Networks and Markets
CAS CS 591 – Fall 2020
The Course Project

1 Project Guidelines

What am I looking for in a strong project?

- First, I am hoping that there will be a component of your project that connects to an interesting, elegant, or under-explored aspect of networks or markets that you are able to drill down upon in your work. Since there is so much work out there that investigates social networks and network connectivity, a cross-cutting study that connects to firms and marketplaces is appropriate, and will make your study more original.

  Recall that I have intentionally left the project open-ended: you may focus primarily on a study that involves data analytics, or you may develop an app that has novel functionality relating to networks or markets, or you can even try to craft a business plan (although this last one will be challenging to do well, and you will need to obtain my explicit approval to proceed).

- One direction that many students have succeeded with in the past is to start from a compelling hypothesis and investigate the question through datasets. Often, the thrust of a good project can be summarized in a single sentence. Here are three such questions that student teams were able to effectively investigate in my E-Commerce class last year.

  1. Farecast and other similar firms provide recommendations on whether to purchase a specific flight now, or wait (also with predicted confidence values). We propose to investigate through a data-driven analysis: how accurate are these recommendations in practice?
  2. Kickstarter campaigns are all the rage, but what causes some campaigns to thrive while others fail? We propose to investigate: what are the attributes of campaigns that are most correlated with success, and if time allows, build a predictive model for the success of hypothetical campaigns.
  3. Using the Yelp academic datasets, we propose to study the extent to which social networks influence restaurant ratings. Specifically, if we know only the time-series of which individuals reviewed each restaurant, how well can we predict the ratings those restaurants received. How much value does this provide over baseline methods in which less information is known?

- Second, I am expecting that there will be a rigorous quantitative evaluation of your approach, regardless of your methodology. This could take the form of an in-depth analysis, or an experimental evaluation of a new method, or quantitative modeling of a new mechanism. What we are shooting for is a rigorously designed evaluation section that both cleanly and elegantly builds the case for the aspect or approach you elected to study. I am not expecting

you to conduct your research in isolation – you should plan on frequent discussions with me, as well as other faculty advisors if you wish, to make sure your project is on track.

2 Deliverables

- Project proposal: by October 27, I ask you to submit a two-page summary (minimum) of your proposed project activities. Begin with a project title, the team members and their roles, and a one to two sentence description of the project that gets to the heart of what you intend to study. Then elaborate the project with an overview of the area your team intends to study and the direction you intend to pursue, along with a preliminary description of your methodology or approach, including any data you intend to use and how you will collect it (if not publicly available). You must include a working bibliography of related work that either connects to your proposed line of work or which you intend to draw from when conducting your study.
- For now, we will reserve the last two classes of the semester, December 8th and December 10th, for presentations of the class projects. I will expect you to produce (pre-recorded videos will be acceptable) a ten-minute presentation highlighting your results. We will talk about guidelines for good presentations as the end of the semester approaches.
- During finals week, a conference-style writeup of your results will be due. The length of the writeup is not as important as the quality of the writeup (but plan on roughly 8-12 pages in single-columned format). We will talk about guidelines for a good paper as the end of the semester approaches. I will also give you a LaTeX template which I will suggest (but not require) you to use for your paper later in the course.

3 Project Ideas

Before I enumerate a few more project pointers, let me first start with a word of caution. As with any piece of research, if you are not intensely interested in getting to the bottom of a problem, then you are not likely to make much progress. Therefore, **do not** choose a topic just because it has been suggested to you — pick a topic you care about deeply. Also, use the resources that are available to you, i.e. talk with me, other students, and even other faculty members about possible project directions that you are considering, and continue to consult with us throughout the semester. Finally, picking a good project can be very challenging, so get started thinking about topics, researching related work, and collecting data early.
Some papers and datasets are listed below, but it is likely that the best projects will follow a line of the students' own initiative, and not one of those mentioned below.

The Yelp Academic Dataset Yelp has provided a ready-made project in its Academic Dataset: https://www.yelp.com/dataset/. But be sure to take the Yelp data in your own new direction – this may not be so easy to do.

Sometimes just seeing what data is out there might spark a new direction, but again, try not to reinvent the wheel. The Stanford Large Network Dataset Collection is one collection of datasets

that might spark your curiosity (but some are getting older). A more recent dataset that I've used in my research is the 2017 Yellow Taxi Trip Data from NYC OpenData.

Multi-Level Marketing: One interesting direction is multi-level marketing, where people at the top of the pyramid get compensated for recruiting others into (typically) a sales effort. A similar model was adopted by the winning MIT-based team that won DARPA's Network (Red Balloon) Challenge. This paper is a really interesting read: Moshe Babaioff, Shahar Dobzinski, Sigal Oren and Aviv Zohar, "On Bitcoin and Red Balloons".

Online Diffusion: Many recent directions in electronic commerce related to social networks, and word-of-mouth marketing. One line is to reason about how information diffuses and propagates through these networks. There is a great deal of work here. Some papers include: Sharad Goel, Duncan Watts, Daniel Goldstein, "The Structure of Online Diffusion Networks"; Flavio Chierichetti, Jon Kleinberg and Alessandro Panconesi, "How to Schedule a Cascade in an Arbitrary Graph"; and Eytan Bakshy and Dean Eckles, "Effects of Social Cues and Tie Strength in Social Advertising: Evidence from Field Experiments".

Voter Models: A different social network direction relates to how individuals' choices influence other individuals. A widely studied model is the so-called voter model, dating back to the early '90s. Investigating this theoretical model in the context of consumer choice could make for a nice project bridging between theory and practice. One starting point is: Durrett et al. "Graph fission in an evolving voter model". Proc. Nat'l. Acad. Sci. 109 (2012).

Online Advertising: Later in the class, we will be coming back to this topic in more depth, but there is a very rich ecosystem here. Just google "online advertising ecosystem kawaja" to see what I'm talking about. Many aspects of this field are open to investigation, for example: Mohammad Mahdian, Arpita Ghosh, Preston McAfee and Sergei Vassilvitskii, "To match or not to match: Economics of cookie matching in online advertising".

Many other examples, firms, and problem domains are in scope. This is just a small, but representative sample to give you some ideas.