See and Read: Detecting Depression Symptoms in Higher Education Students Using Multimodal Social Media Data

Anonymous Author(s)
Affiliation
Address
`email`

1 Motivation

Mental disorders have been alarmingly increasing in the worldwide population [1], characterized by a combination of abnormal thoughts, emotions, and behavior. One of the most common mental disorder is depression, globally estimated as more than 300 million cases [1]. Particularly, Brazil has the highest prevalence of Major Depressive Disorder (MDD)[1] among South American countries, with nearly 5,8% [1]. This is not only true to the general population but has also been progressively observed in the academic environment, where graduate students are more than six times likely to experience depression and anxiety, compared to the general population [2].

However, before an individual with depression can be treated, this disorder must be detected. Besides the statistics showing that roughly only 50% of the cases are detected [3], some individuals might not have money, knowledge, or may have fear of the social stigma to look out for help [4]. Because of that, the disorder may remain undiagnosed, which may further aggravate its symptoms. Thus, although the most reliable way to screen for depression is the clinical diagnosis, it is important to enhance other passive diagnosis methods beyond the ones based in active consultation.

Another common way of detecting MDD is relying on questionnaires, such as the *Beck Depression Inventory* (BDI-II) [5]. It evaluates the severity of depression through a final score obtained from the answers given into the questionnaire. Notwithstanding, given that these questionnaires should also be handled by professionals and the individual with MDD may not always have access to them, one question that arises is if we could use the data generated by the individuals themselves to detect depression. Furthermore, the questionnaire-based criteria have been defined years ago. As the world develops and evolves, the criteria to detect MDD should also change to go along with the new technologies that impact everyday routine and behavior. This is especially the case of online environments, where the individual may express depression symptoms in a way different from the established criteria.

Along these lines, social media such as microblogs and social networking sites poses as a promising environment to investigate depressive symptoms and behavior. Several previous studies have already investigated social media features to characterize a user with a depressive behavior [6, 7, 8, 6, 9, 10]. To accomplish the detection task enriched with the features from either these paths, the vast majority of works benefit from Machine Learning techniques. However, those previous work focused on either text or image features separately, or engineering metadata to fed ML systems. The question that arises is if modern deep learning techniques have the mechanisms to generate classifiers directly from the user-generated data, in a end-to-end fashion [11, 12]. Thus, in this work, we use the data shared by students through Instagram to induce MDD patterns directly from the user-provided content using deep learning methods. We show that our deep multimodal classifier performs better than the unimodal architectures for our dataset at the task of screening depression.

[1] In this work, we use MDD and depression interchangeably.

Submitted to 33rd Conference on Neural Information Processing Systems (NeurIPS 2019). Do not distribute.

2 Methodology and Results

To create the dataset, we posted invitations through several official social media pages, and email lists of University XX[2], Brazil. Students, then, voluntarily answered the questionnaire within the invitation, resulting in a total of 416 answers, in which 221 provided access to their personal Instagram profiles[3], where we downloaded a total of 8188 posts (images and captions). Furthermore, to estimate the impact of observation period in the classification accuracy, we evaluate the dataset for three observation periods: 60, 212 or 365 days back from the answer to the questionnaire, as have been done in previous work [7, 9]. Thus, for each observation period, we generated 10 different stratified datasets, each divided in training, development and test sets.

For textual representation, we use bag-of-words (BoW) as baseline, FastText [13] and ELMo [14], with pre-trained weights on a dump of the Portuguese Wikipedia. To obtain a single caption representation, we employ both simple average, and pmean [15] across words. For visual representation, we use ResNets [16] of different sizes with pre-trained weights on ImageNet, and ResNeXt WSL with pre-trained weights on Instagram pictures [17]. The fusion architecture follows an early fusion where we project both visual and textual embeddings to the same dimensional space, followed by a Hadamard product and a final linear classification layer. We use simple average for the textual representation when using fusion. All models perform a binary classification problem — high severity (59% of examples) vs. low severity of depressive symptoms. The final result targets the student prediction obtained from the average of every post probability belonging to the high severity class.

Figure 1 presents the obtained results using an NVIDIA DGX-1. Four fusion models achieved the best scores, all with an observation period of 212 days. The best model has a F1 score of 0.778 (ELMo + ResNet-18), while the best textual and visual models have 0.746 (ELMo + avg), and 0.708 (ResNeXt), both for 212 days. Furthermore, from the 10 best F1 scores, only one was unimodal: ELMo + avg. ELMo also consistently performed better as the textual representation.

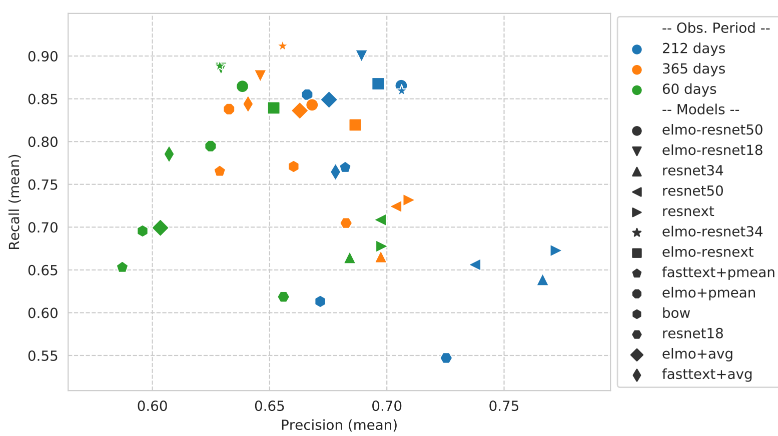

Figure 1: Metrics for prediction of our positive class using our proposed models with different observation periods. All results are for students predictions, not posts, over 10 different datasets

3 Final remarks

Detecting depression from self-generated content in social media is a challenging task, with natural benefits to the society as a whole. Here, we show that we can improve detection scores using a deep multimodal classifier compared to using unimodality by a maximum of 7% in F1 score. We also demonstrate insight on how different observation periods impact on model accuracy, showing that using 212 or 365 days results in better model performance compared to 60 days. In the future we intend to invest on the particularities of the students sample and on explanation methods to help on elucidating this aggravating disorder.

[2]Omitted to respect the double blind process.

[3]The research was conducted under the approval of the ethical committee of the University XX, CAAE: 89859418.1.0000.5243.

References

[1] WHO. Depression and other common mental disorders: global health estimates. 2017.

[2] Teresa M Evans, Lindsay Bira, Jazmin Beltran Gastelum, L Todd Weiss, and Nathan L Vanderford. Evidence for a mental health crisis in graduate education. *Nature biotechnology*, 36(3):283, 2018.

[3] Alex J Mitchell, Amol Vaze, and Sanjay Rao. Clinical diagnosis of depression in primary care: a meta-analysis. *The Lancet*, 374(9690):609–619, 2009.

[4] Atle Roness, A Mykletun, and AA Dahl. Help-seeking behaviour in patients with anxiety disorder and depression. *Acta Psychiatrica Scandinavica*, 111(1):51–58, 2005.

[5] Aaron T Beck, Robert A Steer, and Gregory K Brown. Beck depression inventory-ii. *San Antonio*, 78(2):490–8, 1996.

[6] Guangyao Shen, Jia Jia, Liqiang Nie, Fuli Feng, Cunjun Zhang, Tianrui Hu, Tat-Seng Chua, and Wenwu Zhu. Depression detection via harvesting social media: A multimodal dictionary learning solution. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI-17)*, pages 3838–3844, 2017.

[7] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. Predicting depression via social media. *ICWSM*, 13:1–10, 2013.

[8] Andrew G Reece and Christopher M Danforth. Instagram photos reveal predictive markers of depression. *EPJ Data Science*, 6(1):15, 2017.

[9] Sho Tsugawa, Yusuke Kikuchi, Fumio Kishino, Kosuke Nakajima, Yuichi Itoh, and Hiroyuki Ohsaki. Recognizing depression from twitter activity. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3187–3196. ACM, 2015.

[10] Michelle Morales, Stefan Scherer, and Rivka Levitan. A cross-modal review of indicators for depression detection systems. In *Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology — From Linguistic Signal to Clinical Reality*, pages 1–12. Association for Computational Linguistics, 2017.

[11] Michelle Renee Morales. Multimodal depression detection: An investigation of features and fusion techniques for automated systems. 2018.

[12] Michelle Morales, Stefan Scherer, and Rivka Levitan. A linguistically-informed fusion approach for multimodal depression detection. In *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*, pages 13–24, 2018.

[13] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5:135–146, 2017.

[14] Matthew E Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. Deep contextualized word representations. *arXiv preprint arXiv:1802.05365*, 2018.

[15] Andreas Rücklé, Steffen Eger, Maxime Peyrard, and Iryna Gurevych. Concatenated power mean word embeddings as universal cross-lingual sentence representations. *arXiv preprint arXiv:1803.01400*, 2018.

[16] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[17] Dhruv Mahajan, Ross Girshick, Vignesh Ramanathan, Kaiming He, Manohar Paluri, Yixuan Li, Ashwin Bharambe, and Laurens van der Maaten. Exploring the limits of weakly supervised pretraining. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 181–196, 2018.