

Analyzing User Feedback on a Fan Community Platform

'Weverse': A Text Mining Approach

Thi Thao Van Ho, Mi Jin Noh, Yu Na Lee, Yang Sok Kim

Abstract

This study applies topic modeling to uncover user experience and app issues expressed in users' online reviews of a fan community platform, Weverse on Google Play Store. It allows us to identify the features which need to be improved to enhance user experience or need to be maintained and leveraged to attract more users. Therefore, we collect 88,068 first-level English online reviews of Weverse on Google Play Store with Google-Play-Scraper tool. After the initial preprocessing step, a dataset of 31,861 online reviews is analyzed using Latent Dirichlet Allocation (LDA) topic modeling with Gensim library in Python. There are 5 topics explored in this study which highlight significant issues such as network connection error, delayed notification, and incorrect translation. Besides, the result revealed the app's effectiveness in fostering not only interaction between fans and artists but also fans' mutual relationships. Consequently, the business can strengthen user engagement and loyalty by addressing the identified drawbacks and leveraging the platform for user communication.

Keywords: Weverse | Topic modeling | LDA | Fan community platform | Communication

I. INTRODUCTION

Korean dramas and music have been gaining popularity in Asia for over a decade. Further, the Korean wave, known as 'Hallyu' has now spread to Europe, America, and other countries [1]. It cannot be denied that the globalization of Korean entertainment industry results in the increased demand to keep in touch with idols. Moreover, artists' efforts to engage actively with fans also contribute to the fast growth of this industry [2]. In the context of Hallyu and technological development, fan community platforms or fandom platforms were created to facilitate the communication between fans and idols.

Weverse is a global fan community platform, developed by Weverse Company (formerly beNX) which is a subsidiary of HYBE Entertainment Company. It was launched on June 10, 2019, and later in March 2022 acquired V Live which is also a fan community platform run by Naver corporation [3,4].

This application offers various functions to promote engagement between fans and artists such as direct messaging, exclusive livestream, and online concerts. Besides, it allows users to send fan letters and get access to official content and events [5].

Weverse emerged as the most popular and widely used fan community application, with 100 million downloads by June 2023 and over 120 artists joining this platform

[3,6]. Consequently, we decided to conduct a study on user feedback on this application.

The purpose of this research is to explore user experience and app issues by collecting user reviews of Weverse on Google Play Store with Google-Play-Scraper tool and then analyzing them by utilizing Latent Dirichlet Allocation (LDA), a common topic modeling technique to reveal hidden topics in a corpus. LDA results and visualization are explained to understand the extracted topics. This helps to identify which aspects need to be improved and maintained to enhance user experience as well as accelerate app performance and attract new users.

II. RELATED WORK

1. Fan community platform

Fandom is defined as a group of enthusiastic fans of someone or something [7]. According to the research by Kim and Kim, a fan community platform or fandom platform is an online community for fans who share common interests and preferences for their favorite singers, actors, and actresses [8]. It bridges the gap between fans and idols by allowing them to connect directly with each other, easily purchase celebrities' merchandise, and offering diverse entertaining content, commerce, and new media technologies on which global fans focus.

In previous studies, four main platforms in Korea were identified, including V Live, Weverse, Bubble, and Universe [8,9]. However, only Weverse and Bubble are currently competitors in this market, after

Weverse's acquisition of V Live in 2022 [3] and Universe's termination in 2023 [10].

In the study of Hong and Kim, a survey was conducted to analyze user satisfaction on Weverse and Universe [11]. As a result, both applications had diverse content and included overseas services for global fans, but Weverse provided more available languages than Universe. It also demonstrated the necessity of protecting fandom and artists from hateful comments. Moreover, users were not satisfied with network error and required a dark mode function.

Interviews with users of Weverse were conducted to evaluate its services [12]. Users were satisfied with the artist service but membership and Weverse Shop service needed to be improved. The study also pointed out some strengths of Weverse, such as easily communicating with artists due to their high frequency of access and conveniently linking to Weverse Shop, Rhythm Hive, and Weverse Magazine. This app also served as a common space for fans to share information, empathy and support each other. However, a more complicated registration was demanded to prevent anti-fans and notification mechanism needed to be improved.

Meanwhile, another research was conducted to explore the strengths and limits of Weverse and V live [13]. Notably, these platforms were suggested to decline the subscription fee and attract more artists. Moreover, market research is necessary to improve user experience.

Machado Pereira investigated some South Korean platforms to explore their operation and business strategies,

financial performance, and their prospects in the context of the emergence of metaverse platforms [14]. Moreover, Song uncovered Weverse's effort to invite more artists to join as a means of appealing to a higher number of fans [15].

2. Text mining techniques

Text mining techniques were employed to gain insights into other related fields. Posts on Naver and Daum were analyzed to identify the perceptions of Hallyu fandom, using LDA method [16]. Fandom and fandom platform were found to be the main keywords expressed by users. It was found that subscribing to these fan community platforms helps to generate economic value.

Meanwhile, using text mining techniques to investigate YouTube comments on BTS's videos discovered several different topics where fans expressed their favorite members or focused on the chemistry among BTS members [17]. The results also showed that fans have shared a sense of togetherness with BTS and have empathized with their lives.

III. METHODOLOGY

In this study, we conduct analysis codes in Python. Figure 1 illustrates our research process, where we first collect online reviews of Weverse on Google Play Store, using Google-Play-Scraper tool. After that, preprocessing is performed to get a suitable dataset for topic modeling, including tokenization, stop words removal, lemmatization, and short reviews removal. Finally, we implement LDA topic modeling

and analyze the obtained results. Each phase is described in the next sections.

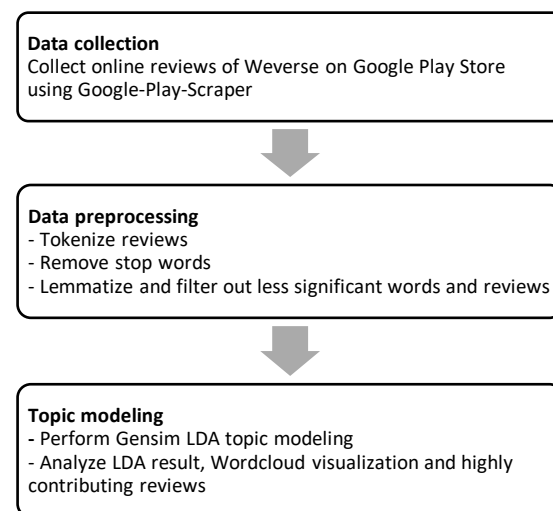


Figure 1. Research process

1. Data collection

Google-Play-Scraper provides APIs to easily crawl the Google Play Store for Python without any external dependencies [18]. Therefore, we use this library to retrieve Weverse's online reviews on Google Play Store and obtain 88,068 first-level English reviews posted from June 6, 2019 to November 28, 2023. Each review has various fields which are appVersion, reviewCreatedVersion, thumbsUpCount, username, reviewId, userImage, repliedAt, replyContent, content, at, and score [18]. But only content, at, score is analyzed in this study. It is important to note that a review might mention both users' good evaluations and app drawbacks, which will be explained in detail later in this study.

2. Data preprocessing

In the preprocessing phase, we conduct preprocessing experiments and propose a final process as shown in Figure 2. And Table 1 presents an

example of the original review and the processed review.

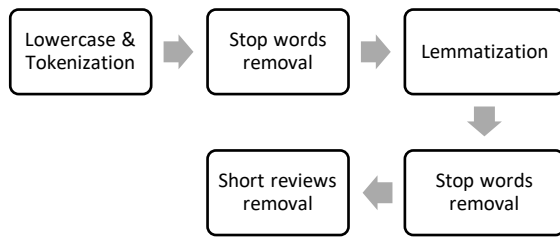


Figure 2. Data preprocessing

Table 1. Example of user review

Original review	Ever since they updated the app it does not work properly. Every time I try to open it show not connected to the network. Even the lives are not translated after a day like they used to on vlive. Is this problem which the app is facing in India only or everyone else is facing the same problem?
Processed review	['ever', 'update', 'work', 'properly', 'time', 'try', 'open', 'show', 'connect', 'network', 'even', 'live', 'translate', 'day', 'use', 'vlive', 'problem', 'face', 'else', 'face', 'problem']

Reviews are converted to lowercase tokens by using Gensim simple preprocess. Next, Natural Language Toolkit (NLTK) library's list of English stop words and our extended stop words are removed from the dataset. The extended stop word list includes 'app', 'application', 'good', 'weverse', 'great', 'amazing', 'awesome', 'love', 'bts', 'txt' and 'gfriend', which are frequent words but do not significantly contribute to analysis results. Then, to get the root words, Spacy's English lemmatization is applied for nouns, verbs, adjectives, and adverbs in reviews which are parts of speech (POS) tags in lemmatization. However, it is observed that a set of lemmatized words is present in the predefined stop word list. Therefore, an additional stop words removal is conducted. Finally, reviews with less than 3 words are filtered out because of their minor

significance. After data preprocessing, the official dataset of 31,861 reviews is obtained.

3. Topic modeling

In this section, Gensim LDA algorithm is applied to unveil the hidden topics of Weverse's online reviews. Gensim is designed to process raw, unstructured digital texts ('plain text') using unsupervised machine learning algorithms [19]. Because of unsupervised learning, the number of topics is unknown and needs to be pre-defined by researchers.

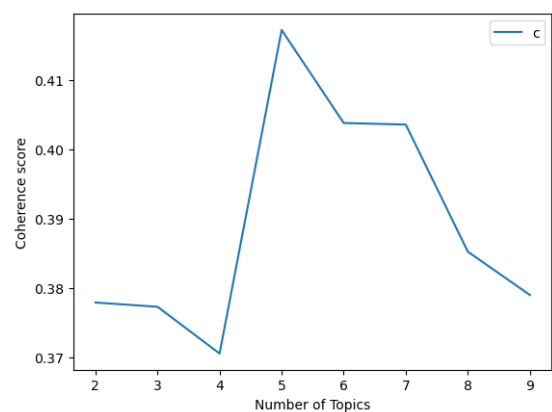


Figure 3. Coherence score

In previous studies, one of the methods to choose the optimal number of topics in topic modeling is coherence score, evaluating the correlation between topics and the level of semantic similarity between words in a topic [20,21]. A higher coherence score implies a better model. Moreover, Röder et al. proposed that the best performing measure is c_v coherence measure [22]. This is obtained after four steps: (i) segmenting the data into word pairs, (ii) calculating of word or word pair probabilities, (iii) calculating

how strongly a word set supports another word set, (iv) aggregating all measurements into an overall coherence score [23].

In this study, coherence score of c_v measure is calculated to determine the optimal number, using Gensim's CoherenceModel [24]. It can be observed from Figure 3 that the best number of topics for this research is 5.

IV. RESULT

In this section, each topic is explained and interpreted, accompanied by an appendix of example user reviews. LDA result in Table 2 shows 5 extracted topics and 10 keywords of each topic with the corresponding weights which indicates each word's importance to the topic. Moreover, Figure 4 is the word cloud visualization of keywords where more important words are displayed in larger font size. Based on these results along with reviews which have high contribution percentages, we identified the topic names as shown in Table 2.

Furthermore, Figure 5 illustrates an inter-topic distance map generated by using pyLDAvis library, where each topic is each circle and a bigger circle implies a more prevalent topic. This visualization provides an overview of topics and helps us understand their differences based on the distance between topics [25]. According to Figure 5, all extracted topics of Weverse's user reviews are non-

overlapping and scatter throughout the chart, which indicates distinct topics in this study.

Table 2. LDA result

Topics	Keywords	Topic names
Topic 0	problem (0.027), work (0.025), cool (0.019), really (0.017), use (0.017), translation (0.016), fix (0.015), say (0.015), update (0.014), help (0.013)	Network connection error
Topic 1	post (0.077), get (0.065), see (0.039), idol (0.031), comment (0.031), translation (0.026), artist (0.025), reply (0.024), notification (0.020), really (0.019)	Notification & translation issues
Topic 2	update (0.032), helpful (0.022), touch (0.021), always (0.019), keep (0.019), boy (0.018), beautiful (0.018), use (0.017), happy (0.016), nice (0.016)	Update frequency & connection to idols
Topic 3	army (0.073), connect (0.068), thank (0.029), star (0.027), give (0.026), much (0.022), make (0.021), really (0.019), video (0.018), wonderful (0.013)	Light stick connection & gratitude
Topic 4	fan (0.059), idol (0.057), artist (0.049), communicate (0.031), really (0.031), interact (0.031), well (0.021), make (0.019), thank (0.018), way (0.016)	Communication

Topic 0 with the keywords of 'problem', 'work', and 'fix' mainly regards negative feedback of application problems, especially the network connection error. Consulting the high contributing reviews, we found that users cannot get access to Weverse via Wi-Fi but via cellular data. They emphasized that this problem needed to be resolved in an upcoming update. Furthermore, the word 'translation' implied the poor quality of subtitles in music videos and live streams. However, with the word 'cool', users generally expressed their compliments on the app where they said Weverse was a cool app for fans.



Figure 4. Word cloud visualization

Figure 5. Intertopic distance map

Topic 1 mentioned both users' negative and positive expressions. It showed that users enjoyed this app but it did not send real-time notifications so they usually missed their favorite artists' new live streams, posts, and comments. Moreover, they cannot

properly understand the content of replayed live videos because of the incorrect and late subtitles.

Topic 2 revealed users' complaints about frequent app updates with the keyword 'update'. Users stated that they were required to update Weverse at short intervals. Whereas 'helpful' and 'touch' refers to compliments on the connection to their idols. They appreciated the app for helping them continuously stay in touch with their idols.

Topic 3 implied users' gratitude when 'thank' is one of the top keywords. Meanwhile other keywords such as 'army' and 'connect' with investigation on reviews led us to the conclusion that users had difficulties connecting to the Army—bomb known as BTS's light stick when watching their music or concert videos. In addition, they expressed their desire for the availability of other groups' light sticks.

Topic 4 including two top keywords 'fan' and 'idol' mainly mentioned positive reviews. Users expressed their satisfaction with Weverse, emphasizing that they could not only interact with their favorite idols but also communicate with other fans on this platform. Weverse emerged as a highly recommended app for meaningful interaction with artists and mutual relationships among fans.

V. CONCLUSION

This study aims to explore the user experience and app issues of Weverse,

a global fan community platform. To achieve this goal, we employ LDA topic modeling to analyze online reviews of Weverse, retrieved from the Google Play Store. Gaining insight into user feedback and identifying common topics contributes to the app advancement and higher user engagement. Notably, our findings provide valuable suggestions for improving app functionality.

As a result, we discovered 5 latent topics of which Topic 4 mostly focused on positive opinions and the rest of topics mentioned both negative and positive aspects. Specifically, Topic 0 emphasized network connection error. Besides, Topic 1 discussed delayed notification and translation issues. Whereas, Topic 2 uncovered app update's high frequency and facilitation of connection to idols. Further, Topic 3 showed users' gratitude as well as light stick connection issue. Finally, Topic 4 emphasized not only interaction between fans and idols but also fans' relationships.

Overall, the topic modeling results showed that users expressed their appreciation and gratitude towards Weverse. This app played an important role in bridging the gap between fans and idols, and among fans which was consistent with the studies' results by Hong and Kim; and by Kim et al. [11,12].

In fact, Weverse significantly contributes to fan engagement by offering exclusive paid entertainment content as well as artists' live streams and features such as connecting to a

light stick, which allows users to synchronize its light color and patterns with the music. This can enhance audiences' experience, especially for those unable to join the offline concert.

However, our uncovered topics revealed that users faced difficulties pairing their light sticks in the app and suggested offering other group's light sticks available. They expressed their dissatisfaction about the network connection error, which was also explored by Hong and Kim [11].

Weverse is a digital community where fans can easily and quickly get updates on their artists but they encountered delayed notifications, which was also derived in the existing research [12]. Moreover, users have been experiencing poor-quality translation and delayed subtitles in replayed live streams. It can be said that with many artists joining Weverse, its content, and features, this platform has been attracting a lot of users. However, there are several existing problems mentioned above which the development team needs to fix to enhance user experience.

The team should ensure a more stable network connection for users accessing Weverse via Wi-Fi. Moreover, it is crucial to improve the notification system to punctually provide users with the latest news. It is additionally expected to balance the frequency of updates. The team should consider notifying users about the availability of new updates with app improvement highlights and providing them with the options of whether to

update the app or not. Besides, improving the translation mechanism in this app is extremely necessary. In the context of globalization of Korean culture, the number of international fans has been dramatically increasing. Therefore, a correct and diverse translation mechanism is essential to foster mutual understanding between Korean artists and international fans who might not speak Korean.

This study provided valuable insights into user feedback towards Weverse and offered some solutions to address the problems faced by users to improve their experience. In the future, sentiment analysis, users' emotions, comparisons of fan community applications should be conducted to gain a more detailed understanding of fandom platform services.

REFERENCES

- [1] Kim, M. S., & Kim, H. M, "The effect of online fan community attributes on the loyalty and cooperation of fan community members: The moderating role of connect hours," *Computers in Human Behavior*, vol. 68, pp.232–243, Mar. 2017.
- [2] Hallyu (Korean Wave). <https://www.korea.net/AboutKorea/Culture-and-the-Arts/Hallyu> (accessed Feb., 10, 2024).
- [3] About the Company. <https://en.weverse.co/about> (accessed Feb., 10, 2024).
- [4] Weverse. <https://play.google.com/store/apps/details?id=co.benx.weverse> (accessed Feb., 10, 2024).
- [5] Weverse service. <https://en.weverse.co/weverse> (accessed Feb., 10, 2024).
- [6] 위버스, 구글플레이 '2023 올해를 빛낸 인기 앱' 후보 올라 (2023). https://hybecorp.com/kor/news/news/4157?companyCode=COM_BENX&page=0 (accessed Feb., 10, 2024).
- [7] Fandom. <https://dictionary.cambridge.org/dictionary/english/fandom> (accessed Feb., 10, 2024).
- [8] Kim, M. S., Wang, S., & Kim, S, "Effects of Online Fan Community Interactions on Well-Being and Sense of Virtual Community," *Behavioral Sciences*, vol. 13, no. 11, pp. 897, 2023.
- [9] Ham, M., & Lee, S. W, "Factors affecting the popularity of video content on live-streaming services: focusing on V Live, the South Korean live-streaming service," *Sustainability*, vol. 12, no. 5, pp. 1784, 2020.
- [10] K-pop fan service Universe to shut down, transfer to Dear U Bubble (2023). <https://koreajoongangdaily.joins.com/2023/01/11/entertainment/kpop/Universe-NCSOFT/20230111105854122.html> (accessed Feb., 10, 2024).
- [11] Hong, S. Y., & Kim, S. I, "A Study on User Experience Design of Fandom Platform Applications in the Media Entertainment Industry: Focused on Weverse and Universe," *Journal of Digital Art Engineering and Multimedia*, vol. 9, no. 4, pp. 451, 2022.
- [12] Hyun-Su Kim, Nam-Mi Kwak, & Da-Yeon Kim, "A Study on Entertainment Platform Service Evaluation Using IPA (Importance-Performance Analysis): Focusing on Weverse," *Journal of the Korea Entertainment Industry Association*, vol. 16, no. 7, pp. 1–17, 2022.
- [13] 장혜리, "A Study on the Fan Community Platform that supports Artists in the Untact Era," *Master's Thesis, 상명대학교*, 2021.
- [14] Machado Pereira, S, "Fan engagement strategies in the K-pop industry," *Doctoral dissertation, Scuola universitaria professionale della Svizzera italiana*, 2022.
- [15] Song, M, "A Study on the Business Model of a Fan Community Platform

- 'Weverse'," *The International Journal of Advanced Smart Convergence*, vol. 10, no. 4, pp. 172–182, 2021.
- [16] Kyung-Yur Lee, & Sang-Hyeon Park, "Social Big Data Analysis of Perceptions and Issues of Hallyu Fandom," *Journal of the Korea Entertainment Industry Association*, vol. 17, no. 4, pp. 1–16, 2023.
- [17] Ko, H. K, "An analysis of YouTube comments on BTS using text mining," *The Rhizomatic Revolution Review [20130613]*, vol. 1, pp. 1–10, 2020.
- [18] Google-Play-Scraper (2024). <https://pypi.org/project/google-play-scraper/> (accessed Feb., 10, 2024).
- [19] What is Gensim? (2009). <https://radimrehurek.com/gensim/intro.html> (accessed Feb., 10, 2024).
- [20] Hasan, M., Rahman, A., Karim, M. R., Khan, M. S. I., & Islam, M. J, "Normalized approach to find optimal number of topics in Latent Dirichlet Allocation (LDA)," *In Proceedings of International Conference on Trends in Computational and Cognitive Engineering: Proceedings of TCCE 2020*, pp. 341–354, 2021.
- [21] Qomariyah, S., Iriawan, N., & Fithriasari, K, "Topic modeling twitter data using latent dirichlet allocation and latent semantic analysis," *In AIP conference proceedings*, vol. 2194, no. 1, Dec.2019.
- [22] Röder, M., Both, A. & Hinneburg, A, "Exploring the Space of Topic Coherence Measures," *In Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, pp. 399–408, Feb. 2015.
- [23] S. Syed and M. Spruit, "Full-Text or Abstract? Examining Topic Coherence Scores Using Latent Dirichlet Allocation," *2017 IEEE International Conference on Data Science and Advanced Analytics*, pp. 165–174, 2017.
- [24] Topic coherence pipeline (2009). <https://radimrehurek.com/gensim/models/coherencemodel.html#gensim.models.coherencemodel.CoherenceModel> (accessed Feb., 10, 2024).

- [25] Sievert, C., & Shirley, K., "LDAvis: A method for visualizing and interpreting topics," *In Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, pp. 63–70, 2014.

APPENDIX

1. Topic 0 (Network connection error)

"I gave my last review on 27 july, it was one star with the prb of weverse not opening for me at all... its been almost 3 months now, but now the prb is that the app only works properly when I'm using my celular data, it shows "try again later" when i use my wifi, my wifi works perfectly fine thou. I can't use my data always for watching the lives, it takes way too much data. I see so many people having the same prb, but its still not fixed. Are the developers sleeping or smth? FIX IT QUICKLY!!!"

"Hi Team, Dont know what happened, weverse is not working for me Showing not connected to network. Please resolve this issue ASAP The app used to work really good but dont know why recently it is not opening and showing " not connected to internet" error even when I use high speed internet. Could you please check and resolve the issue. I am missing lot of updates recently. Hope it works soon"

2. Topic 1 (Notification & translation issues)

"Its a very great app but I am giving 3 stars bcz the translation of english needs a better check and also sometimes i couldn't receive notifications of the artists posts and comments and its a very bad thing plz look forward to it"

"Your notification is worst yesterday Le Sserafim did live but your notification didn't popping up on my phone i missed i always had to check every hour plz fix it we didn't getting any notification about artist post and live notification is not popping up"

3. Topic 2 (Update frequency & connection to idols)

"Why is this to update again.? I've done doing.this at least 3 times everytime i use my load for this. It makes me sick."

"So nice app, I am so happy this app is fantastic, useful, helpful and everything. Very good app"

4. Topic 3 (Light stick connection & gratitude)

"The new update doesn't give you the option to connect an army bomb to any of the music videos (at least not from a Samsung phone). It also doesn't give you the option to put subtitles on the weverse lives."

"user friendly and efficient to use. thank you for creating a platform that brings BTS and ARMY together"

5. Topic 4 (Communication)

"I like this so much!! I have a lot of fun using it, i can communicate with the other fans and also to the idols, and i can also share my thoughts and all i want to share to the others. I only like to add, Weverse can you add the "Edit" in the comment too?. Thank You so much. It's recommended to the all fans in the world!!"

"As a k-pop fan , this app provides all the necessary information and updates about the idols , we can interact our favourite idols , groups , watch them live and much more , i really enjoy weverse.... lot's of love and keep the service up weverse"



Mi Jin Noh

She received her M.S. and Ph.D. degree in Management Information Systems from Kyungpook National University, Korea in 2001 and 2006, respectively. Since 2022, she has been an assistant professor in Department of Business Big Data, Keimyung University, Korea. Her research interests are Big Data Analysis, Text Mining and Mobile Services.



Yu Na Lee

She received her B.S. and M.S. degree in Management Information Systems from Keimyung University, Korea, in 2021 and 2024. Her research interests include Big Data analysis, Data Visualization and Text Mining.



Yang Sok Kim

He has been serving as an Associate Professor at the department of Management Information Systems, Keimyung University, South Korea. He received his Ph.D. from University of Tasmania, Australia. His research interests are Machine Learning, Web Search/Mining, Social Network, and Recommenders Systems

Authors



Thi Thao Van Ho

She received a bachelor's degree in International Economics from Foreign Trade University, Vietnam, in 2020. She is currently pursuing her M.S studies at the Department of Management Information Systems in Keimyung University, South Korea.