Introduction

Educational research has been traditionally focused on the academic performance of students. Academic achievements are usually considered as the most important or even the only outcome of pedagogical interventions and the whole education system. As a result, important aspects of students' well-being, including their emotional well-being, are often neglected. In recent years, the well-being of students has been gaining more attention and interest among researchers and policymakers [1], however, this topic remains underexplored. One reason is that it is difficult to measure both the emotional state of individuals and their behavior. These challenges might be partially overcome with the help of digital traces and, in particular, social media data. For instance, it has been shown that Facebook posts might be used to predict depression [2] and Twitter data has been used to study emotional reactions to terrorist attack [3] and celebrity suicide [4].

A similar approach might be used to measure the emotional pulse of schools. One could collect public posts of students from a particular school on social media, apply sentiment analysis to estimate the sentiment of each post and then average over all students and over a certain time period. The resulting measure could be used to study the dynamics of students' emotions or to compare different schools, e.g. schools located in various parts of the city, schools with different educational outcomes, etc.

In this paper, we test this approach using the data on 20,348 students from 596 schools. We collected 309,586 public posts made by users from Saint Petersburg on a popular Russian social networking site who indicated their schools in profile. We built a model to predict the sentiment of these posts and additionally validate the model to check if the predicted sentiment is related to the actual mood of users. We then study the dynamics of students' emotions over time and compare schools with different educational outcomes.

We computed vector representations of Russian words by training a fastText [5] model on a large corpus of VK posts (2B tokens). We then represent short texts as 300-dimensional vectors by averaging over all their constituent words. We use this representation to train a logistic regression on a data set of Russian tweets with annotated sentiment [6]. For our model, the AUC for detecting positive emotions is 0.77 and for detecting negative emotions is 0.80. Note that unlike tweets VK posts are not limited to 140 characters. However, most of VK posts are short texts (97.5% of the posts are less than 280 characters in our sample). It means that, in practice, a model trained on a Twitter data set might be applied to VK data and vice versa.

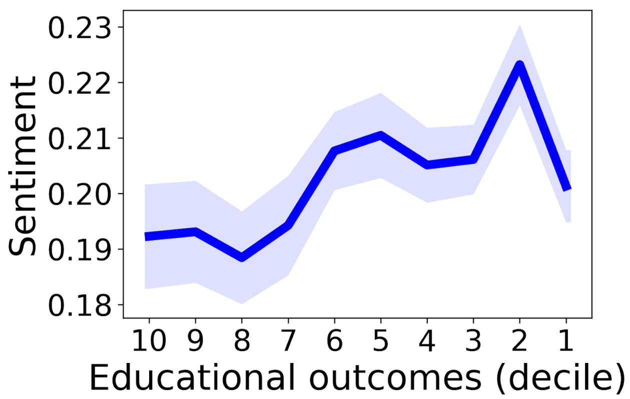
Validation

To check if the sentiment of posts is related to the mood of users we use data on 150 participants of TeenLife project. TeenLife participants provided access to their public posts on VK and also filled in Patient Health Questionnaire (PHQ-9) designed for screening, monitoring and measuring the severity of depression [7]. We find that the average sentiment of all posts that were written by TeenLife participants during the 4 months of the study is correlated with their PHQ-9 score (Pearson's $r = -0.34$). It means that the sentiment of posts measured by the model might not only reflect emotions but also slow-moving feeling states such as moods.
We also find that there is not much variation in sentiment during the day. However, posts become significantly less positive after midnight. The changes in posts sentiment during the week are V-shaped with the minimum on Thursday and additional local minimum on Monday.

January is the most positive month. That is probably due to New Year and long winter holidays. Summer months are more positive while the autumn months are less positive. These patterns are consistent with previous research [8-11] and serves as further validation of our approach.

Results

We compare average posts sentiments for schools with different educational outcomes. We split all school into 10 deciles according to the results of their graduates on USE. Our results suggest a complex relationship between academic performance and the valence of posts. The posts of students from the bottom 40% of schools are the least positive. The sentiment becomes more positive for the next 40% and even more positive for the next 10%. However, for the top 10% of schools, there is a significant drop in the valence of posts. For the bottom 90% of schools the correlation between posts sentiment and their academic performance is 0.19 ($P = 1.3 \times 10^{-5}$), while for the top 20% of schools the correlation is -0.37 ($P = 3.7 \times 10^{-5}$).



Shaded regions are bootstrapped confidence intervals

References

1. Choi, A. (2018). Emotional well-being of children and adolescents.
2. Eichstaedt, J.C., Smith, R.J., Merchant, R.M., Ungar, L.H., Crutchley, P., Preoţiuc-Pietro, D., Asch, D.A. and Schwartz, H.A., 2018. Facebook language predicts depression in medical records. Proceedings of the National Academy of Sciences, 115(44), pp.11203-11208.
3. Garcia, D., & Rime, B. (2018). Collective Emotions and Social Resilience in the Digital Traces After a Terrorist Attack.
4. Niederkrotenthaler, T., Till, B., & Garcia, D. (2019). Celebrity suicide on Twitter: Activity, content and network analysis related to the death of Swedish DJ Tim Bergling alias Avicii. Journal of affective disorders, 245, 848-855.
5. Bojanowski P, Grave E, Joulin A, Mikolov T (2016) Enriching word vectors with subword information. arXiv:1607.04606 Preprint, posted June 19, 2017.
6. Mozetič, I., Grčar, M., & Smailović, J. (2016). Multilingual Twitter sentiment classification: The role of human annotators. PloS one, 11(5), e0155036.
7. Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. Journal of general internal medicine, 16(9), 606-613.
8. Egloff, B., Tausch, A., Kohlmann, C. W., & Krohne, H. W. (1995). Relationships between time of day, day of the week, and positive mood: Exploring the role of the mood measure. Motivation and emotion, 19(2), 99-110.
9. Ryan, R. M., Bernstein, J. H., & Brown, K. W. (2010). Weekends, work, and well-being: Psychological need satisfactions and day of the week effects on mood, vitality, and physical symptoms. Journal of social and clinical psychology, 29(1), 95-122.
10. Golder, S. A., & Macy, M. W. (2011). Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. Science, 333(6051), 1878-1881.
11. Stone, A. A., Schneider, S., & Harter, J. K. (2012). Day-of-week mood patterns in the United States: On the existence of 'Blue Monday','Thank God it's Friday'and weekend effects. The Journal of Positive Psychology, 7(4), 306-314.