Text Mining for Social Science

Term Paper Instructions for Data Science Students

by Ruben Durante, Hannes Mueller and Nandan Rao

This document provides a framework for the term paper for the course Text Mining for Social Science. This framework is set-up to provide students with enough freedom to develop their own research ideas while at the same time giving some common starting position for this work. Important points:

- **Handing in date is June 21 2020, 12.00 (noon).**

- Up to three students can cooperate when writing the term paper.

- The students need to use tools provided in the course. The tools need to be explained when used.

- The term paper has to be between 10 and 15 pages in length. However, appendices are allowed for example for robustness checks or longer mathematical proofs.

- Mandatory sections: Introduction, Related Literature, Theory or Data Section, Results Section, Conclusion

- Bibliography needs to signal that the students have done a proper literature research (see Research Hacks).

- We understand you are under severe time constraints. Please don't react to this constraint by being sloppy and overstretching. Simple, focused, correct arguments are better than wrong, broad claims.

- Remember the course name: Text Mining for Social Sciences. Try to relate to the social science literature.

The key of the term paper is that the teams signal that they understood and are able to use the material and kind of arguments from the lectures correctly. The evaluation of the term paper will take into account this use of the material, the soundness of the arguments made, the originality and effort put into the three main parts of the thesis (literature, theory, data analysis). These three parts need to be linked with each other.

Research Hacks 1) Good literature search and summary is completely underestimated by students. Typically two things arise: a) students come up with sources which are strange (unpublished or odd journals) b) the bibliography looks like the syllabus with a handful of additional sources. This leads to literature overviews which provide detailed summaries of articles but which are not

really helpful in understanding the broader points. What you need to do is to think of issues/topics instead of articles.

2) Snowball like a pro: Don't just read one article but read the bibliography of the most relevant articles and follow it. Mark articles that seem relevant in the bibliography and look for them on Google Scholar. Once you have found them don't just download the articles blindly but read the abstract of each before you do. When you have ten articles start reading the intros. Stop reading if you think it's not relevant - keep reading if you think it is. Don't read more than the intro but make a list of articles that remain interesting after you have read the intro. Finish the article you started with, then go to the articles that seemed relevant. Repeat for several hours. Important trick: if you have old articles that are relevant you can also *forward snowball* by typing them into Google Scholar and then looking at who cites them. Don't print out anything until you are sure you need to understand an article properly! Printing is a productivity surrogate and kills trees.

4) Even though this is a course in text-mining don't underestimate the role played by "theory". Your work should be linked explicitly to some concept you want to test. Descriptive analysis is fine useful but making sense of the data is essential for a good grade. This is impossible without some "theory". Think of what the data-generating process looks like. If you can't model this process explicitly think about which method is a good proxy.

5) Research is non-linear. This means you will not see progress sometimes - reading and looking at the data more will confuse you more or you will simply not have an idea. As long as you stay focused on the question and keep working on it an idea should come up. Don't be impatient - creativity requires the combination of a lot of information. Get out of the non-linearity by doing simple data analysis first and asking yourself why you see the patterns in the data that you see.

1 What are the trolls up to?

1.1 Introduction

Some governments have run political influence campaigns in other countries in which they use fake profiles on social media which imitate local population to spread content. This is done with the explicit goal to influence public opinion and, perhaps, even election outcomes. This is best documented for Russian campaigns in the US but is a more wide-spread phenomenon.

We are interested in whether definitions laid out in a research project at Princeton by Diego Martin and Jacob Shapiro (Trends in Online Foreign Influence Operations) can be used productively to analyze tweet content on twitter.

Martin and Shapiro (2019) identify 53 foreign influence efforts (FIE), in 24 targeted countries, from 2013 to 2019. Fully 70% of the attacks were conducted by Russia, with China, Iran, and Saudi Arabia accounting for the remainder. According to their report the strategies used by trolls can be classified as: defame, persuade, polarize, change the political agenda and weaken the institutions.

1.2 Question

Go to the elections integrity website of twitter and look at their data archive[1] and understand the content provided on this webpage *in the data archive*.

1) Review the literature on political influence efforts (FIE). Discuss the relevant literature on this issue (including social science literature). Some leads are given below. Explain the reason for defamation, persuasion, polarization, changing the political agenda and weakening of institutions in the light of this literature. Conceptualize at least one concept more in detail and use it in your answers to 2), 3) and 4). (20 points)

2) Download some of the tweet data (and content data only if you think you can use it) and produce some useful summary statistics that you think are relevant. Describe and interpret the patterns you see. (20 points)

3) Develop measures for at least one of the strategies used by trolls according to Martin and Shapiro (2019). It will not be immediately clear how to do this and will require some conceptual thinking, creativity and possibly some manual work. You can use the content itself, add content/labels yourself, or look to outside sources of labelled data. Explain your measures theoretically using the definitions provided below and your own conceptualizations and then implement them empirically. It is very important that you link the ideas in 2) to your text features developed in 3). Do so explicitly and explain why you do what you do. (40 points)

4) Present results coming out of the measures you have developed. Come up with additional things to look at that you can only show with the measure you have developed. For example, answer these questions: How are they chang-

---

ing over time? Are some trolls specializing or are they all changing strategies together? (20 points)

2 Definitions

Defame   A direct attack against a person, intended to discredit him or her. Example: IRA create fictitious social-media personas to spread falsehoods and promote messages against Hillary Clinton.

Persuade   An influence effort where the goal appears to be directly shift political views about an issue or actor in an identifiable direction. This affects the median voter in one direction. Example: Trolls create blogs and fake news to incentive people to vote in favor of Donald Trump. These trolls do not push the Hillary Clinton campaign at the same time.

Polarize   An attack aims to create polarization on issues. This is persuasion on both sides of an issue to move individuals to the extremes. This affects the variance of the decision because it looks for pushing one political goal in two opposite directions. Example: Nearly 600 Russia-linked accounts tweeted about the US Affordable Care Act - ObamaCare. The majority of the nearly 10,000 tweets on the Affordable Care Act seemed intended to pit one side against the other, not to advance a particular policy with respect to the ACA.

Shift political agenda   The efforts add something new in the political agenda. Example: Russia plans to eliminate the American energy threat, as an exporter of this source, and to do so by influencing social media users, American voters, and public offcials.

Undermine institutions   The objective is to reduce the credibility/reputation of one or more institutions in the target country. The Democratic Party or the Republican party are not institutions. Example: Russian media outlets circulated a false story about a state prosecutor in Berlin failing to prosecute an alleged rape by immigrants.

References

[1] Allcott, Hunt and Matthew Gentzkow (2017) Social Media and Fake News in the 2016 Election. Journal of Economic Perspectives.

[2] Boxell, Gentzkow, and Shapiro (2020) Cross-Country Trends in Affective Polarization. http://web.stanford.edu/~gentzkow/research/cross-polar.pdf

[3] Bradshaw, S. & Howard, P. N. (2018), 'Challenging truth and trust: A global inventory of organized social media manipulation', The Computational Propaganda Project .

[4] Cagé, Hervé and Viaud (2017) The Production of Information in an Online World: Is Copy Right? Unpublished manuscript.

[5] DellaVigna and Kaplan (2007) The Fox News Effect: Media Bias and Voting. Quarterly Journal of Economics.

[6] Duclos, Esteban and Ray (2004) Polarization: Concepts, Measurement, Estimation. Econometrica.

[7] Esteban and Ray (2011) Linking Conflict to Inequality and Polarization. American Economic Review. https://www.econ.nyu.edu/user/debraj/Papers/ConPolIneq.pdf

[8] Gentzkow and Shapiro (2010) What Drive Media Slant? Evidence from US Daily Newspapers. Econometrica

[9] Hindman, M. & Barash, V. (2018), 'Disinformation, and influence campaigns on twitter'.

[10] Martin and Shapiro (2019) Trends in Online Foreign Influence Efforts. https://scholar.princeton.edu/jns/research-reports

[11] Martin and Yurukoglu (2017) Bias in Cable News: Persuasion and Polarization. American Economic Review.

[12] Mullainathan, Sendhil, Joshua Schwartzstein and Andrei Shleifer (2008) Coarse Thinking and Persuasion. Quarterly Journal of Economics.

[13] Müller, Karsten and Carlo Schwarz (2018) Fanning the Flames of Hate: Social Media and Hate Crime. Unpublished manuscript.

[14] Nimmo, B. & Karan, K. (2018), 'Trolltracker: Favorite russian troll farm sources. Measuring the websites and accounts the internet research agency shared most'.

[15] Stewart, L. G., Arif, A. & Starbird, K. (2018), Examining trolls and polarization with a retweet network, in 'Proc. ACM WSDM, Workshop on Misinformation and Misbehavior Mining on the Web'.