I have to establish clear standards and apply them consistently, so please understand that being close to a cutoff is not the same as making the cut ($3.699 \neq 3.700$). It would be unethical to make exceptions for some and not others.

See the following table for an exact breakdown of the grading per assignment and for the computation of the overall course grade.

Assignment Grade	Grade Value	Course Grade
A+	**4.3**	**N/A**
A	**4.0**	**>= 3.700**
A-	**3.666**	**>= 3.400**
B+	**3.333**	**>= 3.200**
B	**3.0**	**>= 2.700**
B-	**2.666**	**>= 2.400**
C+	**2333**	**>= 2.200**
C	**2.0**	**>= 1.700**
C-	**1.666**	**>= 1.400**
D+	**1.333**	**>= 1.200**
D	**1.0**	**>= 0.700**
D-	**0.666**	**>= 0.400**
F	**0.0**	**< 0.400**