

605.433 – Social Media Analytics

Module 13 Homework

For this assignment, you will obtain social media data and conduct an exponential random graph model (ergm). You will submit a type written paper in word that describes your data and ergm analysis, as well as the data you collected. You will also complete the RSiena example included in the lecture.

Specifically, you will download/collect/obtain an interesting social media network data set. This can be the same data that you downloaded or collected for a previous assignment. Make sure that the data has at least two attributes defined for the nodes, and two edges defined between nodes (hint: this can be mention and retweet or some permutation using relational algebra).

In your write-up, provide

- I. Overview A brief summary of the data. This is effectively an abstract for the dataset.
- II. Data Collection A narrative description of how data were collected, including a high-level description of questions asked (if from online survey/blog) or methods employed to collect data.
- III. Data Files and Formats A description of the structure of the data accompanying this article. The data may be in any format useful to social network programs. You just need to describe that format. For example, are the network data stored as an adjacency matrix, edgelist, or JSON. If a JSON, describe how the JSON is structured.
- IV. Data Details A required table detailing relevant information about the dataset (if applicable), including
 1. response rates/sampling rate (i.e. garden hose/fire hose; targeted search)
 2. non-respondent bias or the impact of the sampling approach
 3. any theoretical grounding for questions or methods employed
 4. any existing publications employing these data
 5. a short description of the context
 6. nature of the respondents
 7. whether the data are longitudinal and, if so, details about collection intervals
 8. temporality of the data (e.g., the extent to which they are specific to the time at which they are collected)
 9. analytic utility of the dataset – aspects others may find interesting in this dataset for teaching or research purposes.
 10. known issues that threaten the validity of the data or anything else other social network analysts using these data for teaching or research should be aware of.

Create an exponential random graph model (ergm) for your data. In your write up, provide the following:

- V. ERGM model summaries of:
 1. Initial restricted model including edges, reciprocity (if undirected), and a triadic term.
 2. Final restricted model with significant terms.
 3. Unrestricted model with at least one nodal covariate included
 4. Unrestricted model with at least one edge covariate included.

Complete the RSiena model using the s50 data set as demonstrated in the lecture. Remove the non-significant model terms and produce an RSiena model summary consisting of only significant terms. Submit a copy of your final model summary.