

A Sentiment Analysis of User's Comments on Higher Education Topics on Social Media

Zhuoyi Zhou^{1,*}

¹Haidian Foreign Language Shiyan School, Beijing, 100195, China

*Corresponding author: underhie@nwfsc.edu

Abstract:

In a digital era where social media platforms have grown to be significant forums for public discourse, it is crucial to comprehend and capitalize on the tone of remarks made about the various fields including education. This paper utilizes the sentiment analysis technique to delve deeper into the diverse opinions and feelings that people have expressed regarding higher education and its subtopics such as graduate employment, educational policies, institution qualities, and the general landscape, aiming to provide insightful recommendations to help educational institutions enhance their services. Based on the comments under 20 trending higher education-related tweets scraped and analyzed by Python on Sina Weibo, the study indicates a predominance of positive sentiments among the overall opinions toward higher education, alongside negative and neutral sentiments. Potential explanations for the varying sentiments towards the subtopics and the role of social media in shaping public opinions are also identified through detailed observations. This study demonstrates, for the first time, how sentiment analysis can be applied to higher education research, shedding light on the complex interplay between communication and education.

Keywords: Higher Education; Social Media; Sentiment Analysis; Weibo

1. Introduction

The accessibility and diversity of the internet, specifically social media, have contributed to the occurrence of internet public opinion that is featured by the loads of attitudes and sentiments expressed by individuals. According to Marr B., there are 293,000 updated status updates and 510,000 posted comments

every minute on Twitter [1]. Nonetheless, these social media platforms challenge the conventional media and political parties' control over public opinion, driving a more divided and challenged public sphere [2]. Internet public opinion gives a stage for genuine public discourse and government accountability while being inclined to emotional volatility and manipulation [3]. These findings collectively illustrate

an urgent need to reveal and respond to the significant trends and patterns, and user sentiments that lie amid the public opinions, especially in the context of education, which is an indicator of the whole society's well-being.

According to Cui Huang, there are five main topics that prevail over time on social media platforms: interactive learning environment and teaching/learning strategies, human capital and educational finance, teacher education, higher education, and equity and social justice [4]. Among the 5 themes, higher education, any post-secondary study that can result in a degree, has undergone a significant shift in its landscape due to rising enrollments, a more diverse student body, the emergence of public community colleges, and the advent of for-profit universities over the past fifty years, gaining widespread concerns and discussion [5]. Globally, there were over 235 million students enrolled in higher education in 2020—more than twice as many as there were in 2000 [6]. Public spending on higher education rose by 64% in a group of 55 nations after inflation between 2006 and 2018; in other words, spending per student increased by 7% as a result [7]. However, there is a decrease in average labor productivity (-2.5%) in China due to resource misallocation after higher education expansion, which is aligned with several other negative effects of the popularization of post-secondary education in different regions [8].

From the above mentioned, the practical significance of investigating the public's sentiments and emotions toward higher education is embodied in the necessity of improvements in services, curriculum, and campus environment. Moreover, it can serve as an early warning system for potential issues like access to higher education, quality of teaching, and affordability, enabling prompt responses and interventions. Academically, by deepening our understanding of the social context in which higher education functions, this study can make contributions to the field of education and communication.

2. Literature Review

Higher education is an engine that propels individuals towards intellectual horizons and societal contributions. Even so, there's an ongoing debate around the field. Specifically, some researchers have underscored its role in eroding the value of a college degree, while others advocate for a sustained investment in higher education, mentioning its merit in economic and social development [5, 9]. More recently, higher education institutions are experiencing transformation toward digitalization and facing manifold barriers [10]. Amongst the public, the perception of higher education as a private good is growing, which could have negative effects on both society and the indi-

vidual [11]. Thus, knowing people's opinions towards it is critical to educational institutions and the government's decision-making, particularly in the contemporary era.

Social media characterized by its high universality and interactivity plays an indispensable role in today's education, offering opportunities for student learning, teacher professional development, and educational research experiences [12]. Taking Facebook as an example, it's widely utilized in higher education settings to facilitate more convenient learning, boost engagement, and increase teacher-student connection [13]. On top of that, it also serves as a repository of public opinions, where netizens exchange thoughts and feelings related to public affairs, including educational events and policies.

Sentiment analysis, a computational study of people's opinions, demeanor, and feelings toward an entity, has been widely used in mining texts on social media platforms to gauge user's opinions relevant to education [14]. Previous papers have illustrated that sentiment analysis (SA) can efficiently capture people's opinions in the context of educational institutions and policies, permitting real-time checking of public emotions and attitudes [15]. These sentiments can be classified into three major categories: positive, neutral, and negative [16]. For instance, Wang and Fikis concluded that Twitter users express overwhelmingly negative sentiment towards the Common Core State Standards by applying sentiment analysis; Similarly, Gong demonstrated that the main emotion characteristic of the discussions on education justice and fairness on Sina Weibo is positivity, under the assistance of the same text mining tool [17, 18]. Besides, another research has utilized SA to indicate a mix of positive and negative sentiments towards online education during COVID-19 [19]. Preliminary literature based on the current application of SA in higher education has suggested that teaching quality evaluation is the most prevalent domain, proving the significance of SA in improving educational outcomes [20].

One of the research gaps currently is the lack of application of sentiment analysis on various specific educational topics, higher education in particular. How public sentiments evolved and what factors drive the formation of emotions still remained uncertain. Furthermore, limited consideration of gender and regional differences in sentiment in preliminary papers hinders the field of education to fully recognizing the issues in terms of higher education. Therefore, this study delves deeper into the distribution and pattern of sentiments of social media users towards post-secondary education topics, aiming to provide useful information for research in the fields of communication, sociology, and psychology. It sheds light on how people express and exchange their opinions about

a significant social institution.

3. Methodology

This study is an exploratory research. To investigate the distribution and pattern of users' sentiment towards higher education-related topics on Weibo, we used the Weibo Application Program Interface (API) to collect tweets with hashtags associated with higher education. Data is cleaned after collection to remove extraneous information, text segmentation, word segmentation, etc. Then, sentiment analysis was applied to examine the emotional tendency (positive, negative, and neutral) of each tweet by using a Python library. Last, demographic information and the content of each particular trending topic were combined with the corresponding sentiment score to analyze the potential factors exerting influence on the distribution and the pattern.

3.1 Data Collection

Among the mainstream Chinese social media, Weibo, a micro-blogging platform, has gained more than 600 monthly active user marks in the autumn of 2023 with its

potential to facilitate issue networks between various organizations and society [8, 9]. Furthermore, rich information about education fairness such as user types, user profiles, date and time of publishing, and amount of likes and followers are all accessible on this microblogging platform, all of which are useful for a methodical examination of the public opinion on education equity in China's networks. Python programming language was employed with appropriate libraries to scrape and extract comments related to higher education on Weibo. By searching keywords such as "higher education", "National College Entrance Examination", "undergraduate", "studying abroad", and "PhD", etc, 20 relevant prevalent topics, covering aspects like college enrollment and admission university management and construction, and educational national policy, between June 1st, 2024 and September 1st, 2024 were identified and selected as shown in Table 1. The scraping process was configured to collect the text of comments under the source tweets of these 20 topics along with relevant metadata like the gender and IP address of the commented user, and any other available information. 3,367 raw data comments in total were collected.

Table 1. Higher education topics on Weibo from June to September 2024

Serial number	Topic content	Popularity	Source	Time
1	The average monthly income of undergraduates in 2023 is 6,050 yuan.	1,644,813	China Business Network	6/13
2	Undergraduate-postgraduate continuous public-funded education in six normal universities directly under the Ministry of Education.	209,392	CCTV News	6/14
3	The Ministry of Education launches the college entrance examination volunteer information system for the first time.	1,394,455	CCTV News	6/18
4	The salary of 54 clinical medical students.	254,002	Toutiao News	6/22
5	Artificial intelligence may become the most popular major.	1,064,739	Shell Finance	6/28
6	The delay graduation rate of doctoral students in China remains high.	326,368	Sanlian Life Weekly	7/9
7	Sichuan University publicly announces the proposed revocation of 31 majors.	2,659,157	The Paper	7/12
8	Is studying abroad still a way out?	608,165	People	7/15
9	The population receiving higher education in China has reached 250 million.	2,578,754	CCTV News	7/19
10	China has built the world's largest education system.	2,978,784	People's Daily	7/19
11	Studying abroad costs 300,000 yuan, but the starting salary after returning home is only 5,500 yuan.	276,720	China News Service Finance	7/22
12	Many colleges and universities have decided to extend the graduate student schooling system.	1,849,850	Sina Finance	7/25

13	The business of guaranteed admission for studying in Hong Kong for graduate programs behind academic credential fraud.	592,498	Sanlian Life Weekly	7/25
14	Why are more and more college students going to vocational schools after graduation?	1,333,801	Workers' Daily	7/29
15	More than 20 colleges and universities have planned to suspend or withdraw more than 100 majors.	612,755	China Newsweek	8/3
16	Why do high-scoring students collectively give up studying abroad?	402,545	China Business Journal	8/3
17	In some colleges and universities, the policy of 'no promotion, then leave' has been perverted into 'managing by forcing people to leave.'	2,140,337	China Comment	8/6
18	How long can 'diploma mill' degrees still swagger around?	496,069	Workers' Daily	8/21
19	With the diminishing benefits of studying abroad, is it still worth going to the United States for an undergraduate degree?	416,460	Sanlian Life Weekly	8/23
20	Why is it getting more and more difficult for doctoral students to enter colleges and universities?	499,980	Southern Weekly	9/1

3.2 Data Analysis

3,356 comments remained after the data cleaning process, and sentiment analysis was performed on each comment using the snowNLP (Snow Natural Language Processing) library in Python. SnowNLP is a simple and efficient library for Chinese text sentiment analysis. For each comment, the library assigns a sentiment score ranging from 0 to 1, where [0-0.4) represents negative sentiment, [0.4-0.6) represents neutral sentiment, and [0.6-1) represents positive sentiment. To determine the proportion of negative, neutral, and positive sentiments, the number of comments in each category was divided by the total number of comments. This provides a quantitative measure of the overall sentiment distribution among the comments under higher education tweets on Weibo.

4. Results

This study conducted sentiment analysis on 3,356 users' comments about higher education topics on social media. The analysis revealed several key findings.

First, when examining the overall sentiment distribution, it was found that 60.86% of comments were positive, 23.12% were negative, and 16.03% were neutral. This indicates a generally positive sentiment among social media users regarding higher education. Second, the 20 topics in Table 1 were categorized into four themes based on contents: a) income and employment, b) educational policies and systems, c) choice, quality, and credibility in higher education, and d) the general education landscape. Analyzing themes revealed nuanced patterns. For instance,

comments relevant to income and employment (topics 1, 4, and 11) tended to have a relatively lower proportion of positive sentiments compared to the overall proportion. To be more specific, 50.47% of comments were positive, 19.98% neutral, and 29.55% negative. On the other hand, comments on educational policies and systems (topics 2, 3, 7, 12, 15, and 17) showed 57.67% of comments were positive, 18.47% were neutral, and 23.86% were negative, reflecting the public's moderately positive view of educational policies and systems and the constant existence of concern, such as the implementation of certain policies and major revocations. In the case of choice, quality, and credibility of higher education (topics 5, 6, 8, 13, 14, 16, 18, 19, 20), the sentiment scores indicated 62.18% positive comments, 15.42% neutral comments, and 22.40% negative comments. Similarly, topics associated with the general education landscape in China (topics 9 and 10) result in 64.45% positive sentiments, 15.42% neutral sentiments, and 20.13% negative sentiments. Last, it is found that there's no noticeable pattern in the geological and gender distribution of the commented users under higher education topics compared to the statistics of the ratio of male to female Weibo users and the regional proportion.

5. Discussion

This paper might be the first one to apply sentiment analysis on public opinion towards higher education. It clearly shows that the overall social media discourse of China's higher education is dominated by positive sentiment, demonstrating a general recognition of its importance in personal development and societal progress. In the

meantime, the coexistence of positive, neutral, and negative emotions is the embodiment of the two-sidedness of higher education discourse and higher education itself. However, a more detailed examination of sentiment distribution across minor themes reveals a complex landscape of public opinion. Although their patterns follow a general trend, each theme elicits varying levels of positive, negative, and neutral sentiments, influenced by the context and specific issues being discussed.

5.1 Sentiment Analysis Under Each Theme

In terms of the minor theme “income and employment”, the relatively decreased proportion of positive sentiment and the increased proportion of negative sentiments compared to that of the overall analysis might be attributed to the growing concerns about job market conditions and the economic value of post-secondary education solely. Topics under this theme are typically featured by authentic statistical data associated with fresh graduates’ unemployment condition, practically reflecting the current labor market confronted by them. Several forces have been identified as shaping such circumstances in China, including the mismatch of university programs and market demand, the inflow of well-educated migrants, and more importantly, the rapid expansion of higher education [21]. The popularization of higher education increases the burden and stress upon the fresh graduate, leading to a soaring number of candidates for a single job while being paid the wage of a manual laborer. Thus, the negative sentiment of the public that emerged in this context cannot be solely ascribed to the stagnant job market itself but also to the root cause—the expansion of higher education. Undoubtedly, it’s still noteworthy that sentiments can be polarized, with users expressing both frustration regarding underemployment or student debt burdens and optimism about potential employment opportunities. Research indicates that social media users often exhibit negative feelings about job insecurity, especially after economic downturns. Conversely, pleasant feelings can emerge from conversations about well-placed jobs or high graduation earnings [22].

When examining “educational policies and systems” and “choice, quality, and credibility in higher education” related topics, sentiment patterns tend to be consistent with the overall trend. Educational policies are highly political and emotionally charged, reflecting broader societal debates. Negative sentiment under this theme usually results from frustration with the perceived ineffectiveness of policy improvements or the slow pace of change; while when policies—like expanding accessibility or affordability—align with user expectations or ideals, positive sentiment arises. With regard to choice, quality, and credibility in

higher education, the sentiment dynamics may be strongly influenced by personal experiences and broader societal narratives about the value of higher education [22].

Lastly, the general education landscape online discourse tends to generate a larger proportion of positive emotions, embodying a majority of hope or excitement over success and innovations in education. The two topics entailed in this theme both mentioned the success of the contemporary Chinese higher education system, consolidating Weibo users’ recognition of the scale and reach of this educational infrastructure, which can empower individuals and drive social mobility. A possible explanation of the negative emotions presented here would be the public’s skepticism towards the quality of education in a large-scale system, considering the disparities in educational resources. For instance, while the overall numbers are impressive, there might be variations in the standard of education between various socioeconomic classes or between urban and rural locations. Negative attitudes may also stem from problems in the educational system, such as intense competition and pressure.

5.2 Social Media’s Role in Public Opinion towards Higher Education.

Social media has exerted an influential force in shaping the public’s opinion toward a variety of issues including higher education. On one hand, by focusing on the content of the 20 trending topics within 3 months, it is conspicuous that social media platforms offer a space for the extensive dissemination of higher education-related information. Policy changes in educational institutions rapidly reach out to widespread users through the internet, allowing them to make informed decisions on major choices and college applications. Moreover, discussions on social media about educational issues such as tuition fees, student debt, and the value of different degrees can raise public awareness and spark debate, potentially impacting policy decisions. In other words, social media like Weibo act as a conduit of information and a facilitator of the interaction between individuals and educational institutions. On the other hand, social media can also exacerbate the dispersal of misinformation and misconception. Misunderstanding and subsequent misleading comments can emerge even under tweets from credible sources, not to mention unreliable ones. Once rumors and confusion exist, the effect of the echo chambers and opinion polarization will appear [23], making it difficult to have a rational and balanced conversation about higher education.

5.3 Limitations and Future Suggestions

There are a few limitations to this study that are worth

cautioning. First, as exploratory research, the result based on a relatively small sample of Weibo comments might lack the ability to be generalized to the larger population. Secondly, there's a dearth of a comprehensive comparison between distinct social media platforms regarding the sentiment of users' comments on higher education topics. Although Weibo is the focus of this study, it is unknown how user behavior and sentiment distribution may vary on other well-known social media sites like Douyin, WeChat, or Twitter (if considered from an international perspective). Additionally, a longitudinal analysis could provide deeper insights into how public sentiment evolves instead of a short-term one focusing on merely 3 months. Future studies should delve deeper into the online sentiment pattern and distribution of higher education discussions by utilizing various social media platforms with a longer time span. Cultural factors, regional differences, and the influence of specific social events should also be taken into consideration when making conclusions.

6. Conclusion

The public opinion on social media relevant to higher education is dominated by positive sentiments, and the presence of neutral and negative sentiments proves the existence of limitations and problems in this field. The dual character of public sentiments and higher education itself is further entrenched through the analysis of the four subordinate themes: a) income and employment, b) educational policies and systems, c) choice, quality, and credibility in higher education, and d) the general education landscape. The findings suggest that different aspects of higher education evoke varying degrees of positive, negative, or neutral sentiment among social media users, providing valuable insights for educational institutions to address specific areas of concern. Furthermore, social media provides a platform for intimate interaction between individuals and educational institutions, potentially shaping both sides' attitudes and actions. Overall, our findings shed light on the emerging intersection of educational policy, politics, and social media.

References

- [1] Marr B. How much data do we create every day? The mind-blowing stats everyone should read [2024-10-1]. Retrieved from <https://bernardmarr.com>.
- [2] Savigny H. Public opinion, political communication and the internet. *Politics*, 2002, 22: 1-8.
- [3] Li X, Li Z. On the advantages and disadvantages of internet public opinion. *Frontiers in Business, Economics and Management*, 2021.
- [4] Huang C, Yang C, Wang S. Evolution of topics in education research: a systematic review using bibliometric analysis. *Educational Review*, 2019.
- [5] Baum S, Kurose C, McPherson M. An overview of American higher education. *The Future of Children*, 2013, 23: 17-39.
- [6] UNESCO. UNESCO Institute for Statistics (UIS) database. [2024-9-25]. Retrieved from <http://data.uis.unesco.org>.
- [7] Williams J, Usher A. 2022 world higher education: institutions, students and funding. Higher Education Strategy Associates, 2022.
- [8] Lai L T. Number of monthly active users of Weibo Corporation from 2nd quarter of 2018 to 2nd quarter of 2024. [2024-10-1] Retrieved from <https://www.statista.com>.
- [9] Huang R, Sun X. Weibo network, information diffusion and implications for collective action in China. *Information, Communication & Society*, 2013, 17(1): 86-104.
- [10] Greenhow C, Galvin S, Willet B. What should be the role of social media in education? *Policy Insights from the Behavioral and Brain Sciences*, 2019, 6: 178-185.
- [11] Horowitz J. Relative education and the advantage of a college degree. *American Sociological Review*, 2018, 83: 771-801.
- [12] García-Morales V, Garrido-Moreno A, Martín-Rojas R. The transformation of higher education after the COVID disruption: emerging challenges in an online learning scenario. *Frontiers in Psychology*, 2021.
- [13] Tilak J. Higher education: a public good or a commodity for trade? *PROSPECTS*, 2008, 38: 449-466.
- [14] Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: a survey. *Ain Shams Engineering Journal*, 2014, 5(4): 1093-1113.
- [15] Rohani V A, Shayaa S. Toward a comprehensive framework of social media analytics. In: Emrouznejad A, Charles V. (eds). *Big Data and Blockchain for Service Operations Management. Studies in Big Data*, 2022, 98.
- [16] Yang S. Sentiment analysis of COVID-19 on weibo text using optimized bi-lstm model, 2023.
- [17] Wang Y, Fikis D. Common core state standards on Twitter: public sentiment and opinion leaders. *Educational Policy*, 2019, 3: 650-683.
- [18] Gong B, Liu H. Subject, emotion, focus and Governance of Internet Public Opinion on Education Equity in China: A Python Analysis based on 2023 Weibo Data. *Educational Development Research*, 2024, 44(12): 76-84.
- [19] Devi S, Dhaval C, Moharkar L, Khanvilkar S. Impact of online education and sentiment analysis from Twitter data using topic modeling algorithms. *International Journal of Applied Sciences and Smart Technologies*, 2022, 4(1): 21-34.
- [20] Baragash R, Aldowah H. Sentiment analysis in higher education: a systematic map review. In: *Journal of Physics: Conference Series*, 2021, 1860(1).
- [21] Chan W K. Higher education and graduate employment

in China: challenges for sustainable development. Higher Education Policy, 2015, 28: 35-53.

[22] Lasri I, Riadsolh A, Elbelkacemi M. Real-time Twitter sentiment analysis for Moroccan universities using machine learning and big data technologies. International Journal of

Emerging Technologies in Learning (Ijet), 2023, 18(05): 42-61.

[23] Baumann F, Lorenz-Spreen P, Sokolov I, Starnini M. Modeling echo chambers and polarization dynamics in social networks. Physical Review Letters, 2019, 124(4): 048301.