Can Social Interactions Support Women in STEM Education? The Data Mining Approach

Junhe **YANG**
University of North Texas, United States
junheyang@my.unt.edu

Scott **WARREN**
University of North Texas, United States
scott.warren@unt.edu

Women are significantly underrepresented in STEM fields. Despite several studies proposed various interventions and measures to address STEM education gender inequality, these efforts are often constrained by traditional research methods and limited participant knowledge. This study addresses these gaps by leveraging data mining to analyze the roles and contributions of individuals using the hashtag #WomeninSTEM on Twitter. Analyzing 101,432 Twitter posts using social network and topic modeling analysis, this study investigates the dynamics of participant interactions and the nature of shared information within the #WomeninSTEM site. The findings reveal the #WomeninSTEM site as a disseminating information network comprising several small communities. Participants actively shared diverse information, including bolstering gender diversity, disseminating female success stories, sharing job opportunities, promoting online events, and collaborating on projects. This study provides novel insights and methodological approaches, emphasizing the critical role of online social networks in advancing women's participation and success in STEM fields.

Keywords: Women in STEM fields, Social interactions, Social network analysis, Topic modeling analysis, Twitter data

Introduction

The science, technology, engineering, and mathematics (STEM) fields are crucial. Developing and expanding the STEM fields is a significant task in the United States, which relies on vast numbers of STEM-trained workers worldwide. STEM education is meant to equip people with problem-solving skills, social communication skills, technology and engineering skills, system skills, and time, resource, and knowledge management skills needed in the STEM job markets (Jang, 2016). STEM education also leads to better public perspectives and understanding of science and prepares people to use technical skills to address the issues society faces today (Lee & Grapin, 2022).

Although STEM education is growing, women studying STEM subjects are underrepresented; only around 20% of bachelor's degrees with STEM majors are women (Burke et al., 2022). It transfers that women are also particularly underrepresented in the STEM workforce. National Science Foundation (2022) show that only 34% of STEM workers are women in 2019. Men occupy more positions in STEM jobs. Therefore, supporting women in pursuing STEM majors should advance gender equity and reduce gender gaps in the STEM fields.

Governments and institutions in the United States have proposed many initiatives to improve STEM education for women and girls. For instance, National Science Foundation funded the National Girls Collaborative Project (NGCP), which sought to ensure all girls have access to resources for enhancing STEM education. This project's goals includes creating and sharing research articles, exemplary practices, webinar recordings, and developing a network of leaders and experts in STEM education (NGCP, 2022). Women in Engineering ProActive Network (WEPAN) is a non-profit organization working on transforming the culture in engineering education to attract, retain and graduate more women since 1990. The project aims to develop a network to make the engineering environment more comfortable for women (WEPAN, 2022). The American Association of University Women (AAUW) sought to provide women equal opportunities in STEM education (Corbett & Hill, 2015). They create an equity network to provide women with the tools and resources they need to be successful in STEM education. Among current initiatives, creating a social networking site is an emerging strategy to support women in STEM education.

X, formally Twitter, is a social media platform where users can post a message with the easy categorization of a conversation using a hashtag (#) (Anger & Kittl, 2011). For example, #WomeninSTEM is a popular hashtag and interaction networking site that involves many participants and provides diverse information for supporting women in STEM fields on Twitter. According to the social media data insights from Linsey (2020), Twitter users tag their tweets with the hashtag #WomeninSTEM to praise the women who contribute critical developments in STEM fields. Some organizations also use this hashtag to promote training and career opportunities for girls in STEM education. #WomeninSTEM hashtag evolve through various discussions, such as individuals sharing inspiring stories and asking questions about women's underrepresentation in STEM fields.

#WomeninSTEM on Twitter is an appropriate data collection site for this study because it has initiated large-scale conversations among professionals with a high volume of participation needed by a data mining approach focused on women in STEM fields. However, it remains unknown what the #WomeninSTEM site looks like as an online social network for women in STEM education, how participants helped women pursuing in STEM education according to their posts, and what information the participants contributed to the site.

Theoretical Framework

To investigate the Twitter participants' engagement and interactions in the #WomeninSTEM site, this study adopts the social media engagement (SME) theory advanced by Di Gangi and Wasko (2016). SME theory predicts that the user experience will influence user engagement on social media. Afterward, high user engagement will positively impact the usage of social media. SME theory emphasizes that higher user engagement leads to greater social media usage. It means that when users frequently engage in the activities, the site will become more valuable to one community or group.

In SME theory, two factors shape the user experience: social interactions and technical features of social media. Social interactions, as a term in this study, refers to the communications among users on social media (Prahalad & Ramaswamy, 2004). Di Gangi and Wasko (2016) proved that when a critical mass of social acquaintances is active on the same site, users perceive their experiences are significant and valuable. Users frequently share personal experiences with acquaintances and others, resulting in the site becoming meaningful. In the Twitter context, users follow their acquaintances, close friends, or family members' accounts to see their updates in their timelines. In such a network, users become engaged in sharing helpful information and discussing their interests through different types of interactions, such as retweets, replies, or mentions (Sultana et al., 2016).

Additionally, Zhao and Rosson (2009) indicated that communicating with different Twitter users can help users acquire new information and knowledge about their professions, gain various perspectives, and even find new collaboration opportunities. How Twitter participants interacted with each other on the #WomeninSTEM networking site remains unknown. Such participants' interactions with others on the site generate supportive and collaborative information for women in STEM education. Hence, examining how and whether the participants' interactions on the #WomeninSTEM site support women in STEM fields is important if improvements are to be made to the approach.

Technical features are "the perceived capabilities of the technology" (Di Gangi & Wasko, 2016) that improve the user experience and shape their interactions. Using SME theory, Di Gangi and Wasko (2016) explained that once users realize one site can meet their information needs, they frequently contribute, retrieve, or search for information, and the site becomes valuable to them and others. Twitter is an instant messaging system that provides various and dynamic information to users. People are engaged in Twitter interactions because they can access recently posted information, share personal information, and express their perspectives (Aladwani, 2015). It thus stores comprehensive information that meets various users' needs and advances their experiences. However, few studies examined what information participants shared on Twitter for supporting women in STEM fields. Therefore, analyzing the information in the #WomeninSTEM site will promote women in STEM fields seeking help on social media.

SME theory is well-suited for this study because it offers and explains how and why participants contributed to the #WomeninSTEM site on Twitter. Examining the #WomeninSTEM site will generate new and diverse insights to sustain women in STEM fields. To guide this research project, this study seeks to answer the following research questions:

- Question 1: How did participants interact within the #WomeninSTEM site on Twitter?
- Question 2: What information did participants share in the #WomeninSTEM site on Twitter?

Methodology

The data associated with the hashtag #WomeninSTEM is collected and analyzed to examine participants' engagement and interactions within this network. The number of participants, times of interactions, density of the interaction, and degree of interaction centrality, are measured to show how participants interacted in the #WomeninSTEM site on Twitter. The figure depicts the #WomeninSTEM site's interactions. Then, topic modeling analysis is applied to analyze the information shared within this network. The following section will introduce the data collection procedure, methods for addressing each research question, and data analysis procedure.

Data Collection

The public tweets that include the hashtag "#WomeninSTEM" were collected during a period. The retrieved method focused on participants who utilized the #WomeninSTEM hashtag and excluded participants who may have engaged with this network without using this hashtag. The raw tweets were obtained through the academic Twitter Application Programming Interface (API) (Twitter Developer, 2022), which enables the retrieval and archiving of tweets for research purpose. Data collection commenced once this study's Human Subjects and Institutional Review Board (IRB) was obtained. The data collection process spanned three weeks due to Twitter data retrieval restrictions. Only English tweets was kept in the raw data because English is the common language among the authors. The total volume of raw data collected amounted to 101,432 tweets. Each entry in the raw data comprises the author's ID number, the main content of the tweet, an indication of whether it is a retweet, mentions, replies, hashtags, time zone, and profile information, including profile description, image, and text color. All collected raw data were merged into one JavaScript Object Notation (JSON) document for further analysis.

Social network analysis (SNA)

Social network analysis (SNA) is a method of analyzing social networks such as Twitter. A social network considers of a set of social entities, which can be people, groups, or communities, exhibiting a pattern of interactions between them (Oliveira & Gama, 2012). Within a social network, a node represents a social entity, and an edge signifies the interaction between two nodes (Scott, 2012). SNA serves the purpose of mapping and measuring the interaction patterns among these social entities within social networks, especially who knows whom and who shares what information with whom. On Twitter, interactions, including retweets, replies, user mentions, and follows, take place among nodes or participants. An edge is established to denote these interactions and connect the involved nodes. In this study, the participants who used the #WomeninSTEM hashtag on Twitter were considered as nodes, and the edges were representative of the interactions among participants, which encompassed retweets, replies, user mentions, and follows.

To answer the first research question (Q1), which examined the overall interaction pattern of the #WomeninSTEM site, this study employed NetworkX (Hagberg et al., 2008), a Python language software package, to measure and plot Twitter data from several metrics. NetworkX is a tool for generating various types of networks, analyzing network structure, visualizing networks, among other functions (Akhtar, 2014). In this study, NetworkX was applied to analyze the #WomeninSTEM site on Twitter, enabling the description of how participants interacted within this network.

First, the number of nodes and edges presented in the figure were computed. The degree of a node refers to the number of interactions (edges) for each node, describing the number of interactions a particular node has. This study also examined the maximum degree, minimum degree, average degree, and the most frequent degree of the nodes in the graph. Density was examined to assess how well the site interacts, representing the ratio of interactions to the figure of possible interactions.

Then, to describe the overall interactions among participants in the #WomeninSTEM site, this study analyzed groups of participants using modularity, a metric that quantifies the tendency to form communities of tightly connected nodes within the network. Higher modularity indicates denser interactions within the same community and less dense connections with other communities. The more modularity that exists, the more disconnected the network is. This measure aids in determining whether the #WomeninSTEM site represents a single, tightly connected conversation community or several smaller, distinct communities.

Next, the centrality metrics were calculated to assess how often a participant interacted with others on the site. Centrality values can identify the key and active nodes in the network. For instance, indegree, which measures the number of interactions received (e.g., being mentioned or retweeted), helps determine the influence a node has based on the messages they tweet or their perspectives. A higher the indegree indicates that the node receives more retweets and mentions. Outdegree, which measures the number of interactions sent (e.g., sending a tweet), signifies the variety of nodes that a participant retweeted or mentioned within the network. Research by Bastos et al. (2013) showed that Twitter users with high outdegree significantly impact social conversations and even influence other users' perceptions on a subject. In this study, the top 10 active nodes, their roles, as well as their indegree and outdegree values were determined based on degree centrality. These results provides insights into the most influential participants in the network and their interaction patterns.

Topic modeling analysis

The collected tweet data underwent preprocessing to address the second research question (Q2). This process involved removing duplicate tweets and meaningless words inside tweets. Duplicate tweets, resulting from retweets, were removed to ensure that redundant content was not re-analyzed. During preprocessing, all duplicates' original content was kept, and all retweets were removed. The dataset was thereby reduced to 9,892 unique tweets. Subsequently, regular expressions were utilized to tokenize the text into individual words and remove all "stopwords" (e.g., "I," "me," "they," "he," "she," etc.) that are meaningless to content analysis numbers, and symbols, tags, URLs, hashtags before topic modeling by following the standard approach in topic modeling analysis (Johri & Bansal, 2018). Additionally, all uppercase letters were converted to lowercase to ensure that all identical terms were detected by the computer. This adjustment accounted for cases where participants used all capitalized letters in their tweets, such as "WOMEN IN STEM", which was considered equivalent to "women in stem". This standardized preprocessing enhanced the data for subsequent topic modeling analysis.

Following preprocessing, this study proceeded to identify several topics using Latent Dirichlet Allocation (LDA), a method introduced by Blei et al. (2013). This approach facilitated the determination of the kind of information participants shared on the #WomeninSTEM site. LDA is a statistical method used to cluster words based on their frequency and similarity. By analyzing the contents of these topics, the study aimed to uncover the information that users shared on Twitter using the #WomeninSTEM hashtag.

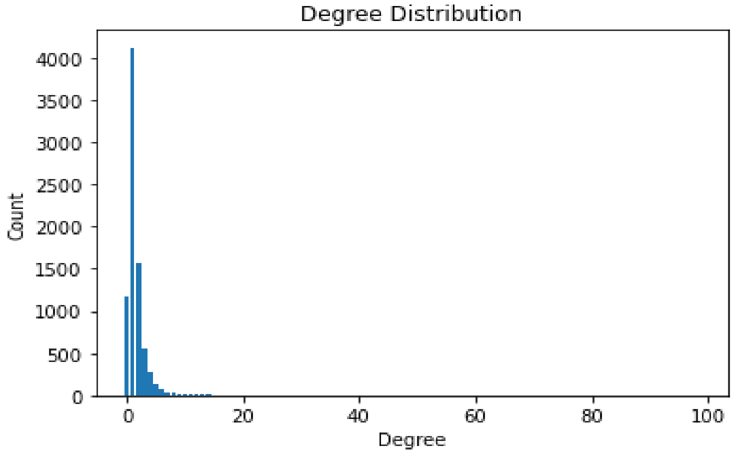
Results

In this section, the findings of this study are presented below:

Patterns of Interaction in the #WomeninSTEM Twitter Community

Based on the results derived from the social network analysis conducted within the #WomeninSTEM site on Twitter. A total of 8,037 Twitter participants were identified actively engaged within the #WomeninSTEM network. These participants collectively engaged in interactions, including replies, user mentions, retweets, and follows, which amounted to 11,581 interactions. The highest number of interactions recorded for one participant was 98 interactions. On average, each participant within the network had 2.88 interactions, providing an insight into the typical engagement level of people involved in the #WomeninSTEM site. Many participants had exactly three interactions with others, making this the most frequent degree in the network. Figure 1 displays the distribution of degrees in the #WomeninSTEM site. Specifically, 2,411 participants exhibited a degree of 1, implying that they had only one interaction with others. 1169 participants were isolated nodes without any interactions (edges). This observation suggests the existence of participants who may have limited and few interaction in the #WomeninSTEM site. Further, the network density, a measure of interaction extent, was remarkably low at only 0.0036. This finding signifies that the participants in the #WomenisSTEM site are loosely connected by replies, user mentions, retweets, and follows.

Figure 1.
Degree distribution in the #WomeninSTEM site



Moreover, this study assessed the #WomeninSTEM network's modularity value. The modularity value was 0.91, a significant indicator of robust community structures within the network (Newman, 2006). The #WomeninSTEM site exhibited a distinctive pattern of tightly interconnected communities. Within these distinct communities, participants actively engaged and connected with others, fostering dense internal interactions (e.g., reply, user mention, retweet or follow). It is noted that participants belonging to the same community share similar information and interests, while interactions between participants from different communities are comparatively sparse. The subsequent part will present the top 10 active participants and their subgraphs in the #WomeninSTEM site.

The top 10 active nodes exhibiting high degree centrality within the #WomeninSTEM network are listed in Table 1. These influential nodes play a pivotal role in bridging different communities by facilitating the dissemination of information across the network. Notably, they encompassed a diverse spectrum of roles and interests, ranging from data scientists such as Node 2 to organizations like Node 3, all of which champion the cause of women in STEM fields. This remarkable diversity underscores the inclusive character of the #WomeninSTEM network. It is worth highlighting that these top 10 nodes possess high outdegree values while displaying relatively lower indegree values. This pattern may suggest that these participants primarily focus on content creation and information sharing, serving as vital conduits for connecting disparate communities within the #WomeninSTEM network. Their influence lies in their capacity to disseminate perspectives and resources, thereby empowering and advancing women's participation in STEM education.

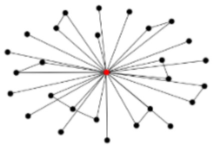
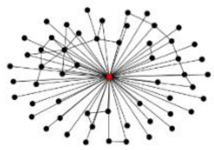
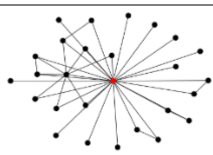
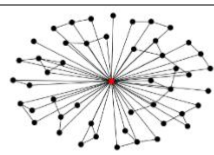
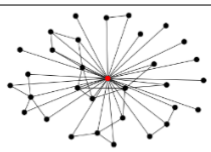
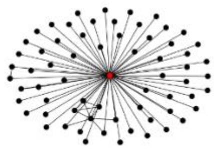
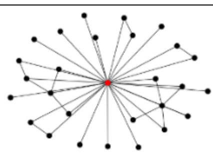
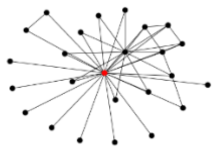
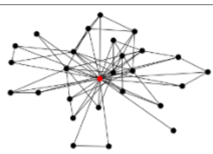
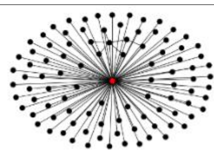
Table 1
The top 10 nodes with high degree values

Node ID	Type	Degree Centrality	Indegree	Outdegree
Node 1	N.A.	98	1	97
Node 2	Data scientist	67	2	65
Node 3	Support young women organization	57	3	54
Node 4	Researchers, Epidemiologist	48	1	47
Node 5	Founder & CEO	33	4	29
Node 6	Support women scientists' organization	28	3	25
Node 7	U.S. government administration	27	2	25
Node 8	Support girls' organization	27	0	27
Node 9	CEO	26	4	23
Node 10	Support girls in STEM organization	24	0	24

Additionally, Table 2 lists the top 10 active nodes' subgraphs and follower counts to examine their dynamic of interactions and impacts within the network. These subgraphs reveal distinct communities centered around the active participants. In each subgraph, the active participant is highlighted by a red node. These active participants play pivotal roles in connecting and influencing others in their community. The black nodes are their neighbors, which means the active node has interactions with them. The edges (black lines) between these nodes in their subgraph signify interactions, including replies, user mentions, retweets, or follows. In the context of #WomeninSTEM site, the subgraphs may center around different topics to promote and support women in STEM fields. Furthermore, the follower counts of these active participants reflect the extent of their reach and influence in their community. Some participants amassed a substantial and diverse audience, amplifying their impact on the network. For example, organizations related to STEM fields, such as Node 7 and Node 3, exhibit a massive Twitter followers. Their significant present underscores their important role in advocating for women in STEM, making them influential voices within the network. Some individual participants, such as female scientist (Node 4) and CEO in the company (Node 9), boast substantial follower counts. Their expertise and engagement contribute to supporting women in STEM education and careers. These results shed light on how different communities connect and collaborate to empower women in STEM fields. This study will further explore the impact of these active participants within the #WomeninSTEM site in the discussion section.

Table 2
The subgraphs and follower count of the top 10 active nodes

In sum, the social network analysis showcases the #WomeninSTEM site with a diverse and influential set of participants from various roles and backgrounds. The top 10 active participants are instrumental in sharing

Node ID	Follower Count	Subgraph	Node ID	Follower Count	Subgraph
Node 7	77M		Node 3	54.8K	
Node 9	33.7K		Node 4	24.6K	
Node 5	4,875		Node 2	3,742	
Node 8	1,472		Node 10	643	
Node 6	594		Node 1	295	

information, facilitating connections, and advocating for important causes related to women in STEM fields. Understanding the roles and influence of these participants can inform targeted engagement, information sharing, and community-focused initiatives for women in STEM education. Moving to the next research question, this study will seek what information the participants shared in the #WomeninSTEM site.

Information Dissemination in #WomeninSTEM Twitter Network

This study used topic modeling analysis to identify the popular topics shared on the #WomeninSTEM site. Five topics were labeled using the LDA model based on each topic's top 20 most relevant terms. To determine the optimal number of topics, the author experimented with various options, including 8, 10, 15, 20, 25, and 30. To ensure the reliability of the coding process, another coder, who is also a co-author of this study, was involved. The coder holds the doctoral degree in education and have extensive experience in educational qualitative research. Prior to the analysis, this coder underwent thorough training on the study's coding framework and methodology to ensure consistency and alignment with the research objectives. The training included the discussions on the coding scheme and practice calibration exercises. Consequently, the interpretations of the five topics were consistent and aligned, as presented in Table 3.

Table 3
The five topics in the data corpus

Topic ID	Terms	Label
Topic 0	way, help, time, barriers, meet, share, breaking, diversity, looking, science, well, great, new, know, love, tech, gender, #womenintech, international, global	Supporting gender diversity
Topic 1	engineering, celebrate, generate, great, highlights, science, happy, #womeninscience, contributions, first, girls, make, role, need, achievements, meet, progress, annual, inspirational, join	Sharing success stories and achievements
Topic 2	industry, science, #stem, LinkedIn, new, key, congrats, work, honored, join, career, see, space, opportunity, business, engineers, young, great, female, students	Great job opportunities sharing
Topic 3	amp, amazing, twitter, next, great, exciting, new, event, look, empowering, life, talk, engineer, love, like, hear, science, female, world, presentation	Online events and presentations
Topic 4	today, join, research, us, #womeninscience, phd, project, first, want, science, excited, new, dr, work, latest, school, know, fantastic, student, discuss	Collaborative projects

In the topic modeling analysis, topics are labeled numerically, following the indexing convention of the software used (e.g., 'Topic 0,' 'Topic 1'). 'Topic 0' represents the first topic generated by the model. We retained these labels as they align with the indexing convention used in Python programming, as shown in Table 3. The findings of the topic modeling analysis indicate the diverse range of information shared on Twitter to support women in STEM fields. The identified topics focused on supporting gender diversity, sharing success stories and achievements, great job opportunities sharing, online events and presentations, and collaborative projects. These categories offer insights for women in STEM education, equipping them with tailored information and support available on the #WomeninSTEM site. By categorizing the content into these thematic areas, this site can become a more effective resource for women pursuing STEM fields, offering them a structured and accessible platform for information and inspiration. After getting the above results, it is important to discuss the implications of this study. The following section will present how such interactions and information support women pursing in STEM education.

Discussion

The results from the social network analysis reveal insights into the nature of interactions and community structures within the #WomeninSTEM site on Twitter. The data demonstrates that the interactions among participants in the #WomeninSTEM site varies significantly. While the highest number of interactions for one single participant reached 98, the average participant engaged in only 2.88 interactions. It is noteworthy that 1,169 participants were identified as isolated nodes, lacking any interactions or edges. This observation suggests that a few participants are active and enthusiastic about discussing the topic related to women in STEM with others, while many participants may not have interactions. Overall, the social network of the #WomeninSTEM site is sparse, with a low density of 0.0036. This signifies that participants within the #WomeninSTEM site are loosely connected, indicating that the interactions within the network are scarce, and participants may not be extensively interconnected. However, Miño-Puigcercós et al.'s study (2019) demonstrated that online social networks serve as a safe space, particularly for young women, to better understand their personal situation, express their concerns about patriarchal society, and share their experiences cross various social contexts, such as their workplaces. Consequently, the #WomeninSTEM site could emerge as a supportive online social network for women pursuing in STEM education. Efforts should be made to encourage more participants' engagement and promote interactions within the #WomeninSTEM site. One effective strategy is to encourage participants to regularly post relevant content attached to a hashtag #WomeninSTEM, which may include success stories, educational resources, and pertinent news. Additionally, collaborating with large educational institutions and organizations to integrate the topic of women in STEM fields into their social media account and outreach efforts is essential. Such endeavors could help create an even more nurturing and empowering online space for women in STEM disciplines.

Although the network of #WomeninSTEM site was low density, its modularity value was high, measured at 0.9. The high modularity value suggests that the #WomeninSTEM site comprises many distinct and tightly interconnected communities. Within these communities, participants actively engaged with one another, fostering dense internal interactions through replies, user mentions, retweets, or follows. They shared similar interests, common goals, or a sense of belonging within the same community. Notably, interactions between participants from different communities were comparatively sparse, indicating that information flow and engagement between distinct communities are limited. To foster a more diverse range of perspectives and ideas within the broader #WomeninSTEM network, it is imperative to facilitate increased interactions between these distinct communities, encouraging various collaboration opportunities and information sharing. Co-hosting online events, webinars, and panel discussions focused on women in STEM fields on social media can be a strategy. It is recommended encouraging participants to use the #WomeninSTEM hashtag during these events, exchanging diverse insights, posing questions, and connecting with others who support women in STEM education.

The identification of the top 10 nodes with high degree centrality underscores the network's dynamics and its impact on supporting women in STEM fields. The presence of influential nodes highlights their active role in different communities within the #WomeninSTEM network. They frequently facilitated the exchange of information and interactions across different communities of the network. Based on their Twitter profile descriptions, these top nodes represented the diversity of roles, including scientists, advocacy organizations to supporting women in STEM, and female industry leaders. It may emphasize that the #WomeninSTEM network site is inclusive, containing information from various organizations and industries supporting women in STEM fields. However, these top nodes had high outdegree values and relatively lower indegree values. It suggests they mainly focus on content creation and information sharing. Although some nodes, such as Node 7, or Node 3 has thousands of followers, they may not receive as much engagement or interactions with others. Popescu and Badea (2020) indicated that both interactions and knowledge-sharing behaviors on social media have significant impacts on facilitating collaborative learning and improving students' learning performance. Hence, it is necessary to encourage influential participants to increase interactions in the #WomeninSTEM site, not only by posing informative content but also by actively engaging in discussions and responding to questions or mentions. This can lead to stronger bonds and a more supportive environment for women pursuing STEM education.

To examine what information contributes to supporting women in STEM education, this study used the LDA model to identify five prevalent topics within the #WomeninSTEM site on Twitter. The topic 0 primarily revolved around fostering gender diversity for women in STEM, particularly in the technology sector (#womenintech). The information shared by participants around this topic encompassed various barriers that women encountered in STEM fields, including issues like stereotype threat, gender bias, or lack of representation. Furthermore, participants shared their strategies to overcome these barriers. In Kricorian et al.'s study (2020), they found that exposure to diverse STEM professionals through social media platform can provide support and encouragement to underrepresented students to pursue STEM studies. Falco and Summers (2019) demonstrated that addressing gender barriers can be an effective strategy for enhancing the development of women in STEM fields. Thus, by sharing information related to gender equity in the #WomeninSTEM site, women pursuing STEM education can access targeted assistance and advice to navigate challenges and make informed decisions about their STEM education and future career paths.

Topic 1 was closely associated with the sharing of success stories and the celebration of achievements. This topic seems to focus on highlighting and honoring the accomplishments and contributions of women in STEM fields, especially within the realm of science (#womeninscience). Information shared under this topic likely revolved around showcasing success stories of people who made significant progress in STEM fields. It may involve the information related to annual activities that celebrate women's achievements, or discussions about the importance of role models for girls and young women in STEM fields. González-Pérez et al.'s study (2020) examined the influence of actual female role models in enhancing young girls' preferences in STEM studies. They indicated that sharing professional and personal experiences of real-life female role models with young girls can significantly motivate them to acquire STEM skills. Drawing from this observation, the sharing of success stories and achievements on social media can serve as inspirational sources for women what are contemplating or actively pursuing STEM education.

Topic 2 focused on the sharing of information regarding job opportunities within STEM fields, with a specific focus on opportunities targeting women. The terms encompassed by this topic suggest that the #WomeninSTEM site on Twitter provides a wide range of new STEM job opportunities for women. One characteristic of using Twitter for job searches is its usefulness in providing frequent real-time updates, making it easier for people to access a large volume of job information (Zhao & Rosson, 2009). However, it is worth noting that social media platforms, including Twitter, are susceptible to issues like misleading information and fake job postings (Amaar et al., 2022). Hence, although the dissemination of STEM job opportunities on social media is beneficial to women who are graduating in STEM education and preparing to enter STEM careers, it is imperative for them to exercise caution in detecting potential fake job postings on social media.

Topic 3 centered around the sharing of information regarding online events and presentations. The terms associated within this topic indicate that the participants held positive attitudes toward event and presentation information related to science and engineering on Twitter. Online events have proven to be effective in facilitating the sharing of knowledge (Kharouf et al., 2020). These online talks and presentations play a crucial role in keeping women who are pursuing STEM education updated on the latest advancements, fostering continuous learning in the realms of engineering and science fields. Research by Gangwani et al. (2021) and Odine (2013) illustrated that online events and presentations held on social media can empower women from different regions to combat society, education, and professional inequalities. Through participation in online events and presentations on social media, women pursuing STEM education can enhance their STEM skills and awareness, equipping them to overcome challenges within STEM fields. Consequently, by sharing information about online events and presentations on social media, women pursuing STEM education can access more opportunities to broaden their perspectives and gain empowerment in STEM fields.

Topic 4 revolved around the dissemination of information related to collaborative projects, with a focus on women in the science field. This topic likely encompassed the introduction of specific Ph.D. research projects tailored for women in science or the sharing of recent projects aimed at attracting women students to participate and engage in discussions. Research conducted by Atkins et al. (2023) underscored the significance of collaborative research experiences, particularly for underrepresented groups in STEM. Such experiences can facilitate the establishment of connections among STEM students with peers, mentors, and professionals. This collaborative network, in turn, serves to amplify the academic and professional growth of underrepresented STEM students. Furthermore, a study conducted by Kelley et al. (2020) examined the efficacy of collaborative projects, specifically the Teachers and Researchers Advancing Integrated Lessons initiative. This project not only enhances STEM teachers' comprehension of technology but also bolsters their confidence in teaching. Collaborative project experiences have been shown to positively influence participants' self-efficacy, contributing to the augmentation of their professional knowledge and practices. Hence, the exposure of collaborative projects information on social media can foster connections, enhancing STEM skillsets, and boosting self-confidence among women in STEM education.

Conclusion

This study aims to investigate how the #WomeninSTEM site on Twitter supports women in STEM education. Employing social network analysis and topic modeling, this study explored interaction patterns and information sharing within this network. The findings highlight that the #WomeninSTEM site on Twitter functions as an information dissemination network, comprising numerous small communities dedicated to discussions related to women in STEM fields. These communities exhibited a high degree of cohesion, with members sharing common interests and information to advance the cause of women in STEM. However, a notable observation is the limited interaction between these distinct communities. Future efforts should prioritize promoting interactions not only within but also between these communities. Additionally, this study identified the top ten active participants in the network, each representing diverse roles in supporting women in STEM fields. While serving as information conduits, these active participants have the potential to further enhance their impact by fostering more interactions and engagement within the network. Moreover, this study also summarized five topics that encapsulate the nature of information sharing within the network, including supporting gender diversity, sharing success stories and achievements, great job opportunities sharing, online events and presentations, and collaborative projects. These categories provide insights for women in STEM education, equipping them with readily accessible and tailored information and support available within the #WomeninSTEM site. This study significantly contributes to the understanding how interactions on social media network, such as #WomeninSTEM, support and empower women in STEM education. It also underscores opportunities for improvement in facilitating more extensive collaboration and information sharing among participants within the network. Future study can apply similar research methods to analyze various social networks on different platforms, potentially providing diverse insights to further support women in STEM fields and reduce gender gap in STEM education.

Limitation

This study has several limitations that should be acknowledged. Firstly, it was observed that some highly active participants within the #WomeninSTEM network lacked Twitter profile descriptions or potentially provided inaccurate information on their social media profiles. Although the author conducted a manual review of all identified profile descriptions and tweets posted by these active participants, addressing this issue proved to be challenging. This study acknowledges this limitation it may pose in the data analysis process.

Secondly, the data collection for this study was confined to Twitter, and the analysis focused exclusively on tweets and interactions within the #WomeninSTEM site on Twitter. This limits the scope of the study, as other social media platforms or networks may also contribute to the support of women pursing in STEM education. Examining different social media platforms or networks could provide different insights to support women in STEM education.

Thirdly, it is essential to acknowledge that the data collection period for this study was constrained to several weeks. This limitation stemmed from the evolving accessibility of data on Twitter for academic research purposes. Twitter increasingly restricted access to its data for research endeavors during this study's data collection period. To continue to access Twitter data for academic analysis, researchers need to pay substantial costs for getting data licenses. Given these constraints, the data collection period in this study was restricted.

Finally, this study exclusively analyzed English-language tweets. While English is a widely used language on social media, particularly in the academic and professional context, it is essential to acknowledge that discussions and support for women in STEM education occur in multiple languages. Inclusivity could be enhanced by incorporating data from different regions and considering tweets in various languages, which would offer a broader view of the support for women in STEM globally.

Future Studies

Future research work could address the limitations of this study and build upon its findings by considering the following aspects. Future research should consider a more extensive exploration of various social media platforms and online communities. Different platforms such as Treads, LinkedIn, Instagram, or other professional networking sites offer unique spaces where discussions and support for women in STEM education take place. Each platform may serve distinct functions, attract diverse user demographics, and offer varying forms of engagement for supporting women in STEM education. Additionally, researchers can use other popular hashtags, such as #womeninscience, #womenincode, #womenintech, to examine such networks how to support women in STEM fields. It enables a more thorough understanding of the diverse networks and support structures available to women in STEM fields, which, in turn, can lead to different and new strategies for empowering and advancing women in these disciplines.

Future research efforts, if adequately funded or equipped with data access from alternative social media sources, present an opportunity to conduct more extensive and long-term investigations. Such research can delve into the enduring impact of participation in online communities and engagement on social media platforms on supporting women in STEM fields. By studying these long-term effects, researchers can offer insights into the sustainability and lasting benefits of these support networks in advancing women's educational pursuits in STEM. This approach has the potential to inform strategies for promoting gender equity in STEM fields over the long run.

Furthermore, a broader examination can provide a more inclusive view of global efforts to support women in STEM by considering platforms that are popular in various regions and among different linguistic groups. This diversity in languages may help tailor support to specific cultural contexts and address the needs of women pursuing STEM education in different parts of the world.

References

Anger, I., & Kittl, C. (2011, September). Measuring influence on Twitter. In *Proceedings of the 11th international conference on knowledge management and knowledge technologies* (pp.1-4).

Aladwani, A. M. (2015). Facilitators, characteristics, and impacts of Twitter use: Theoretical analysis and empirical illustration. *International Journal of Information Management, 35*(1), 15-25.

Akhtar, N. (2014, April). Social network analysis tools. In *2014 fourth international conference on communication systems and network technologies* (pp. 388-392). IEEE.

Atkins, K., Dougan, B. M., Dromgold-Sermen, M. S., Potter, H., Sathy, V., & Panter, A. T. (2020). "Looking at myself in the future": How mentoring shapes scientific identity for STEM students from underrepresented groups. *International Journal of STEM Education, 7*, 1-15.

Amaar, A., Aljedaani, W., Rustam, F., Ullah, S., Rupapara, V., & Ludi, S. (2022). Detection of fake job postings by utilizing machine learning and natural language processing approaches. *Neural Processing Letters*, 1-29.

Burke, A., Okrent, A., Hale, K., & Gough, N. (2022). *National science board science & engineering indicators*. National Science Foundation. https://files.eric.ed.gov/fulltext/ED619338.pdf

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research, 3*(Jan), 993-1022.

Bastos, M. T., Raimundo, R. L. G., & Travitzki, R. (2013). Gatekeeping Twitter: Message diffusion in political hashtags. *Media, Culture & Society, 35*(2), 260-270.

Corbett, C., & Hill, C. (2015). *Solving the equation: The variables for women's success in engineering and computing.* Institute of Education Sciences.

Di Gangi, P. M., & Wasko, M. M. (2016). Social media engagement theory: Exploring the influence of user engagement on social media usage. *Journal of Organizational and End User Computing, 28*(2), 53-73.

Falco, L. D., & Summers, J. J. (2019). Improving career decision self-efficacy and STEM self-efficacy in high school girls: Evaluation of an intervention. *Journal of Career Development, 46*(1), 62-76.

González-Pérez, S., Mateos de Cabo, R., & Sáinz, M. (2020). Girls in STEM: Is it a female role-model thing?. *Frontiers in Psychology, 11*, 2204.

Gangwani, S., Alruwaili, N., & Safar, S. A. (2021). Social media usage and female empowerment in Saudi Arabia. *Academy of Strategic Management Journal, 20*(4), 1-8.

Hagberg, A., Swart, P., & S Chult, D. (2008). Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th annual Python in science conference*.

Jang, H. (2016). Identifying 21st century STEM competencies using workplace data. *Journal of Science Education and Technology, 25*(2), 284-301.

Johri, V., & Bansal, S. (2018, January). Identifying trends in technologies and programming languages using topic modeling. In *2018 IEEE 12th international conference on semantic computing* (pp. 391-396). IEEE.

Kricorian, K., Seu, M., Lopez, D., Ureta, E., & Equils, O. (2020). Factors influencing participation of underrepresented students in STEM fields: Matched mentors and mindsets. *International Journal of STEM Education, 7*, 1-9.

Kelley, T. R., Knowles, J. G., Holland, J. D., & Han, J. (2020). Increasing high school teacher's self-efficacy for integrated STEM instruction through a collaborative community of practice. *International Journal of STEM Education, 7*, 1-13.

Kharouf, H., Biscaia, R., Garcia-Perez, A., & Hickman, E. (2020). Understanding online event experience: The importance of communication, engagement and interaction. *Journal of Business Research, 121*, 735-746.

Lee, O., & Grapin, S. E. (2022). The role of phenomena and problems in science and STEM education: Traditional, contemporary, and future approaches. *Journal of Research in Science Teaching, 59*(7), 1301-1309.

Linsey (2020). *International women's day: The status of #womenintech online.* https://www.obi4wan.com/en/blog/women-in-tech/

Miño-Puigcercós, R., Rivera-Vargas, P., & Cobo Romaní, C. (2019). Virtual communities as safe spaces created by young feminists: Identity, mobility and sense of belonging. *Identities, Youth and Belonging: International Perspectives*, 123-140.

Newman, M. E. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences, 103*(23), 8577-8582.

National Girls Collaborative Project (2022, June). *Our strategic direction.* https://ngcproject.org/resources/our-strategic-direction

National Science Foundation. (2022). *Science and engineering indicators 2022: U.S. and global STEM education and labor force*. National Science Foundation. https://ncses.nsf.gov/pubs/nsb20221/u-s-and-global-stem-education-and-labor-force

Oliveira, M., & Gama, J. (2012). An overview of social network analysis. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2*(2), 99-115.

Odine, M. (2013). Role of social media in the empowerment of Arab women. *Global Media Journal, 12*(22), 1-30. Prahalad, C. K., & Ramaswamy, V. (2004). Co-creation experiences: The next practice in value creation. *Journal of Interactive Marketing, 18*(3), 5-14.

Popescu, E., & Badea, G. (2020). Exploring a community of inquiry supported by a social media-based learning environment. *Educational Technology & Society, 23*(2), 61-76.

Sultana, M., Paul, P. P., & Gavrilova, M. (2016). *Identifying users from online interactions in Twitter*. Springer.

Scott, J. (2012). *What is social network analysis?*. Bloomsbury Academic.

Twitter Developer. (2022). *Academic research access*. https://developer.twitter.com/en/products/twitter-api/academic-research

Women in Engineering ProActive Network. (2022). *About WEPAN*. https://www.wepan.org/aboutwepan

Zhao, D., & Rosson, M. B. (2009, May). How and why people Twitter: The role that micro-blogging plays in informal communication at work. In *Proceedings of the 2009 ACM international conference on supporting group work* (pp. 243-252).