FAHA – Assignment 01

Anonymized

2021-09-15

Question 1: Correlations – Recall

The Russian diaries

In case of the Russian diaries the correlation is between the diary's using style and the historical time. The diary's using style is derived through machine learning (ML) that focuses on the word patterns surrounding word 'diary'. ML tells us how the writer used the diary – ML distinguish 5 main categories of diary keeping/writing (spirit, routine, literary, interpersonal, material form). The research used archive of 4,000 Russian diaries written in the course of 100 years (since 1900 until 2000).

Such a research helps to answer question: 'How do people understand the act of diary-writing?' The authors of the research found approximately stable proportions of 5 diary-writing categories over 10 decades. There were not any clear changes of any category proportion over the course of the time and the proportions of each category were more or less same in every decade that research covered. The conclusion is that people understood and did the act of diary-writing in the same way in the course of the whole 20^{th} century.

Question 2: Correlations – Apply

Introduction

I focus on opinion dynamics by introducing the classical agenda-setting theory (AST) into new communication context. The AST focuses on what the public is thinking about, not exactly what the public thinks – AST shifts from *predicting public's position on certain issue* to *predicting what issues gain public attention*. AST defines issue as a problem publicly recognized as a public problem and the agenda as the set of the most important issues. The classical applications of AST compare the results of public opinion surveys on most important issues with the numbers of news-pieces mentioning each issue in the most respected newspapers and TV newscasts. The classical AST correlation is between the amount of media content on issues and the public attention to these issues.

Correlations in my focus

I want to study communication behavior of users on social network sites (SNS), e.g. Twitter or Facebook. I want to predict the issues the users mention in their own SNS statuses according the previous mentions of issues by the users' connections (accounts they follow, interact with). I also want to predict users' issues mentions by the previous issues' mentions in online media the users follow (e.g. cite them in their statuses, follow their SNS accounts). Finally, I also want to predict the online discussions on issues the users enter (e.g. commenting a status of account they follow, re-tweet) according the previous issues' mentions by their connections and online media. Schematically:

$$mentions_{connections} + mentions_{online\ media} => p(issue\ in\ status_{user}) + p(discussion_{user})$$

Scholarly claims

Classical AST shows that the more the old media covered an issue the more important the issue was for the public. *Correlations in my focus* (see above) could tell us whether the same holds true for the online media and connected accounts on the SNS. Analogically to the example of the French Revolution transcripts analysis (video 2.1) we could distinguish between agenda leaders and agenda followers – leaders use more the mere statuses and focus the attention of the rest of SNS users on the issue, followers use more discussions and amplify the public attention started by leaders. We could also find which users are more influenced by other users and which by online media. Finally, we could also investigate the flows of issues through agendas of SNS users over the course of time and whether the issues enter SNS from online media or the users.

What could we learn?

By the content analysis we could measure both type of agendas – media and personal. Content analysis of issues covered in online media done via machine learning (ML) tells us about online media agenda, about the set of the most covered issues and amount of content devoted to each of them. The communication behavior on SNS of the user (i.e. statuses, discussions) tells us about user's personal agenda, about issues so important to her that she writes statuses on them and enters discussions on them.

Correlation between media agenda on one hand and personal agendas on the other tells us whether the AST holds true also in the realm of the new media. We could find how many personal agendas correlates with agendas of the followed media, we could also find whether correlates the agenda of the online media as a whole and the agenda of the public as a whole. In case of successful finding of strong correlation we learn that AST still holds true.

We could distinguish *leaders* from *followers* by the correlations. Present agenda of *leaders* should correlate with future agenda of *leaders' connections*, the present leaders' agenda could correlate with past agenda of *online media*, but not with the past agenda of *leaders' connections*. Present agenda of *followers* should correlate with past agenda of *online media* or *followers' connection* – correlation with future agendas is not important.

We could distinguish types of *leaders* and *followers* by the power of the correlations. Agenda of *social leaders/followers* would correlate more with agenda of other users, agenda of *institutional leaders/followers* would correlate more with the agenda of online media, and agenda of *mixed leaders/followers* would strongly correlate with both agendas (of other users and online media).

Finally, we could describe complete flow of issues in the communication network. We could find which type of users set on the public agenda the issues from media agenda. We could investigate dynamics of agenda-setting process – how and which type of users initiate focusing of public attention, which propagate issues through the network, which amplify the issues etc.