Using Social Media Content to Identify Mental Health Problems: The Case of #Depression in Sina Weibo

Review of Integrative Business & Economics Research

Danyang Lyu*
Wenzhou-Kean University

Zixuan Wang
Wenzhou-Kean University

Yutong Du
Wenzhou-Kean University

Rob Kim Marjerison
Wenzhou-Kean University

Rongjuan Chen
Wenzhou-Kean University

ABSTRACT

This paper aims to propose a logistic regression model that would predict depression by relying on content-related factors observed through social media. Specifically, we propose that by examining: 1) emotions embedded in Weibo posts, 2) user's tendency of self-disclosure, and 3) user characteristics, it could be determined whether the person is likely to be suffering from depression or not. Data was gathered from 48 individual users of Sina Weibo, a popular microblogging platform in mainland China, and was then subjected to quantitative analysis based on several factors. The sample consisted of 5,354 Weibo posts from the 48 individuals, 11 of whom were suffering from depression. We found that some, but not all, negative emotions (e.g., sadness, disgust) were positively related to depression. Furthermore, the two indicators of self-disclosure, number of followers and length of self-introduction in user profile, were negatively related to depression. In addition, user characteristics such as gender and location could determine the likelihood of depression. Finally, we addressed some theoretical and practical implications and suggested several directions for future study on related topics.

Keywords: Social Media, Content Analysis, Depression, Sina Weibo

1. INTRODUCTION

On a global scale, the number of people suffering from depression has been increasing steadily for several decades. With increasing social pressure, a wide range of mental health problems, such as depression, insomnia and suicide, have emerged. For example, there are various factors causing depression. According to a report released by the World

Health Organization (https://www.who.int/) in 2017, depression will be the second most common disease in the world by 2020, next to heart disease (WHO, 2017). In China, the prevalent rate of depression is 6.1% and the incidence of depression has increased year over year (WHO, 2015). The WHO report issued in 2017 has also addressed that, while there are more than 54 million patients with depression in China, the treatment rate of depression is less than 20% (WHO, 2017). Depression has become a common mental disease worldwide and particularly in China.

Importantly, there may be many undetected potentially depressed patients in the world. Less than a half of depression patients in the world, and in many countries less than 10%, receive effective treatment (WHO, 2017). Many potential patients have not been diagnosed, but they are expected to behave normally. In public, those striving to, or wishing to appear normal, may want to show their love of life and that they are sociable and energetic. When those living with depression disguise their mental problems, they tend to feel more depressed and exhausted. A research team sampled Facebook messages from 683 users, few of which have been diagnosed with depression. After a mental health examination, it was found that 114 of them have been identified to be depression sufferers (Eichstaedt, 2016). This somewhat surprising result suggests that it would be important for us to know how depression could be detected at an early stage for timely support and effective intervention.

The relationship between depression and social media has become an interesting topic. Many believe that social media could be a contributing factor in depression, arguing that social media has a negative impact on individuals' well-being. Not only depression, adolescents' sleep problems could be a result of the overuse of social media (Woods & Scott, 2016). Those authors claim that social media is linked to young people's poor self-esteem, which could potentially lead to poor mental health. Meanwhile, other researchers suggest that the use of social media does not lead to depression, but individual differences, such as personality, play a greater role in how social media impacts mental health (Heffer et al., 2019). Nevertheless, social media is considered by many as a useful tool for people to enhance social relationships and receive social support through information exchange (Ellison, 2013). Furthermore, the advantages of using social media could help individual users maintain friendships, express thoughts and emotional identities (Best, 2014). The use of social media has been linked to reduced loneliness, improved self-esteem and life satisfaction (Seabrook, Kern, & Rickard, 2016). For people with depression or anxiety, in particular, use of social media may increase the exposure of individual social interactions and affect emotional and mental health (Steger & Kashdan, 2009).

Although the impacts of social media on our people and society have been complicated and debatable, past researchers have widely recognized that social media content helps in identifying mental problems, such as depression. For example, Cavazos-Rehg and colleagues (2016) researched depression-related tweets and gained some insights of social networking about this mental problem on Twitter. Through content analysis, they found that several symptoms of depression were frequently addressed in a random sample of

2,000 tweets discussing depression. But, to our best understanding, these results are mostly descriptive and may not be able to explain whether certain content is more likely to indicate depression, as opposed to other content. Thus, it becomes important to explore if social media content can be further examined, in terms of its relation to depression. To address this research interest, the current study aims to examine if social media content can be used to predict depression. Specifically, we study individuals' depression-related posts collected from Sina Weibo, which is a Twitter-like microblogging website originated from mainland China.

Based on the literature, we will propose and test an empirical research model, in which the relationships between content-related factors and depression are addressed. To specify, we are hoping to gain a better understanding of depression, which is a typical mental problem, as being disclosed by oneself in social media. For a prediction of depression, several content-related factors will be examined, including: 1) emotions embedded in a Weibo post, 2) self-disclosure, and 3) user characteristics, all of which are likely to be related to depression. More details regarding past literature and theoretical development are discussed in the following section.

2. LITERATURE REVIEW

2.1 Depression

Depression is a common mental disorder problem that comes along with depressed mood, loss of interest or pleasure, decreased energy, feelings of guilt or low self-worth, disturbed sleep or appetite, and poor concentration (WHO, 2012). Hankin et al. (1998) claim that depression may affect the feeling and thinking of individuals. This mood disease is usually associated with serious clinical symptoms that can cause other more severe symptoms (Robinson et al., 2019). According to a recent study, depression has become a prevalent problem in life (Papageorgiou et al., 2015). Furthermore, depression, as a widely recognized mental problem, may add a high burden on an individual's well-being and life quality (Ustun et al., 2004). It is noteworthy that the severity of depression is not only about the incidence and mortality rate, but also depression affects one's physical health status (Kessler et al., 2010).

In order to better understand depression as such a common mental problem, we reviewed the literature addressing depression and noted that nowadays people would be more inclined to use social media to discuss depression. A past study examined the use of social media in relation to depression, and those researchers found that depression could be affected by a variety of factors, including living environment, people around, internet information, personal personality, and sleep quality (Woods & Scott, 2016). People would be more willing to declare their depression status in social media anonymously. Therefore, it becomes a promising approach that information is gathered from social media in order to extract content-related factors, which are potentially associated with self-disclosed depression status. For example, Andalibi and colleagues (2017) studied whether disclosing negative emotions would signal significant depressive symptoms on Instagram. Other researchers conducted content analysis to better understand people's thoughts and ideas about depression communicated through Twitter (Cavazos-Rehg et al., 2017). However, to our best knowledge little has been explored by far, about predicting depression by using social media content. The empirical evidence supporting a predictive model of depression is particularly insufficient. The current study will address this limitation in existing literature.

2.2 Emotions in Social Media Content

Emotion refers to an emotional state including thoughts, physiological changes, and an outward expression or behavior (Jalonen, 2014). While using social media, people tend to express emotions and share experiences publicly in an unprecedented manner. For social media researchers, classifying emotion is a useful method to analyze social media content and differentiate users by showing individual behavior (Schweitzer & Garcia, 2010). Past research claims that the emotion types of different users can be examined by social media content and user linkages, in order to better understand individuals' perception and behavior (Chung et al., 2015).

Particularly, negative emotions, such as anger and fear, depress memory and produce inefficient information processing (Baron, 1990). Based on Lazarus and Folkman's (1984), negative emotions depend on how an individual evaluates transactions with the environments and their effects on humans may be contagious. Considering that social media provides such a convenient way for people to share real-time information and immediate responses, negative emotions expressed online can be more widely spread. That is, people can catch on some extreme emotions, such as sadness, anger, anxiety, immediately and directly from others through social media. Social media researchers should understand individuals' emotions behind what is being posted online.

Furthermore, some negative emotions are believed to be more closely related to depression. Hankin and Abramson (2001) proposed that people's tendency to ruminate in response to negative emotions could place them at risk for depression. Also, past research claimed that negative emotions would lead to an inability to effectively communicate emotional experiences with others, resulting in a lack of social support for managing negative emotions, and a failure to develop skills for regulating negative emotions once they are experienced and expressed (Keenan et al., 2009). Since depression is characterized by difficulties in controlling negative emotions and thoughts, people with long-term negative emotions are more likely to suffer from depression (Disner et al., 2011; Keenan et al., 2009). Past studies also indicated that the tendency of an individual to experience depression could be detected through social media by extracting negative emotions, as he or she may share difficult feelings and signal a need for help to others (Andalibi, Ozturk, & Forte, 2017). In the current research, we explore if negative emotions expressed online could help identify depression.

RQ1: How do emotions embedded in Weibo posts relate to depression?

2.3 Self-Disclosure

Self-disclosure refers to the tendency of individuals to share their sensitive information, such as photos, personal experiences, thoughts, and feelings, with others (Jeske, Lippke, & Shultz, 2019). In other words, self-disclosure is the process of revealing personal information to someone else (Greene, Derlega, & Mathews, 2006). Self-disclosure is a conscious act, usually conveyed through verbal behaviors that describe a person's experiences and feelings (Choi & Bazarova, 2015). Such act has become a basic phenomenon of communication on the Internet, and also a major feature of computer-mediated communication (El Ouirdi, Segers, & Pais, 2015). A number of past studies have collected publicly accessible data containing self-disclosed information as well as self-reported mental health status, in order to research how related topics have been discussed in social media.

Interestingly, self-disclosure can indicate how likely one is suffering from depression as usually expressed by oneself. There are several studies that relate self-disclosure to depression in social media. A past study suggested that a person who was willing to express himself or herself would be more likely to seek support through social media, whereas those unwilling to reveal information were less likely to do so (Andalibi, Ozturk, & Forte, 2017). As a result, those who did not seek social support would become more likely to be in a difficult situation and experience depression. According to past research, user-disclosed content includes explicit identity statements, such as autobiographic descriptions, which are narrative claims in the form of users' verbal descriptions of themselves (El Ouirdi, Segers, & Pais, 2015; Zhao, Grasmuck, & Martin, 2008). Visual appeal in the form of photos or pictures uploaded by users themselves, or posted by others on their pages to show themselves as social actors is also a means of disclosure (El Ouirdi, Segers, & Pais, 2015). Furthermore, self-claims on Weibo are publicly shared among a complete network of friends or followers, composed of a large number of different audiences, ranging from strangers and far-away acquaintances to close friends and family members (Bazarova & Choi, 2014). In the current study, we research the relationship between self-disclosure and depression.

RQ2: How does self-disclosure of a person relate to depression?

2.4 User Characteristics

D'Ambra and Rice (2001) propose that username, gender, and background information could present the characteristics of individuals surfing on the Internet. Besides, Aapola (2002) states that age is a significant variable in contemporary social science research to distinguish individuals and explain differences among them. Furthermore, Sajilan and colleagues (2015) demonstrated evidence showing that males and females would have significant differences in personal independence and self-confidence, which may in turn affect their online behaviors. As social media provides the functions to connect people nearby but also those from faraway, geographical factors would probably build many different ways of dealing with depression (Papakostas, 2004). When people are communicating about depression, their individual differences would lead to the differences in their acceptance of things (Griffiths et al., 2019).

In social media, the information that reveals user characteristics is publicly accessible, for example in one's user profile, so that researchers can use this kind of information to indicate the likelihood of depression. Indeed, little has been done in past research, regarding the prediction of depression by directly referring to user information. Although user characteristics have long been studied in social media research (Gao et al., 2013; Juan & Jin, 2016; Zheleva & Getoor, 2009), the link between individuals' characteristics and mental health, such as depression, has been overlooked. In the current work, we want to address this interesting question.

RQ3: How do user characteristics relate to depression?

2.5 Social Media

Social media is a combination of media and social communication in the form of online tools, which can facilitate interaction and communication among a wide audience (As' ad et al., 2014). It creates a communication environment in which people discuss depression. In social media, users generate and share content with others, create their own personalized profiles, and develop online social networks (Obar & Wildman, 2015). Social media users may choose the information of their particular interests and share the information with others through social media platforms at any convenient time (Weeks & Holbert, 2013). The platforms of social media have a wide range of objectives and usage options, ranging from text-based platforms focused on information exchange (e.g., Twitter) and image-centric platforms focused on sharing videos and photos (e.g., Instagram, Snapchat), to platforms with multiple usage functions (Vannucci & McCauley, 2019). According to Jamison-Powell and colleagues (2012), social media services are being used as tools to disclose various personal health information. Depression-related information is one of this kind. Considering the rapid development and great popularity of social media, this research focuses on how depression is discussed through social media.

2.6 A Conceptual Model of Predicting Depression

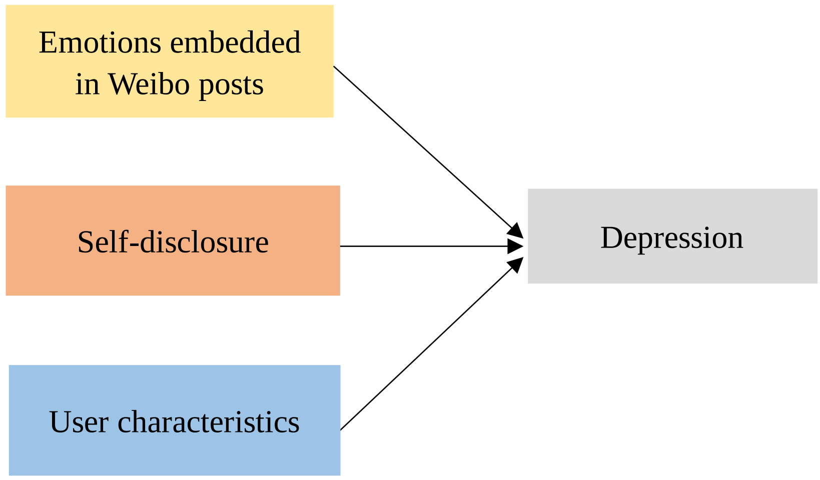

Figure 1. A conceptual model of depression indicated by social media content

3. METHOD

3.1 Data Collection

In order to predict an individual's likelihood of depression, we developed three independent variables, including negative emotions embedded in Weibo posts, self-disclosure of users, and user characteristics extracted from the user profile. A logistic regression model was utilized to examine how these factors could help identify Weibo users' depression status. Data was collected from Sina Weibo (www.weibo.com), one of the most popular social media platforms in mainland China. We randomly sampled 5,354 Weibo posts from 48 individuals, by following a thread named "Depression" on Weibo's website. Among these individuals, 11 of them were identified as depression or potential depression sufferers, as being self-reported. For each individual user, the latest 110 posts along with user profile information were collected through the web scraping package from Python x64 3.6. We used this full sample of data for further analysis.

3.2 Statistical Treatment

Our dependent variable is a person's depression status (0 = non-depression, 1 = depression). Since a logistic regression model is commonly used to predict a probability between 0 and 1, we consider it appropriate for this research. Specifically, a binomial model should be employed to investigate the discrete prediction of a Weibo user's likelihood of depression. Furthermore, in this study there are three sets of independent variables for a prediction of depression, including: 1) Weibo-embedded emotions, 2) self-disclosure, and 3) user characteristics. The details of setting up these independent

variables will be discussed in the next sub-section.
A regression model was used as a statistical tool to investigate the relationship between the dependent variable and independent variables (Gujarati, 2003). Logistic regression, as one type of the non-linear regressions, is usually adopted to model a two-dependent variable system (Stock & Watson, 2007). Furthermore, we linearized the non-linear regression model by applying statistical transformations (Bal & Gulse, 2013). When the dependent variable is a dichotomous choice, the logistic regression model should be treated as a "binary logistic regression model" (Gujarati, 2003). In the current work, we used the binary logistic model and the cumulative logistic probability function as follows:

$$P_i \;= F\left(\alpha + \sum^{m} \beta_j X_{ij} + \varepsilon\right) = \,{}^{1}\!\big/_{\left(1 + e^{-(\alpha + \sum^{m} \beta_j X_{ij} + \varepsilon)}\right)}$$

As shown above, P is the probability of a Weibo user i identified to be depressed or not; m is the total number of indicators; β_j is the parameter of indicators; j is the number of indicators; X_{ij} is the independent variable corresponding to each indicator and each observation; α is the intercept; ε is the error.

3.3 Variable Settings

As for the independent variables, Weibo-embedded emotions consist of several emotions extracted from the sample Weibo posts. Since the main themes of these posts were mostly negative, indicating depression and related mental health issues, the appearance of positive emotions was extremely rare. So, we decided to focus on negative emotions only. We then used a lexicon to identify negative emotional words and computed the frequencies of these emotions appearing in each Weibo post. In particular, we first tokenized a post by using "Jieba tokenizer" in Python. Next, removed the stop words, representing commonly used words, and obtained several emotions that were most frequently mentioned in the posts. Using a sentiment lexicon provided by Dalian University of Technology (Xu et al., 2008), we concentrated on four types of negative emotions, including: sadness, disgust, fear, and anger. Finally, based on the frequency of each emotion expressed in a post, we gave a score to each post, such as "SAD" for sadness and "DIS" for disgust. Thus, a post should have four scores, one for each of the four negative emotions as being extracted and classified.

As for self-disclosure, for each user we counted the total number of pictures from his or her sample posts. Another indicator of self-disclosure was the number of followers. The third indicator of self-disclosure was the length of a user's profile description, for which we counted the words in a person's self-introduction on his or her profile page. For user characteristics, we used the gender and location of each individual. Figure 2 presents the empirical model of predicting depression, and Table 1 details the operationalization of all the variables.

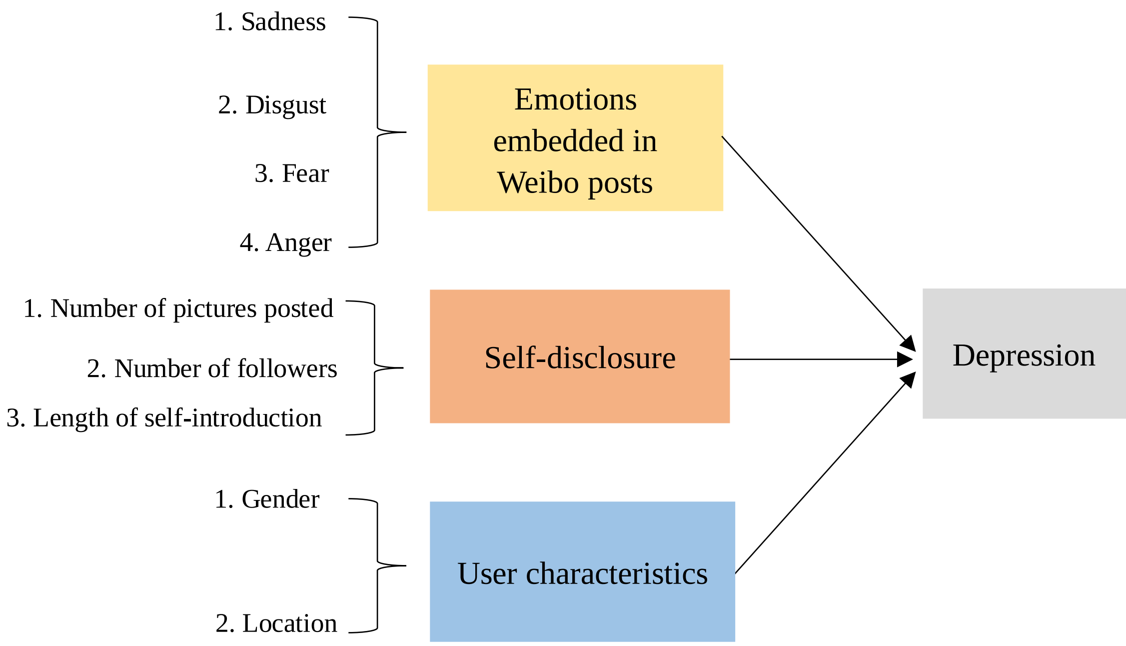

Figure 2. An empirical model of predicting depression using social media content

Variable	Operationalization	Data type
Independent variable		
Independent variables #1: Negative emotions		
Sadness (SAD)	Frequency of 'sadness' appearing in a post	Continuous
Disgust (DIS)	Frequency of 'disgust' appearing in a post	Continuous
Fear (FER)	Frequency of 'fear' appearing in a post	Continuous
Anger (ANG)	Frequency of 'anger' appearing in a post	Continuous
Independent variables #2: Self-disclosure		
Picture (PIC)	The number of pictures	Continuous
Followers (FLW)	The number of followers	Continuous
Description (DCP)	The length of profile description	Continuous
Independent variables #3: User characteristics		
Gender (GEN)	0: female, 1: male	Discrete
Location (LOC)	Location information (1: N/A, 2: Beijing, 3: Guangzhou, etc.)	Categorical
Dependent variable		
Depression	0: non-depression, 1: depression	Binary

Table 1. Variable operationalization

4. RESULTS

Our results suggest that negative emotions, self-disclosure, and user characteristics are promising indicators of depression. First, Weibo users being potentially a depression patient is more likely to address "sadness" and "disgust" in his or her posts, as opposed to non-depression users. For sadness, depression versus non-depression is corresponding to a mean score of 0.138 versus 0.070 ($p < .001$). Likewise, for disgust depression versus non-depression are associated with a mean score of 1.964 versus 0.820 ($p < .001$). However, not all negative emotions expressed by Weibo users could differentiate depression from non-depression. No significant difference in "fear" and "anger" is found ($p > .1$). So, the results indicate that the mention of "sadness" and "disgust" would signal one's depression status but the appearance of "fear" and "anger" would not work in the same manner. Table 2 shows the descriptive statistics of negative emotional words extracted from our sample Weibo posts.

Variable	Mean	Std.Dev.	Range (min ~ max)
Sadness	0.085	0.749	0 ~ 14
Disgust	1.073	2.981	0 ~ 43
Fear	0.037	0.493	0 ~ 12
Anger	0.016	0.311	0 ~ 9

Table 2. Negative emotional words

Figure 3 shown below highlights the words that have frequently appeared in sample posts, including: "sorry", "why", "pain", "survive", "unable", and "tortured".

Figure 3. Word cloud showing frequent words in sample posts

Next, our results show that on average a Weibo user has posted 80 pictures in his or her

110 total posts. Depression users versus non-depression users demonstrate a significant difference in the number of posted pictures, 31 versus 95 respectively ($p < .001$). This is indicating that, by solely relying on the number of pictures in a person's posts, we could possibly differentiate depression and non-depression. Since self-disclosure is multi-dimensional, we then look through the results regarding the other two indicators of self-disclosure. For the number of followers, potential depression sufferers have an average of 107 followers, as compared to those non-depression with an average of 432 followers ($p < .001$). The result suggests that an individual experiencing depression is less sociable and more likely to receive rejection from others, perhaps because of his or her negative attitudes toward things. For profile description, depressed individuals use fewer words than others in their introduction, 7.41 versus 12.72 ($p < .001$). The result shows that those suffering from depression are more hesitant to express themselves. Table 3 presents the descriptive statistics of self-disclosure of sampled Weibo users.

Variable	Mean	Std.Dev.	Range (min ~ max)
Number of pictures	80.46	90.41	0 ~ 379
Number of followers	359.89	293.23	0 ~ 1099
Length of self-introduction	11.55	14.79	0 ~ 73

Table 3. Self-disclosure

Finally, using a Chi-square test we found no significant difference in gender. Females and males have equal likelihood to suffer from depression, 23.1% and 22.2% respectively ($p > .1$). As for location, we further created a subsample involving five locations. Other locations originally existing in the full sample were found to have no depression sufferers, therefore they should be ruled out before we conducted a Chi-square test. In terms of the proportion of Weibo users potentially experiencing depression, the result shows significant differences across the following locations, Beijing, Guangzhou, Guizhou, oversea, and other, with X^2 (4, N = 3007) = 569.15, $p < .001$ (see Figure 4).

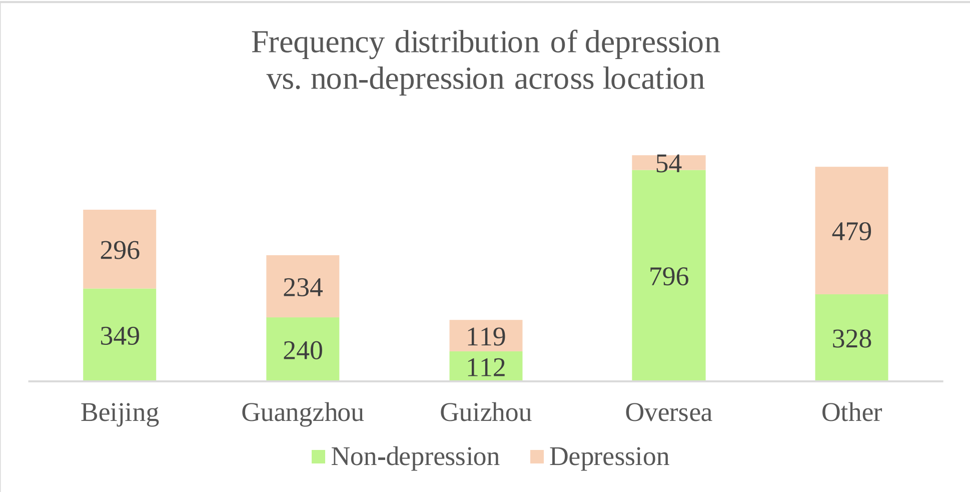

Figure 4. Depression vs. non-depression across the location

Based on the results of logistic regression, Table 4 below presents the statistic of the variance inflation factor (*VIF*) for each variable. Mathematically, *VIF* for the i^{th} independent variable is defined as follows:

$$VIF_i = \frac{1}{1 - R_i^2},$$

Here, R_i^2 denotes the coefficient of determination of the model that regresses the i^{th} independent variable against the remaining variables. Practically, it sheds light on the degrees of multicollinearity in a linear model. Multicollinearity could inflate the estimated variance of coefficients and thus make statistical inference spurious. Moreover, severe multicollinearity makes the estimation of coefficients biased and unstable in inverting a nearly singular matrix in modern statistical packages. In view of its implication, we then adopt a rule of thumb to consider $VIF < 5$ as an acceptable threshold for deciding whether to retain one variable in the model. According to the statistical results, the frequency of emotional words including sadness, disgust, fear, and anger in a post, number of pictures posted by Weibo users, number of followers, length of self-description, and users' gender and location information should have significant associations with depression. Furthermore, the model constructed enjoys a Mcfadden R^2 of .389, indicating a well-performed prediction.

SAD	**DIS**	**FER**	**ANG**	**PIC**	**LOC**	**GEN**	**FLW**	**DCP**
1.003	1.017	1.005	1.010	1.189	3.347	1.802	2.560	1.326

Table 4. Variance inflation factor (VIF)

Table 5 shows the indicators that are found to be significant indicators of depression. These indicators are: negative emotions including sadness and disgust, self-disclosure including the number of followers and length of self-introduction, gender, and location.

Variable	**Coefficients**	**Std.Error**	**z value**	**Pr(>|z|)**
SAD	.148	.046	3.253	< .01 **
DIS	.150	.012	12.155	< .001 ***
FER	.081	.068	1.189	.234
ANG	-.028	.137	-.202	.840
PIC	-.001	.001	-.785	.432
FLW	-.004	.000	-13.321	< .001 ***
DCP	-.021	.005	-4.024	< .001 ***
GEN	.310	.096	3.216	< .01 **
LOC	-.213	.014	-15.702	< .0001 ***
McFadden R^2	.3892			

Note: Significance level: *** < .001, ** < .01, * < .05.

Table 5. The results of logistic regression analysis

To sum up our findings, in Figure 5 we present our empirical test results.

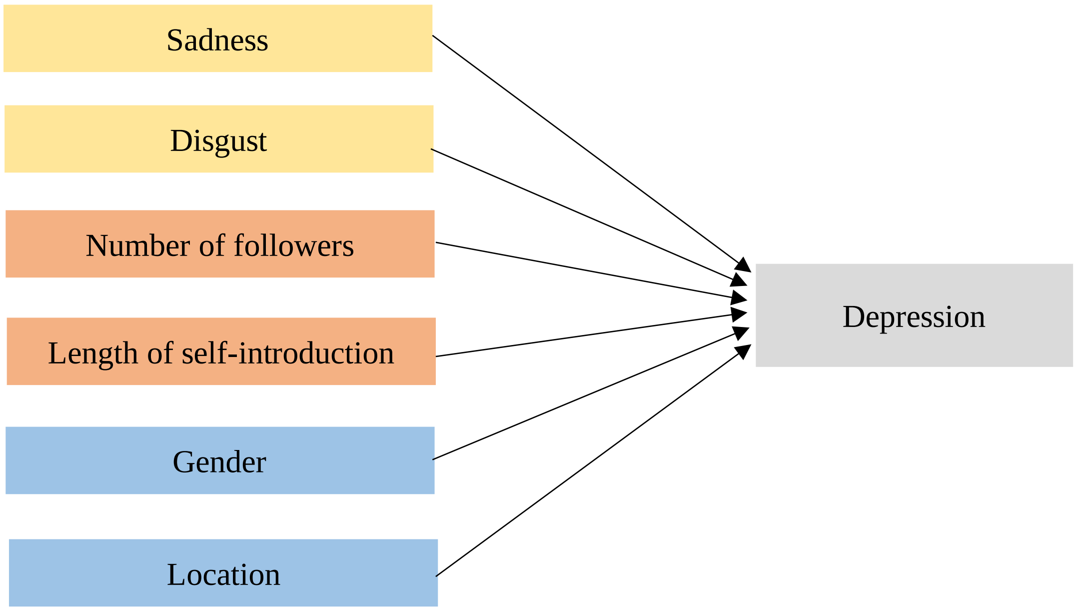

Figure 5. The empirical results of predicting depression using social media content

We summarize our results regarding the three research questions as follows.

RQ1: How do emotions embedded in Weibo posts relate to depression?

- Sadness and disgust are positively related to depression, but fear and anger are not significantly related to depression.

RQ2: How does self-disclosure of a person relate to depression?

- The number of followers and length of self-introduction are negatively related to depression. However, we have found mixed results regarding the number of pictures posted. This factor alone seems to be able to differentiate depression from non-depression. Meanwhile, according to the results of logistic regression, it is not a good indicator of depression likelihood.

RQ3: How do user characteristics relate to depression?

- Mixed results were obtained, regarding the relation of gender to depression. Gender as a single factor is not significantly related to depression, but it survives in a logistic regression model. In an additional test, we found that the contribution of gender to the overall prediction remains marginal. We found significant differences in the proportion of depression across location. For "other" location, meaning an unknown location, the depression rate is significantly higher than any other locations. "Guizhou" is a location with relatively high depression probability, while "Oversea" is along with a much lower chance of depression.

5. DISCUSSION

In this study, we focus on three content-related factors: 1) Weibo-embedded emotions, 2) self-disclosure, and 3) user characteristics, to indicate depression likelihood. According to the results, social media content can be used to effectively indicate a person's tendency to experience depression. Specifically, the results present several significant indicators of depression, including: sadness and disgust as expressed in Weibo posts, number of followers and length of self-description of Weibo users, gender and location.

The contributions of this research are twofold. First, empirical evidence is provided in this research, regarding the association between social media content and depression. For social media researchers, it remains an interesting topic to model individual behavior using social media data (Aldarwish & Ahmad, 2018; De Choudhury et al., 2013). Our study follows this line of research and shows that public social media data is a good source of information that enables researchers to conduct quantitative analysis. Second, in addition to user characteristics and self-presented profile information, in this research emotions are extracted for a prediction of depression. The results demonstrate a promising approach to exploring the underlying psychological status of a person who claims to be depressed. Past research has widely recognized the role of emotions in thinking processes, but whether an individual's certain emotions link to his or her depression symptoms has been underexplored (Cavazos-Rehg et al., 2016; Woods & Scott, 2016). This work is considered one of the early studies following this direction.

Since depression is a common mental problem, the results are meaningful to not only social media researchers but also those who are concerned about individuals' well-being. Based on the results, we would like to suggest that depression can be detected at an early stage so that timely support can be offered. Also, through social media people can feel comfortable to reveal their mental status as well as seek support from others. More importantly, for those who care about depression sufferers they can disseminate valuable information among potential depression patients, by recognizing differentiate social media content. For example, one can target a person feeling depressed by detecting his or her online posts mentioning some negative emotions such as "sad", "disgusted", and similar keywords. In addition, this study proposes that social media technologies should be leveraged to enhance individual and social well-being.

There are several limitations of this research to be addressed and the directions for future research to be offered. First, given the vast amount of information accessible in social media, a larger data sample can be obtained and used in a similar study. Other interesting factors can be explored in a predictive model of depression, such as hashtag and length of a post. Second, other social media platforms can be further studied by taking a similar approach, to generalize the current results. Finally, since depression in this research is self-reported, a future study may need to validate one's mental health status, by surveying social media users beforehand, for ground truth data collection.

ACKNOWLEDGEMENTS

This work is supported by Wenzhou-Kean University under Grant WKU201819011. The authors would like to thank the faculty and students involved in the project for their comments and encouragement.

REFERENCES

[1] Aapola, S. (2002). Exploring dimensions of age in young people's lives: A discourse analytical approach. *Time & society, 11*(2-3), 295-314.

[2] Aldarwish, M. M., & Ahmad, H. F. (2017, March). Predicting depression levels using social media posts. In *2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS)*, 277-280.

[3] Andalibi, N., Ozturk, P., & Forte, A. (2017, February). Sensitive self-disclosures, responses, and social support on Instagram: The case of #depression. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 1485-1500.

[4] As' ad, H. A. R., & Alhadid, A. Y. (2014). The impact of social media marketing on brand equity: An empirical study on mobile service providers in Jordan. Review of Integrative Business and Economics Research, 3(1), 334-345.

[5] Bal, H. S. G., & Gulse, S. (2013). Consumer characteristics influencing organic milk consumption preference in Tokat, Turkey. *Journal of Food, Agriculture & Environment*, *11*(1), 159-164.

[6] Bazarova, N. N., & Choi, Y. H. (2014). Self-disclosure in social media: Extending the functional approach to disclosure motivations and characteristics on social network sites. *Journal of Communication*, *64*(4), 635-657.

[7] Best, P., Manktelow, R., & Taylor, B. (2014). Online communication, social media and adolescent wellbeing: A systematic narrative review. *Children and Youth Services Review*, *41*, 27-36.

[8] Cavazos-Rehg, P. A., Krauss, M. J., Sowles, S., Connolly, S., Rosas, C., Bharadwaj, M., & Bierut, L. J. (2016). A content analysis of depression-related tweets. *Computers in Human Behavior*, *54*, 351-357.

[9] Choi, Y. H., & Bazarova, N. N. (2015). Self-disclosure characteristics and motivations in social media: Extending the functional model to multiple social network sites. *Human Communication Research*, *41*(4), 480-500.

[10] D'Ambra, J., & Rice, R. E. (2001). Emerging factors in user evaluation of the World Wide Web. *Information & Management, 38*(6), 373-384.

[11] De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013, June). Predicting depression via social media. In *Seventh International AAAI Conference on Weblogs and Social Media*.

[12] Eichstaedt, J. C., Smith, R. J., Merchant, R. M., Ungar, L. H., Crutchley, P., Preoţiuc-Pietro, D., Asch, D. A., & Schwartz, H. A. (2018). Facebook language predicts depression in medical records. *Proceedings of the National Academy of Sciences*, *115*(44), 11203-11208.

[13] Ellison, N. B., & Boyd, D. M. (2013). Sociality through social network sites. In *The*

Oxford Handbook of Internet Studies.

[14] El Ouirdi, M., Segers, J., El Ouirdi, A., & Pais, I. (2015). Predictors of job seekers' self-disclosure on social media. *Computers in Human Behavior, 53,* 1-12.

[15] Gao, R., Hao, B., Bai, S., Li, L., Li, A., & Zhu, T. (2013, October). Improving user profile with personality traits predicted from social media content. In *Proceedings of the 7th ACM Conference on Recommender Systems,* 355-358.

[16] Greene, K., Derlega, V. J., & Mathews, A. (2006). Self-disclosure in personal relationships. *The Cambridge Handbook of Personal Relationships*, 409-427.

[17] Griffiths, K. M., Carron-Arthur, B., Reynolds, J., Bennett, K., & Bennett, A. (2017). User characteristics and usage of an open access moderated internet support group for depression and other mental disorders: A prospective study. *Internet Interventions*, *7,* 9-15.

[18] Gujarati, D. N. (2009). *Basic Econometrics*. Tata McGraw-Hill Education.

[19] Hankin, B. L., Abramson, L. Y., Moffitt, T. E., Silva, P. A., McGee, R., & Angell, K. E. (1998). Development of depression from preadolescence to young adulthood: emerging gender differences in a 10-year longitudinal study. *Journal of Abnormal Psychology*, *107*(1), 128-140.

[20] Heffer, T., Good, M., Daly, O., MacDonell, E., & Willoughby, T. (2019). The Longitudinal Association Between Social-Media Use and Depressive Symptoms Among Adolescents and Young Adults: An Empirical Reply to Twenge et al. (2018). *Clinical Psychological Science*, *7*(3), 462-470.

[21] Jamison-Powell, S., Linehan, C., Daley, L., Garbett, A., & Lawson, S. (2012, May). I can't get no sleep: discussing #insomnia on twitter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1501-1510.

[22] Jeske, D., Lippke, S., & Shultz, K. S. (2019). Predicting Self-Disclosure in Recruitment in the Context of Social Media Screening. *Employee Responsibilities and Rights Journal*, *31*(2), 99-112.

[23] Juan, Y. F., & Jin, K. X. (2016). *U.S. Patent No. 9,444,907*. Washington, DC: U.S. Patent and Trademark Office.

[24] Kessler, R. C., Birnbaum, H. G., Shahly, V., Bromet, E., Hwang, I., McLaughlin, K. A., ... & Haro, J. M. (2010). Age differences in the prevalence and co-morbidity of DSM-IV major depressive episodes: results from the WHO World Mental Health Survey Initiative. *Depression and Anxiety*, *27*(4), 351-364.

[25] Obar, J. A., & Wildman, S. S. (2015). Social media definition and the governance challenge: An introduction to the special issue. *Telecommunications Policy, 39*(9), 745-750.

[26] Papageorgiou, K., Vermeulen, K. M., Schroevers, M. J., Stiggelbout, A. M., Buskens, E., Krabbe, P. F., ... & Ranchor, A. V. (2015). Do individuals with and without depression value depression differently? And if so, why?. *Quality of Life Research*, *24*(11), 2565-2575.

[27] Papakostas, G., Petersen, T., Mahal, Y., Mischoulon, D. Nierenberg, A., & Fava, M. (2004). Quality of life assessments in major depressive disorder: A review of the literature. *General Hospital Psychiatry, 26*(1), 13-17.

[28] Robinson, A., Bonnette, A., Howard, K., Ceballos, N., Dailey, S., Lu, Y. and Grimes, T. (2019). Social comparisons, social media addiction, and social interaction: An examination of specific social media behaviors related to major depressive disorder in a millennial population. *Journal of Applied Biobehavioral Research, 24*(1), e12158.

[29] Sajilan, S., Ul Hadi, N., & Tehseen, S. (2015). Impact of entrepreneur's demographic characteristics and personal characteristics on firm's performance under the mediating role of entrepreneur orientation. *Review of Integrative Business and Economics Research*, *4*(2), 36-52.

[30] Seabrook, E., Kern, M. and Rickard, N. (2016). Social Networking Sites, Depression, and Anxiety: A Systematic Review. *JMIR Mental Health, 3*(4), e50.

[31] Steger, M. F., & Kashdan, T. B. (2009). Depression and everyday social activity, belonging, and well-being. *Journal of Counseling Psychology*, *56*(2), 289-300.

[32] Stock, J. H. and Watson, M. W. (2007). Introduction to Econometrics. Pearson Addison Wesley, Boston.

[33] Ustun, T., Ayuso-Mateos, J., Chatterji, S., Mathers, C., & Murray, C. (2004). Global burden of depressive disorders in the year 2000. *British Journal of Psychiatry, 184*, 386-392.

[34] Vannucci, A., & Ohannessian, C. M. (2019). Social Media Use Subgroups Differentially Predict Psychosocial Well-Being During Early Adolescence. *Journal of Youth and Adolescence*, 1-25.

[35] Weeks, B. E., & Holbert, R. L. (2013). Predicting dissemination of news content in social media: A focus on reception, friending and partisanship. *Journalism & Mass Communication Quarterly, 90*, 212-232.

[36] WHO. (2017, March). WHO CHINA OFFICE FACT SHEET. Retrieved from http://www.wpro.who.int/china/topics/mental_health/1703mentalhealthfactsheet.pdf.

[37] WHO. (2015). Mental Health Atlas 2014. Retrieved from https://www.who.int/mental_health/evidence/atlas/mental_health_atlas_2014/en/.

[38] WHO. (2012). Depression – A Global Public Health Concern. Retrieved from https://www.who.int/mental_health/management/depression/who_paper_depression_wfmh_2012.pdf.

[39] Woods, H. C., & Scott, H. (2016). # Sleepyteens: Social media use in adolescence is associated with poor sleep quality, anxiety, depression and low self-esteem. *Journal of adolescence, 51*, 41-49.

[40] 徐琳宏, 林鸿飞, 潘宇, 任惠, & 陈建美. (2008). 情感词汇本体的构造. *情报学报*, *27*(2), 180-185.

[41] Zhao, S., Grasmuck, S., & Martin, J. (2008). Identity construction on Facebook: Digital empowerment in anchored relationships. *Computers in Human Behavior, 24*(5), 1816-1836.

[42] Zheleva, E., & Getoor, L. (2009, April). To join or not to join: the illusion of privacy in social networks with mixed public and private user profiles. In *Proceedings of the 18th International Conference on World Wide Web*, 531-540.