Project 3
Social Network Simulation

Description To better understand the dynamics of how information bubbles form and how political polarization can occur in social networks, this assignment provides a small-scale simulation of a social network to investigate these phenomena. The simulation will allow you to see how choices with regard to trying to maximize revenue generated by user activities in the network may lead to a polarization of users. Through this investigation we want you formulate a strategy for how a social network might make choices regarding how it recommends content to users to achieve what you believe are the right outcomes for the social network and for society.

The Data and Code

Information on the data and the code can be found in the file titled `Project3 - data explanation.pdf`. You should be sure to read all of that information carefully before starting with the task below.

The task

Your task in this assignment is to run a social network simulation to understand the dynamics of the network under different conditions (e.g., how recommendations are optimized in the system, different numbers of users in the network, etc.). Note that it is possible to complete this assignment without having to modify any of the actual code for the simulation as the simulation allows for setting various critical parameters through the text-based user interface. Still, we provide the code for the simulation to allow you to make changes if they help you, for example, to gather more data for your write-ups, test ideas for suggestions on how you might address some of the issues encountered under different conditions in the network, etc. Presently, we provide an explanation of what is happening in the social network simulation and what results are presented. Then, we detail the specific work required for this assignment.

You are employed at `SocialNewsBook.net`, an up-and-coming online social network that is focused on engaging users with news. The company is interested in how it might best engage its users in both the short and long-term. It has built a simulation of its social network (i.e., the simulation in this assignment) to allow you to determine how the users in the social network would behave in a variety of situations, controlled by various parameters that can be set through the user interface of the simulation. The code for the simulation is also provided to you just in case you would like to look at the internals of the simulation or even modify the code to help you gather data, but you are under no obligation to do so. In other words, it's fine for this assignment if you just use the simulator, modifying parameters through the text-based interface without actually changing any of the underlying code.

The executive staff of the company has tasked you with helping them determine how they should be recommending news articles to users in the network (both the initial article that is presented to each user each day as well as how articles are posted to users' news feeds). More specifically, the executives want you to answer the questions below. You should write up the answer to all these questions in a single PDF file titled "Project3" and submit it in Gradescope.

Questions to answer:

Note: Regarding word counts in the questions below, statistics you cut/paste from the simulation runs or any tables/charts of numerical data you create and include, are not included in the word count. You should feel free to include such data in your write-up if it is useful for making a point, without worrying about it impacting the word count.

1. Run the social network simulation with various numbers of left/right leaning users for 500 days with the "Probability to explore for one user" parameter set at 0.1 and the "Probability of diversity among users" parameter set at 0.1. In a paragraph (about 200 words, though this is neither a

strict minimum nor maximum), explain what is happening in the network as it evolves during the 500 days of the simulation (with respect to the network structure, other statistics reported for the simulation, and any other interesting observations you made).

2. Run the social network simulation with various numbers of left/right leaning users for 500 days with "Probability to explore for one user" parameter set at 0.9 and "Probability of diversity among users" parameter set at 0.9. In a paragraph (again, approximately 200 words), explain how and why the results in this case (both with respect to the structure of the network, other statistics reported in the simulation, and any other interesting observations you made) are different than the results from Question 1.
3. Consider the change in revenue generated by the social network simulation in Question 1 versus that in Question 2 (for 20 each of left/right leaning users in both cases). In approximately 300 words, explain what drives this revenue difference, especially with respect to what is happening in the social network simulation in Question 1 versus Question 2. Include an explanation of the dynamics in the social network. To answer this question, you are encouraged to run the simulation with a variety of parameter values to give you a better sense of the dynamics in the network and the impact on revenue generation. That will also give you more data to help you better explain/justify your answer.
4. Write an approximately 600 word memo to the executive team of `SocialNewsBook.net` justifying how you believe they should make decisions with respect to how news articles should be recommended to users in their social network. You can assume the executive team has a solid technical background—they all know how to code well —and have done all the readings and attended all the classes in CSC-395, so they are familiar with both the technical and nontechnical issues discussed in class.

   The executive team is interested in both the short and long-term impacts of your recommendations for the company, its user base, and the relationship with news content providers (i.e., news sources) that it recommends articles from. Moreover, the executives are interested in public policy considerations of potential actions that they (or their competitors) might take. For example, if/how government regulators might react to potential actions taken by social networking companies such as `SocialNewsBook.net`.

   Utilize data you obtain from running the simulation as well as concepts from the class lectures and readings (and external sources, if desired) to justify your position, including clearly specifying criteria for how you believe recommendations should be made in the social network. Your memo should specify how the company should make exploration/exploitation decisions when making recommendations (including, but not limited to, user polarization impacts), how the amount of revenue generated should impact these choices, and what other criteria, if any, beyond revenue generation should be considered and why. As mentioned previously, you may modify the code for the simulation if it helps you capture additional data for your memo, but you are under no obligation to do so. As a side note, we point out that the profit margin for many large successful online platforms is approximately 30%. Profit margins are much lower than that while the platforms are earlier in their lifecycle. Keep that in mind while discussing revenue implications in your memo.

   If it matters in your write-up, with respect to the simulation provided here, you can think of the revenue generated when the parameters are set for maximal "exploration"(i.e., when the "Probability to explore for one user" and "Probability of diversity among users" parameters are both set close to 1.0)) as indicative of the base cost of running the social network (i.e., the network is generating little or no profit in that case). Any revenue generated above that amount (using different parameter settings) could then be considered profit for the company.

Learning Outcomes Completion of this assignment will contribute to your ability to fulfill the following learning outcomes:
5. Identify the impact of different technologies on various identities.
6. Identify strategies for creating more equitable and inclusive technical environments and software for diverse identities.
8. Engage in current ongoing discussions of algorithms, ethics, and society.
9. Understand advanced algorithms in the context of their application(s).

Criteria for Success Assignments are graded on a Satisfactory/Not scale, which I hope encourages you to marinate in the learning rather than being overly concerned by grades. I will ask for further revisions of any assignment that receives a grade of Not Satisfactory (*NS*). Assignments that meet the following criteria will be graded Satisfactory (*S*):

- All questions are answered with the approximate word suggestions.
- All answers to questions indicate thoughtful reflection of the simulations.
- Data is used to support answers to questions.
- Resources are informally or formally cited (you do not need to cite the code for this assignment, but any course readings or other documents you use should be cited).

Acknowledgments This assignment was written and designed by Rob Reich, Jeremy Weinstein, and Mehran Sahami, all at the Stanford University for their CS182 course. Student Michael Dworsky also categorized the recidivism data and developed an early prototype for the assignment. They have given me permission to use their materials.
Any changes made have been minor with the aim of providing context, or adapting to the specifics of CSC-395 at Grinnell College.