



Machine Learning Assessment of Factors Associated with Social Media Addiction Among Chinese College Students

Shuang Luo^(✉) and Mengke Zhu

School of International Education, Xiamen University of Technology, Xiamen, China
luoshuang@xmut.edu.cn

Abstract. Research has showed that the excessive use of social media has an impact on the user's functioning. Young students are among the most vulnerable to the impacts of social media addiction. This study aims to determine the status of social media addiction amongst Chinese college students and access the factors associated with such an addiction. An online survey was conducted among Chinese college students and the collected data were analyzed in IBM SPSS Statistics 29 descriptively. A machine learning algorithm, the random forest, was deployed to identify and Pearson correlation analysis was implemented to confirm important variables related to social media addiction. Results show that female respondents had higher scores than males and students in their third or fourth year of study had higher score than those in the first or second year of study. Participants majoring in Arts and Humanities also had higher addiction levels than those majoring in STEM disciplines. Fear of missing out and stress of online neglect, as well as hours spent on social media daily, had a clear correlation with social media addiction.

Keywords: social media addiction · college students · machine learning · random forest · Pearson correlation · predictive factors

1 Introduction

During the past century, technological advancements have fundamentally transformed the way we live [1] and significantly improved our lives by providing us with easy access to information, various forms of communication, increased productivity, cost-effectiveness, and time-saving benefits [2]. Social media has been playing a significant role. As of January 2022, there were a total of 4.62 billion people actively using social media around the globe [3], and 983.3 million in China, accounting for 68% of the Chinese population [4]. While social media addiction has not been officially recognized as a mental disorder, research has shown that excessive social media use can have negative impacts on health and physical functioning.

Adolescents and young people are particularly at risk for developing social media addiction due to their heavy usage of the internet and their vulnerability to the potential negative effects of social media. This age group tends to spend more time online than

other age groups and are therefore more likely to be exposed to the risks associated with social media [5]. Some of the negative effects of excessive social media use on young people may include sleep disruption, reduced attention span, negative self-perception, and cyberbullying [6, 7]. It is important not only to investigate the prevalence of addictive behaviors related to social media among young people, but also to develop effective treatment programs to address and reduce these rates. This study aims to determine the prevalence of social media addiction among college students and identify factors that may be associated with this addiction, such as peer pressure and the fear of missing out.

2 Methods and Materials

2.1 Data Collection

An online cross-sectional study was conducted from November to December 2022 in China among college students. The structured questionnaire used in this study included four different sections: (1) social characteristics; (2) the Bergen Social Media Addiction Scale (BSMAS); (3) stress of online neglect and negative reactions and fear of missing out (FOMO scale); and (4) status of social media platform usage.

In order to verify the reliability of the survey and evaluate the phrasing and arrangement of the questions, the questionnaire was initially administered to a small group of college students. After revising the questionnaire, it was uploaded to the online survey platform. Data collection commenced after the online survey was verified for accuracy and all technical issues were resolved. The convenience sampling method was used for data collection, and the Tencent Questionnaire Platform (wj.qq.com) was utilized to host the online survey questionnaire. The total number of participants who completed the questionnaire was 467, with a response rate of 88.5%.

2.2 Variables

Outcome Variable. The BSMAS scale was utilized to assess the severity of social media addiction among college students in this study. It consisted of 6 items, each of which was scored on a Likert scale with a range of options from 1 (very rarely) to 5 (very often). The psychometric properties of the BSMAS have been confirmed in different languages in various populations, including English [8], Italian [9], Persian [10], Portuguese [11], as well as Chinese [12]. In this study, the Cronbach's alpha for the BSMAS scale was 0.85.

Covariates. Four covariates were included in the questionnaire. Participants were asked about gender, age, year of study, major category, and living location. Then four questions were asked to assess participant's status of social media platform usage, specifically their favorite social media platforms, years of using social media, hours spent on social media daily, and purposes of using social media. Finally, we assessed the participant's stress of online neglect and negative reactions and fear of missing out. For the former, four Likert-scale questions with a range of options from 1 (disagree completely) to 5 (agree completely) were asked for the degree of stress the participant would feel when their online posts were not commented, not liked, or commented negatively, and when their

number of followers decreased or they were blocked by online friends. For the latter, the Chinese version of FOMO scale was used [13]. It consisted of 8 items with a range of options from 1 (disagree completely) to 5 (agree completely). The Cronbach's alphas of the two scales were 0.91 and 0.884 respectively.

2.3 Data Analysis

IBM SPSS Statistics 29 was used to analyze the data in this study. Continuous variables were presented as mean and standard deviation, while categorical variables were presented as frequencies with percentages. The Wilcoxon rank-sum test, Kruskal–Wallis tests for continuous variables, and χ^2 were used to compare differences between the status of social media addiction among college students and certain characteristics. A p-value (p) < 0.05 was thought to be significant statistically.

In terms of factor assessment, a random forest algorithm was employed, which is a type of machine learning technique that utilizes multiple models to make predictions. This method is nonparametric and aims to identify the optimal combination of variables to accurately predict new observations [14]. To analyze the data, the original dataset was randomly divided into two subsets: a training set, which comprised 80% of the total, and a testing set, which comprised the remaining 20%. Variables of importance were identified in the random forest model and confirmed with Pearson coefficients. Analyses were conducted in Python, where the random forest algorithm was implemented with the scikit-learn module.

3 Results

3.1 Descriptive Characteristics

Table 1 shows the demographic characteristics of total 467 participants, of whom the average age was 19.4. Most respondents were female (61.5%) and majoring in STEM disciplines (76%). The majority of the respondents were living in urban areas. Wechat Moments and QQ Zone were the most popular social media platforms among college students (95.9% and 73.4% respectively). Staying connected with friends and obtaining information were their major purposes of using social media (95.5% and 88% respectively). The spent an average of 5.91 h on social media per day.

Figure 1 shows six characteristics of social media addiction among college students. 27.6% of the participants had often/very often tried to cut down the use of social media but failed; 32.1% often/very often spent time thinking or planning to use it; 26.3% used social media to forget about personal problems; 27.2% became anxious or restless when they were denied access of social media; and 27.8% felt that overuse of social media had negative impact on their studies.

3.2 Status of Social Media Addiction Among Chinese College Students

Table 2 shows the average score of social media addiction among college students. Overall, female respondents had higher scores than males (17.45 and 16.67 respectively).

Table 1. Demographic characteristics of participants ($n = 467$).

| Characteristics | Male | | Female | | Total | | <i>p</i> |
|--|-------------|-----------|-------------|-----------|-------------|-----------|-----------------|
| | <i>n</i> | % | <i>n</i> | % | <i>n</i> | % | |
| Total | 178 | 38.1% | 289 | 61.9% | 467 | 100.0% | |
| Year of study | | | | | | | 0.022 |
| First Year | 72 | 40.4% | 114 | 39.4% | 186 | 39.8% | |
| Second Year | 38 | 21.3% | 96 | 33.2% | 134 | 28.7% | |
| Third Year | 45 | 25.3% | 54 | 18.7% | 99 | 21.2% | |
| Fourth Year | 23 | 12.9% | 25 | 8.7% | 48 | 10.3% | |
| Major category | | | | | | | 0.003 |
| STEM | 152 | 85.4% | 207 | 71.6% | 359 | 76.9% | |
| Humanities | 21 | 11.8% | 65 | 22.5% | 86 | 18.4% | |
| Arts | 5 | 2.8% | 17 | 5.9% | 22 | 4.7% | |
| Living location | | | | | | | 0.029 |
| Rural | 60 | 33.7% | 66 | 22.8% | 126 | 27.0% | |
| Suburban | 14 | 7.9% | 21 | 7.3% | 35 | 7.5% | |
| Urban | 104 | 58.4% | 202 | 69.9% | 306 | 65.5% | |
| Favorite social media platforms | | | | | | | |
| Wechat Moments | 165 | 35.3% | 283 | 60.6% | 448 | 95.9% | 0.005 |
| QQ Zone | 130 | 27.8% | 213 | 45.6% | 343 | 73.4% | 0.874 |
| Xiaohongshu | 67 | 14.3% | 224 | 48.0% | 291 | 62.3% | < 0.001 |
| Sina Weibo | 82 | 17.6% | 202 | 43.3% | 284 | 60.8% | < 0.001 |
| Douban | 13 | 2.8% | 27 | 5.8% | 40 | 8.6% | 0.444 |
| Douyin | 113 | 24.2% | 200 | 42.8% | 313 | 67.0% | 0.202 |
| Kuaishou | 15 | 3.2% | 29 | 6.2% | 44 | 9.4% | 0.564 |
| Others | 58 | 12.4% | 55 | 11.8% | 113 | 24.2% | < 0.001 |
| Purposes of using social media | | | | | | | |
| Stay connected with friends | 164 | 35.1% | 282 | 60.4% | 446 | 95.5% | 0.006 |
| Show lifestyle | 61 | 13.1% | 113 | 24.2% | 174 | 37.3% | 0.294 |
| Express emotions | 52 | 11.1% | 125 | 26.8% | 177 | 37.9% | 0.002 |
| Seek attention | 37 | 7.9% | 50 | 10.7% | 87 | 18.6% | 0.347 |
| Obtain information | 154 | 33.0% | 257 | 55.0% | 411 | 88.0% | 0.436 |
| Others | 38 | 8.1% | 67 | 14.3% | 105 | 22.5% | 0.645 |
| | Mean | SD | Mean | SD | Mean | SD | <i>p</i> |

(continued)

Table 1. (continued)

| Characteristics | Male | | Female | | Total | | p |
|---|------|-------|--------|-------|-------|--------|---------|
| | n | % | n | % | n | % | |
| Total | 178 | 38.1% | 289 | 61.9% | 467 | 100.0% | |
| Age | 19.8 | 1.9 | 19.2 | 1.2 | 19.4 | 1.5 | < 0.001 |
| Years of using social media | 10.3 | 4.0 | 9.4 | 3.6 | 9.8 | 3.8 | 0.009 |
| Hours spent on social media daily | 5.44 | 3.943 | 6.2 | 3.19 | 5.91 | 3.5 | < 0.001 |
| Fear of missing out (range: 8–40) | 20.5 | 7.8 | 21.1 | 7.1 | 20.9 | 7.4 | 0.313 |
| Stress of online neglect and negative reactions (range 4–20) | 10.2 | 4.5 | 10.8 | 4.5 | 10.6 | 4.5 | 0.126 |

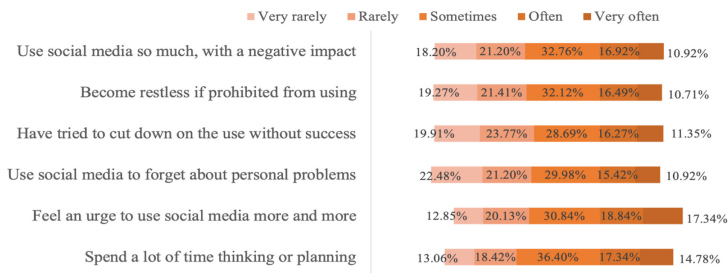


Fig. 1. Social media addiction characteristics.

Students in their third or fourth year of study had higher score than those in the first or second year of study. Participants majoring in Arts and Humanities also had higher addiction levels than those majoring in STEM disciplines. Moreover, students from urban and suburban areas, as well as those using more than one social media platforms, exhibited a slightly higher degree of social media addiction.

3.3 Assessment of Factors Associated with Social Media Addiction

Variable importances were provided by a built-in attribute of the scikit-learn module and they were computed as the mean and standard deviation of accumulation of the impurity decrease within each tree. Figure 2 shows the important variables that strongly impacted the performance of the random forest model for predicting social media addiction. They are sorted in a descending order by degree of importance. As is demonstrated, fear of missing out, stress of online neglect, and hours spent on social media daily were three most important correlated features of the model, while gender, living location, and year of study were the least important variables to the model.

The Pearson correlation analysis was conducted with Python to confirm the relationship between individual variables with the degree of social addiction. As we can see

Table 2. Status of social media addiction among Chinese college students ($n = 467$).

| | Social Media Addiction (Range: 6–30) | | |
|---------------------------------|--------------------------------------|------|----------|
| | Mean | SD | <i>p</i> |
| Total | 17.16 | 5.66 | |
| Gender | | | 0.192 |
| Male | 16.67 | 6.27 | |
| Female | 17.45 | 5.23 | |
| Year of study | | | 0.003 |
| First Year | 16.04 | 5.45 | |
| Second Year | 17.62 | 5.41 | |
| Third Year | 18.16 | 6.03 | |
| Fourth Year | 18.10 | 5.76 | |
| Major category | | | 0.147 |
| STEM | 16.89 | 5.63 | |
| Humanities | 17.85 | 5.66 | |
| Arts | 18.82 | 5.82 | |
| Living location | | | 0.033 |
| Rural | 16.80 | 5.25 | |
| Suburban | 19.77 | 6.59 | |
| Urban | 17.00 | 5.65 | |
| Number of platforms used | | | 0.305 |
| Only one | 16.36 | 6.78 | |
| More than one | 17.22 | 5.57 | |

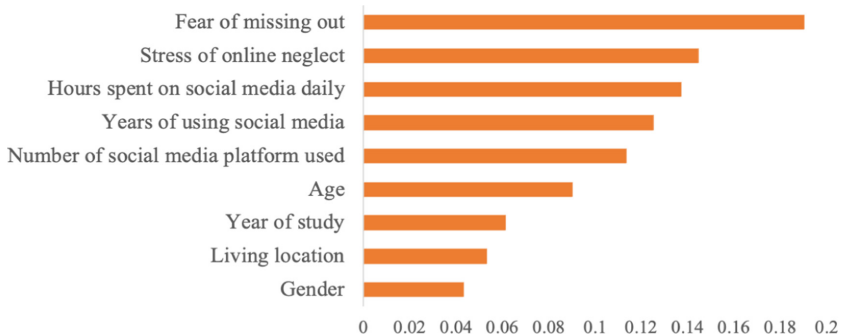
**Fig. 2.** Variable importance measure with the random forest model

Table 3. Correlation of individual characteristics with social media addiction

| Characteristics | Social Media Addiction |
|--|------------------------|
| | Pearson Coefficients |
| Gender | 0.067 |
| Age | 0.020 |
| Year of study | 0.149 |
| Major category | 0.093 |
| Living location | -0.001 |
| Number of social media platforms used | 0.140 |
| Years of using social media | 0.064 |
| Hours spent on social media daily | 0.286 |
| Fear of missing out (FOMO) | 0.601 |
| Stress of online neglect and negative reactions | 0.552 |

from Table 3, fear of missing out and stress of online neglect and negative reactions were clearly correlated with the score of the BSMAS scale (coefficients were 0.601 and 0.552 respectively). The number of hours spent on social media daily was to a lesser degree correlated with social media addiction too.

4 Discussion and Conclusions

This study characterized investigated the prevalence of social media usage among Chinese college students. We found that female students, students of third and fourth year, students living in urban and suburban, and students majoring in Arts and Humanities tended to have higher scores of social media addiction. Another finding was that the perceived stress of online neglect and negative reactions and the fear of missing out were associated with the overall level of social media addiction, which is consistent with findings of previous research works about perceived stress and problematic social networking use [15]. This study is significant because of the diversity of its sample. We were able to recruit different participants among Chinese students during the COVID-19 pandemic and investigate their level of social media addiction. The online survey platform was conducive in that it enabled us to reach various participants from different regions and with different demographic characteristics, thus increasing the representativeness of the results.

However, there remain several limitations. First, our sample size was not big enough, considering the extremely large number of college students in China. Only a small cross-section of students was involved, which may impact the degree of external validity. Second, the results of this study might be biased, because most questions in the survey were about personal preferences and feelings. Lastly, we did not take into account the existence of underlying distress, which might impact the relationship between perceived stress and social media addiction. With the prevalence of network devices and social

network services among college students, social media addiction and its consequences are increasingly evident and undeniable. This study contends that there is a need for higher institutions to deal with excessive use as well as for researchers to conduct further research on social media addiction and its association among college students.

References

1. Dragan, I. F., Dalessandri, D., Johnson, L. A., Tucker, A., & Walmsley, A. D. (2018). Impact of scientific and technological advances. *European Journal of Dental Education*, 22(S1), 17–20. <https://doi.org/10.1111/eje.12342>
2. The Benefits of Social Technology Use Among Older Adults Are Mediated by Reduced Loneliness. (2016). *Cyberpsychology, Behavior, and Social Networking*, 19(9), 551-556. <https://doi.org/https://doi.org/10.1089/cyber.2016.0151>
3. DataReportal (2022), “Digital 2022 Global Digital Overview,” retrieved from <https://datareportal.com/reports/digital-2022-global-overview-report>
4. DataReportal (2022), “Digital 2022 China,” retrieved from <https://datareportal.com/reports/digital-2022-china>
5. Pantic, I. Online social networking and mental health. *Cyberpsychol. Behav. Soc. Netw.* 2014, 17, 652–657.
6. Schmeichel, M.; Hughes, H.E.; Kutner, M. Qualitative research on youths’ social media use: A review of the literature. *Middle Grades Rev.* 2018, 4, 4.
7. Swist, T.; Collin, P.; McCormack, J.; Third, A. *Social Media and the Wellbeing of Children and Young People: A Literature Review*; Western Sydney University: Penrith, Australia, 2015.
8. Andreassen CS, Billieux J, Griffiths MD, Kuss DJ, Demetrovics Z, Mazzoni E, et al. The relationship between addictive use of social media and video games and symptoms of psychiatric disorders: a large-scale cross-sectional study. *Psychol Addict Behav.* 2016;30(2):252–62. <https://doi.org/https://doi.org/10.1037/adb0000160>.
9. Monacis L, De Palo V, Griffiths MD, Sinatra M. Social networking addiction, attachment style, and validation of the Italian version of the Bergen social media addiction scale. *J Behav Addict.* 2017; 6(2): 178–86.
10. Lin, C.Y.; Broström, A.; Nilsen, P.; Griffiths, M.D.; Pakpour, A.H. Psychometric validation of the Persian Bergen social media addiction scale using classic test theory and Rasch models. *J. Behav. Addict.* 2017, 6, 620–629.
11. Pontes H.M., Andreassen C.S., Griffiths M.D. Portuguese validation of the Bergen Facebook addiction scale: an empirical study. *Int J Ment Heal Addict.* 2016;14:1062–73.
12. Liu, C., & Ma, J. (2018). Development and validation of the Chinese social media addiction scale. *Personality and Individual Differences*, 134, 55-59. <https://doi.org/https://doi.org/10.1016/j.paid.2018.05.046>
13. Li, Y.-Y., Huang, Y.-T., & Dou, K. (2021). Validation and Psychometric Properties of the Chinese Version of the Fear of Missing Out Scale. *International Journal of Environmental Research and Public Health*, 18(18), Article 9896. <https://doi.org/10.3390/ijerph18189896>
14. Caruana R., Niculescu-Mizil A. An empirical comparison of supervised learning algorithms. Paper presented at: 23rd International Conference on Machine Learning; June 25–29, 2006; Pittsburgh, P.A. doi:<https://doi.org/10.1145/1143844.1143865>.
15. Hou, X.L.; Wang, H.Z.; Hu, T.Q.; Gentile, D.A.; Gaskin, J.; Wang, J.L. The relationship between perceived stress and problematic social networking site use among Chinese college students. *J. Behav. Addict.* 2019, 8, 306–317.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

