Identifying depression related behaviour in Facebook - An experimental study

Zoltán **Kmetty**
Károly **Bozsonyi**

Email: kmetty.zoltan@tatk.elte.hu

Introduction

In the recent years, an increasing tendency of suicide has been observed in some western countries. One of the research streams who picked up on the increasing trends, started to analyze the online footprints of suicide and depression. In recent years Twitter and Instagram were the most studied platform for investigate depression and suicide related behaviour. Most of these studies focused on word usage and used only single source data. Another approach used by Cheng et al. as, they conducted an on-line survey among Weibo users (n=974), where beyond a classic questionnaire participants' Weibo posts published in the public domain were also downloaded. As it can be seen, Facebook is rarely used for these kinds of analysis.

Data and methods

Our study uses a novel joint data source of combined Facebook and survey data. After an informed consent obtained, respondents were asked to log-in to FB on the interviewers' notebook and to download their FB profile archive. 150 respondents took part in our study. the data covers a wide range of Facebook activities: posts, comments, likes and reactions, pages, friends, profile, and ads data. Besides sharing their Facebook data, participants had to fill out an online questionnaire. Questions about politics, media usage, self-representation. spare-time activities and music preferences were asked from the participants. Above that, we asked the participants to fill out a modified version of Patient Health Questionnaire (PHQ-9). Two indicators of depression are extracted by ML Factor Analysis based on the PHQ-9 questions, a cognitive depression scale (CDS) and a psychosomatic depression scale (SDS) one.

For this study we used the temporal dynamic of Facebook activity of the users and the ads interest categories. Facebook categorizes every user for sales for advertising. This is an algorithmic machine learning classification of the users based on their own likes, activities, and used keywords and also based on their friends' preferences.

For the temporal analysis we selected the 2019 and 2018 years of Facebook usage, and calculated the ratio of days with Facebook usage. Then we calculated two indicators: the first one was the difference between active day ratio in 2019 and 2018. The second indicator was the absolute value of the previous indicator. The calculations were done separately for posts, reactions on friends and reactions on pages.

Selected results

The analysis found a moderate but significant correlation (0.33) between cognitive depression level and absolute change of Facebook posting activity. The detailed analysis showed that the effect was stronger for males than females (see figure 1 for the marginal predictions)
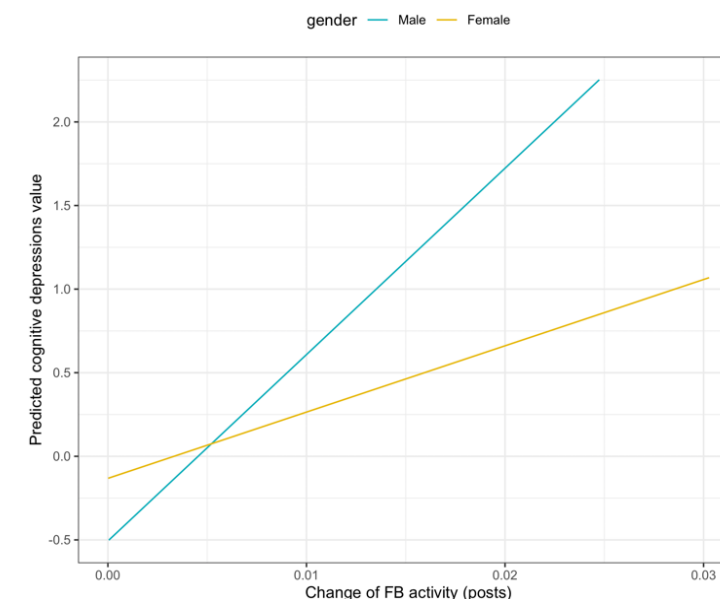
The effect was much smaller in the case of reactions on friend and we didn't find significant relationship when we studied the reactions on groups data. Psychosomatic depression didn't correlate with any of the measures.

The second part of the study focused on the possible correlation between FB generated ads interest categories and depression level. We selected those ads categories where at least 10 percent of our sample were categorized in (N=1561). We calculated the relationship of each category belonging and the two depression scales (based on ANOVA and ETA statistics). Then we searched for patterns within the results. One interesting thing came out. Those who scored low on psychosomatic depression scale had higher probability to categorized into finance related categories by Facebook, like banking, finance or car renting.