BlueMemo: Depression Analysis through Twitter Posts

Pengwei Hu[1*] , **Chenhao Lin**[2*] , **Hui Su**[3] , **Shaochun Li**[4] , **Xue Han**[1] , **Yuan Zhang**[1] and **Jing Mei**[1]

[1]IBM Research, China
[2]Xi'an Jiaotong University, China
[3]Pattern Recognition Center, Wechat AI, Tencent Inc, China
[4]IBM Watson Health, China
hupwei@cn.ibm.com, linchenhao@xjtu.edu.cn, aaronsu@tencent.com,
{lishaoc,bjhanxue,zngyuan,meijing}@cn.ibm.com

Abstract

The use of social media runs through our lives, and users' emotions are also affected by it. Previous studies have reported social organizations and psychologists using social media to find depressed patients. However, due to the variety of content published by users, it isn't effortless for the system to consider the text, image, and even the hidden information behind the image. To address this problem, we proposed a new system for social media screening of depressed patients named BlueMemo. We collected real-time posts from Twitter. Based on the posts, learned text features, image features, and visual attributes were extracted as three modalities and were fed into a multi-modal fusion and classification model to implement our system. The proposed BlueMemo has the power to help physicians and clinicians quickly and accurately identify users at potential risk for depression.

1 Introduction

Social media services (such as Twitter, Facebook, Weibo) have become popular platforms for content sharing and information dissemination. The widespread use of social media also provides a new convenient medium for users to express their emotions. Studies have shown that social media has become a powerful platform for people with depression to self-disclose psychological problems and seek social support [Manikonda and De Choudhury, 2017]. Users' posts are an expression of their inner emotions, and discovering the clues offers a promising new opportunity to spot depression. More than 350 million people worldwide are affected by depression [Organization, 2010], and without proper treatment, people with depression will increasingly develop severe symptoms. Early detection and timely treatment may benefit the prognosis of most adult patients with mental illness [Picardi *et al.*, 2016]. Therefore, it is necessary to carry out early detection and intervention for patients with depression to minimize its impact on public health [Reay *et al.*, 2011]. The traditional approaches for depression detection require the

---

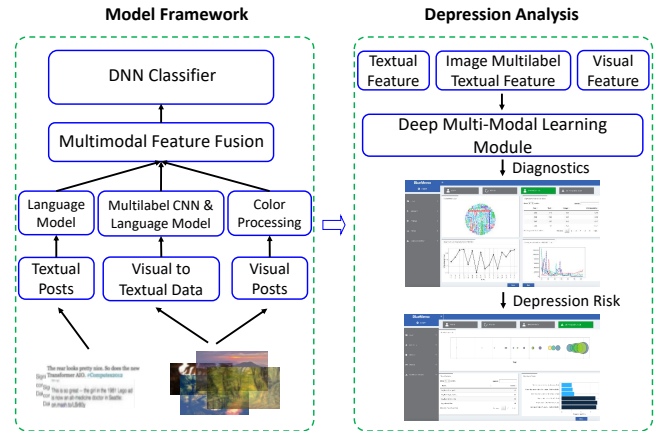*Both authors contributed equally to this research.

Figure 1: Illustration of proposed model for depression analysis.

psychiatrist to measure the severity of depression in a face-to-face interview. With limited resources, the effectiveness of this clinical diagnosis is not satisfactory. Encouragingly, advances in artificial intelligence have made it possible to detect depression-related indicators from between the lines automatically. Many clinical studies have shown that text and image features extracted from video signals in interviews can be used to detect depression objectively [Stolar *et al.*, 2015; Gillespie *et al.*, 2017]. Other studies of social networks have found that users can be divided into groups based on their profiles or information posted [Ding *et al.*, 2019; Hu *et al.*, 2017; He *et al.*, 2018]. Inspired by such work, the new study has used visual social media, such as images and text on twitter, to encode predictors of depression with widespread success. Shen et al., for the first time [Shen *et al.*, 2017], attempted to detect users' depression by combining users' various demographic information and visual features of pictures. Gui et.al [Gui *et al.*, 2019] used features such as text features posted by users and color saturation of images to check the depression of Twitter users.

However, as users vary greatly in their use of text and become more inclined to publish image content, a well-trained model will lose its power by failing to capture the subtle expressions of emotion in text and the target information im-

plicit in photo content. Fortunately, the latest research on multi-modality [Liu *et al.*, 2018] shows that by incorporating more feature expressions of users into the learning process, the model can fully understand users' mental state and significantly improve performance. As such, there is an urgent need for developing a system to learn from users' emotions and desires <mark>in a multi-modal way to help us find patients with depression on social media</mark>. In this paper, on the basis of the multi-modal user posts, we demonstrated a new analysis system of depression: BlueMemo[1]. <mark>Our system feeds textual feature, image multi-label textual feature, and visual feature of social media users into the learning process, then uses tensor transformation to fuse multi-modal features to optimize the detection strategy.</mark> BlueMemo is presented as a browser-based service interface, where users can freely customize the detection target and training population on the interactive page, and obtain a patient analysis experience that resembles clinical diagnosis. Figure 1 illustrates the architecture of our system.

2 Blue Memo

2.1 Proposed Approach

Feature Extraction. Three types of features are considered and extracted in our system, textual feature, visual to textual feature and visual feature. We extracted textual feature representation from users posted tweets by using a pre-trained Bert[Devlin *et al.*, 2019], which has been proved effectively in contextualized sentence understanding and representation. In order to take full advantage of user posted information in social media, users posted images were also acquired for visual and visual to textual feature extraction. A deep CNN-based multi-label classifier was trained using COCO datasets [Lin *et al.*, 2014], which contains 80 common objects and can be effectively used to describe the images. For each posted image, its multi-label text information can be predicted by using this classifier. Finally, the predicted multi-label text information was converted into textual feature representation by using the Bert model. In such way, visual to textual feature can be acquired. In addition, the visual color feature representation was extracted by using posted images. We adapted conventional image processing method to generate color-related attributes, including the image histograms in RGB and HSV color space. This color-related attribute was then normalized and fused as the visual feature representation.

Feature Fusion and Classification. To generate more robust feature for depression detection and analysis, similar to the feature fusion method in [Lin and Kumar, 2017; Lin and Kumar, 2018] ,we combined the extracted multi-modal features. Let f_t be the textual feature extracted from the tweets, f_v be the visual features calculated from the posted images and f_{vt} be the visual to textual feature acquired from the image multi-label. The multi-feature f_m was combined and trained by using a tensor fusion network [Liu *et al.*, 2018], which can efficiently learn the features from different modalities. During the network training process, the

multi-feature f_m was combined by using a tensor multiplication operation and soft-max function was used for classification. Finally the input multi-features are classified into two classes, i.e. depression user and non-depression user.

2.2 System Main Features

BlueMemo is a system designed to detect people with depression from social media and further assist psychologists in their analysis. It goes beyond the clinical diagnosis of depression based on doctor-patient interviews and gives psychologists the ability to retrieve and visualize users' past social media posts. BlueMemo can assess a user's risk of depression based on a customized scope, and psychologists can receive systematic diagnostic support as well as more detailed understanding of a user's psychological expression through interactive functions. More specifically, BlueMemo provides the following facilities for psychologist:

User/Posts. provides a quick preview for system users. To some extent, users' posts have replaced the traditional Q&A record of interview diagnosis. Therefore, we set up the preview function of the user's homepage information and the user's posts, respectively.

Diagnostics. presents four visual analysis results for system users. Once we have identified the user to test, we can analyze by clicking diagnostics-button. The $posted word cloud$ is an intuitive display of the modal data of the text. $Histogram comparison in HSV color space$ is a visual display of the image visual color feature. The $depression on an annual basis$ can be used to look at users' depression risk scores yearly for different modality. $Diagnosed with depression for the first time$ provides a time when the first depression risk score exceeds the threshold, facilitating the psychologist to pinpoint the onset time.

Depression Risk. mainly presents the final assessment score. $Trend of user posts$ are used to show the overall trend in the number of posts from subjects. Considering that visual feature has reference basis in clinic, we show the $visual feature$ count table on this page. Depression score provides a user's final risk assessment score and compares it with the mean of all users.

2.3 Evaluation

We followed the rules in [Shen *et al.*, 2017; Shen *et al.*, 2018] to collect the twitter dataset, and used it to train and evaluate the model. The data collection rule is to label users' self-reports as depression. Specifically, if a user ever posted a message on twitter claiming to be diagnosed or experiencing depression, we marked it as depression user and took the posts before the time of this message as a training sample. The dataset consists of two subsets, D_1(depressed users) and D_2(non-depressed users), which mainly includes tweets and images posted by users. During the offline training, all 1,402 depressed users and 292,564 tweets in D_1 were enlisted. Non-depressed users were randomly selected from D_2 with more than 300 million non-depressed users to combine with D_1 into training, testing, and evaluation set. When evaluating the model, we conducted 5-fold cross-validation experiments on the proposed method. We used F1 scores for

---

[1] A demo video can be viewed at: https://youtu.be/SkLR3kIB8dg

evaluation, and the F1 scores of MDL [Shen *et al.*, 2017], COMMA [Gui *et al.*, 2019] and proposed models were 84.9, 90.0 and 96.1, respectively. The detection model of Blue-Memo had better performance than previous work.

3 Conclusion

We propose a new depression analysis system for people at risk for depression on Twitter. People are increasingly inclined to express their feelings, to show their life state, so this system is designed to fully explore the multi-modal data in social networks and detect users' abnormal mental state. The proposed model extracted text features, image multi-label textual feature, and visual feature and used tensor transformation based classifier to detect those depression groups. The system realizes the practical detection with high precision and carries on the visual analysis of the detection result. The system is promising to screen online users for depression and has the potential to help psychologists review patient histories and support diagnosis.

References

[Devlin *et al.*, 2019] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, 2019.

[Ding *et al.*, 2019] Kaize Ding, Jundong Li, Shivam Dhar, Shreyash Devan, and Huan Liu. Interspot: interactive spammer detection in social media. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pages 6509–6511. AAAI Press, 2019.

[Gillespie *et al.*, 2017] Stephanie Gillespie, Elliot Moore, Jacqueline Laures-Gore, Matthew Farina, Scott Russell, and Yash-Yee Logan. Detecting stress and depression in adults with aphasia through speech analysis. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5140–5144. IEEE, 2017.

[Gui *et al.*, 2019] Tao Gui, Liang Zhu, Qi Zhang, Minlong Peng, Xu Zhou, Keyu Ding, and Zhigang Chen. Cooperative multimodal approach to depression detection in twitter. In *Proc. AAAI*, 2019.

[He *et al.*, 2018] Tiantian He, Lun Hu, Keith CC Chan, and Pengwei Hu. Learning latent factors for community identification and summarization. *IEEE access*, 6:30137–30148, 2018.

[Hu *et al.*, 2017] Pengwei Hu, Tiantian He, Keith CC Chan, and Henry Leung. Deep fusion of multiple networks for learning latent social communities. In *2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 765–771. IEEE, 2017.

[Lin and Kumar, 2017] Chenhao Lin and Ajay Kumar. Multi-siamese networks to accurately match contactless to contact-based fingerprint images. In *2017 IEEE International Joint Conference on Biometrics (IJCB)*, pages 277–285. IEEE, 2017.

[Lin and Kumar, 2018] Chenhao Lin and Ajay Kumar. A cnn-based framework for comparison of contactless to contact-based fingerprints. *IEEE Transactions on Information Forensics and Security*, 14(3):662–676, 2018.

[Lin *et al.*, 2014] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014.

[Liu *et al.*, 2018] Zhun Liu, Ying Shen, Varun Bharadhwaj Lakshminarasimhan, Paul Pu Liang, AmirAli Bagher Zadeh, and Louis-Philippe Morency. Efficient low-rank multimodal fusion with modality-specific factors. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2247–2256, 2018.

[Manikonda and De Choudhury, 2017] Lydia Manikonda and Munmun De Choudhury. Modeling and understanding visual attributes of mental health disclosures in social media. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 170–181. ACM, 2017.

[Organization, 2010] World Health Organization. *World health statistics 2010*. World Health Organization, 2010.

[Picardi *et al.*, 2016] A Picardi, I Lega, L Tarsitani, M Caredda, G Matteucci, MP Zerella, R Miglio, A Gigantesco, M Cerbo, A Gaddini, et al. A randomised controlled trial of the effectiveness of a program for early detection and treatment of depression in primary care. *Journal of affective disorders*, 198:96–101, 2016.

[Reay *et al.*, 2011] Rebecca Reay, Stephen Matthey, David Ellwood, and Maureen Scott. Long-term outcomes of participants in a perinatal depression early detection program. *Journal of affective disorders*, 129(1-3):94–103, 2011.

[Shen *et al.*, 2017] 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 *IJCAI*, pages 3838–3844, 2017.

[Shen *et al.*, 2018] Tiancheng Shen, Jia Jia, Guangyao Shen, Fuli Feng, Xiangnan He, Huanbo Luan, Jie Tang, Thanassis Tiropanis, Tat-Seng Chua, and Wendy Hall. Cross-domain depression detection via harvesting social media. In *IJCAI*, pages 1611–1617, 2018.

[Stolar *et al.*, 2015] Melissa N Stolar, Margaret Lech, and Nicholas B Allen. Detection of depression in adolescents based on statistical modeling of emotional influences in parent-adolescent conversations. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 987–991. IEEE, 2015.