**Computational Research Methods**

**Typically Taught in Spring**



**Natalie Lambert, Ph.D.**

**Brian Lamb School of Communication**

**Email: njlambert@purdue.edu**

**Office: BRNG 2160**

**Computational Research Methods**

**Syllabus**

**Course Description**

Computational research methods utilize new advances in computing such as algorithms, models, simulations, and systems in order to understand complex social, biological, technological, and endless other patterns and behaviors. These research methods have the ability to:

* Investigate the world on many levels, from the individual to group and system levels
* Automatically extract and categorize information of interests from large datasets
* Model the emergence of new behaviors from individual-level interactions
* Identify meaningful structural patterns within webs of relationships
* Recognize important thresholds within a complex system
* Explain the emergence of a system or network
* Find patterns within datasets that are too large to be analyzed using other methods

The purpose of this course is to provide students with an introduction to four cutting-edge computational research methods: network analysis, text mining, agent-based modeling, and machine learning. The course will be structured as a series of hands-on workshops designed to introduce students to the basic mechanics of each method. Students will complete hands-on methods tutorials in class and as part of their homework. Together we will discuss assigned readings that showcase applications of these methods. This seminar will culminate in a research proposal in which students design a research study implementing one of these methods in a context related to their personal areas of research. Students who have taken a computational methods course with me in the past can collect and analyze data for this project. No programming skills are required, and this course is structured to be applicable to students of any disciplinary background who have an interest in data exploration. By the end of the semester you will have learned: 1) the mechanics and underlying assumptions of each method, 2) how data is collected and formatted for each method, 3) how to perform basic analyses using each method, and 4) how to apply each method within a social scientific or other context of personal interest.

**Required Readings**

* Alvarez, M. R. (2016) *Computational Social Science: Discovery and Prediction.* New York, NY: Cambridge University Press.
* Articles and chapters that I make available to you on Blackboard.

**Software Tools Introduced in this Course**

**AutoMap**: Carley, K. (2016). AutoMap (version 3.0.10.41) [Computer software]. Pittsburg, PA: CASOS, Carnegie Mellon University. Available from http://www.casos.cs.cmu.edu/projects/automap/index.php

**NetLogo**: Wilensky, U. (1999). NetLogo. <http://ccl.northwestern.edu/netlogo/>. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

**NodeXL**: Smith, M., Milic-Frayling, N., Shneiderman, B., Capone, T., Mendes Rodrigues, E., Leskovec, J., Dunne, C. (2012) Network Overview, Discovery and Exploration Add-In for Microsoft Excel. http://nodexl.codeplex.com

**Weka:** Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. ACM SIGKDD explorations newsletter, 11(1), 10-18.

**Grade Breakdown**

Participation: 10%

Take-home Tutorials: 20%

Method Reports: 30%

Research Proposal: 40%

**Participation**: Students should attend and participate in every class meeting. Students are responsible for completing weekly readings (posted on Blackboard) prior to class. Participation in class discussions is essential for learning how to apply and debate computational research methods. You should participate in class discussions frequently enough that I and your classmates know your name and your perspective about the readings and class concepts. I will notify students by February 28th if their level of in-class participation is insufficient for full credit.

**Take-home Tutorials**: You will be assigned four take-home tutorials over the semester. They will be online videos, book chapters, or written instructions that explain step-by-step how to conduct an analysis using each of the methods. These tutorials are designed to give you experience using the software as well as an understanding what decisions must be made during analysis. Students will provide screenshots of their work and answer accompanying questions. No tutorial is perfect, and in fact, available tutorials for these methods are often not geared towards social scientists. However, part of learning these methods is to work to translate concepts embedded in these methods from other fields into our own. Please therefore leave yourself plenty of time to complete each tutorial. You may work with one partner on these tutorials, but both partners must write their names at the top of the assignment before turning it in. Turn in assignments to me via Blackboard by the due dates. Assignments turned in late will be penalized 33% off the graded score for every 24 hours they are late.

**Method Reports**: For each method you will write a report wherein you describe in your own words the mechanics of each method, its underlying assumptions, and the fundamentals of how each method is performed. Students will also review a small number of studies that apply the method (other than those that were introduced through assigned course readings), and answer a few questions about the method. The format for these review papers will be posted in Blackboard. The method reports should be about 6 pages double-spaced in length including citations.

**Research Proposal**: You will write a research proposal for this class wherein you design a research study utilizing one of the methods introduced in this class. The paper need not be more than 15 pages, but must clearly explain the study design, how the method will be used, how data will be collected and formatted, and how data analysis will take place. Students must also include a literature review that combines research articles from their personal area of interest with research that utilizes the computational method they have chosen for their design. If you have taken a computational research methods course with me in the past, you may use this paper to further an ongoing research project. If you are collecting human subject data or wish to publish the results of this paper, you must secure IRB approval prior to data collection. **You must receive written approval from me prior to data collection.** Further instructions for the research proposal will be provided on Blackboard.

**Class Policies**

You can expect me to give you feedback on any assignment, to respond to your emails within 24 hours except over the weekend and during holidays, and I will provide you with further learning resources when you are interested.

I expect all students to adhere to Purdue University’s academic code of conduct. As Purdue states:

*Purdue University values intellectual integrity and the highest standards of academic conduct. To be prepared to meet societal needs as leaders and role models, students must be educated in an ethical learning environment that promotes a high standard of honor in scholastic work. Academic dishonesty undermines institutional integrity and threatens the academic fabric of Purdue University. Dishonesty is not an acceptable avenue to success. It diminishes the quality of a Purdue education, which is valued because of Purdue's high academic standards.*

Purdue’s Academic Integrity Policy can be found here: https://www.purdue.edu/odos/academic-integrity/

The Purdue Student Conduct Statement can be found here:

http://www.purdue.edu/studentregulations/student\_conduct/

Additionally, I expect students to be respectful of each other’s opinions during class discussions. This class is comprised of students from different departments and disciplines, which creates an exciting opportunity for us to learn about the different contexts in which computational research methods can be applied. I expect students to listen to each other during discussion without interjection or interruption, and to make time during conversation for quieter students to contribute to the conversation.

**Emergency Preparedness**

* If the fire alarm sounds, we will immediately exit the building via a stairwell and gather together in the courtyard between BRNG and UNIV for a head count. We will not use the elevators.
* During a tornado or other storm event, we will shelter in the basement of BRNG.
* During a shooter event or other dangerous event, we will shelter in place in our classroom unless otherwise directed by campus authorities. You may wish to sign up for the campus emergent text alert system.

**Class Schedule**

The class schedule contains each class meeting topic and the due dates for homework and research proposal deliverables. The required reading assignments are posted on Blackboard. I reserve the right to make changes to the schedule but will always notify you in advance of changes. No change will reduce the time allotted to you to complete an assignment.

I’m looking forward to a great semester!

|  |  |  |
| --- | --- | --- |
| **Meeting** | **Topic** | **Deliverables** |
| **1** | Introduction to Each Other and the Class |  |
|  |   |  |
| **2** | Overview of Computational Research Methods |  |
|  |   |  |
| **3** | Computational Datasets and Research Questions |  |
|  |   |  |
| **4** | **Readings Discussion** |  |
|  |   |  |
| **5** | Overview of Network Analysis |  |
|  |   |  |
| **6** | Data Collection and Formatting for Network Analysis |  |
|  |   |  |
| **7** | Network Analysis in NodeXL: Part I | **Network Analysis Tutorial Due** |
|  |   |  |
| **8** | Network Analysis in NodeXL: Part II |  |
|  |   |  |
| **9** | **Readings Discussion and Lab Day** |  |
|  |   |  |
| **10** | **Visit from IRB**: Collecting Publicly Available Online Data | **Network Method Report Due** |
|  |   |  |
| **11** | Overview of Text Mining |  |
|  |   |  |
| **12** | Text Mining in AutoMap: Part I | **Text Mining Tutorial Due** |
|  |   |  |
| **13** | Text Mining in AutoMap: Part II |  |
|  |   |  |
| **14** | Social Media Research Design |  |
|  |   |  |
| **15** | Ethical Collection of Large-scale Datasets |  |
|  |   |  |
| **16** | **Readings Discussion and Lab Day** | **Text Mining Method Report Due** |
|  |   |  |
| **17** | Crash Course in Programming |  |
|  |   |  |
| **18** | Overview of Agent-based Modeling |  |
|  |   |  |
|  | **SPRING VACATION** |  |
|  |   |  |
|  | **SPRING VACATION** |  |
|  |   |  |
| **19** | Agent-based Modeling in NetLogo: Part I | **Agent-based Modeling Tutorial Due** |
|  |   |  |
| **20** | Agent-based Modeling in NetLogo: Part II |  |
|  |   |  |
| **21** | **Readings Discussion and Lab Day** | **Agent-based Modeling Method Report Due** |
|  |   |  |
| **22** | Overview of Machine Learning |  |
|  |   |  |
| **23** | Data Collection and formatting for Machine Learning | **Machine Learning Tutorial Due** |
|  |   |  |
| **24** | Machine Learning in Weka: Part I |  |
|  |   |  |
| **25** | Machine Learning in Weka: Part II | **RESEARCH PROPOSAL DRAFT DUE (OPTIONAL)** |
|  |   |  |
| **26** | **Readings Discussion and Lab Day** |  |
|  |   |  |
| **27** | The Future of Computational Research Methods |  |
|  |   |  |
| **28** | **LAB HELP DAY** | **Machine Learning Method Report Due** |
|  |   |  |
| **29** | **RESEARCH PROPOSAL PRESENTATIONS** |  |
|  |   |  |
| **30** | **RESEARCH PROPOSAL PRESENTATIONS** |  |
|  |   |   |
|  |   | **FINAL RESEARCH PROPOSAL DUE** |