# Development of Real-time Soccer Match Comment Sentiment Analysis and Emoji Conversion System

Ye-Hyun Kim<sup>a</sup>, Jungwon Cho<sup>a</sup>, Jaechoon Jo<sup>a,\*</sup>

<sup>a</sup> Department of Computer Education, Jeju National University, Jeju, Republic of Korea Corresponding author: \*jjo@jejunu.ac.kr

*Abstract*—This study proposes an innovative system designed to automatically analyze emotions in real-time comments during live sports broadcasts, particularly soccer matches, and convert them into appropriate emojis. This approach aims to overcome the limitations of viewer interaction and enable seamless emotional communication across language barriers. The system utilizes the KoBERT model, optimized for Korean natural language processing, for accurate text-based sentiment analysis. The real-time emoji conversion and display functionality is implemented using the React framework, and web socket technology is employed to achieve low-latency data processing in real-time. The model was trained using a large Korean conversation dataset and achieved an emotion classification accuracy of 71.12%. In terms of performance, the system can process 308.71 comments per second with an average latency of 37.52 milliseconds, proving its effectiveness in a live sports broadcast environment. The proposed system enhances the viewing experience by allowing users to express emotions intuitively, thus breaking the limitations of text-based communication. This system introduces a new paradigm for audience interaction in live sports broadcasts, promoting a more inclusive and engaging experience. Using emojis, which transcend language barriers, viewers can share their emotions without text, fostering real-time emotional exchange among a global audience. The research highlights the practical application of real-time text and sentiment analysis technologies. It provides a foundation for future enhancements, such as support for multiple languages and more advanced real-time data processing capabilities.

Keywords-Sentiment analysis; KoBERT; emoji conversion; sports broadcasting.

Manuscript received 11 Nov. 2023; revised 9 Mar. 2024; accepted 12 Sep. 2024. Date of publication 31 Dec. 2024. IJASEIT is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



## I. INTRODUCTION

Live sports broadcasting, especially soccer matches, has a massive global audience and is crucial in modern entertainment [1]. Today's sports viewers have evolved from passive spectators to active participants, with a strong desire to express and share their emotions and thoughts in real time during matches. This is primarily realized through real-time chat or comment features provided by social media platforms[2], and such interactions enrich the sports experience and strengthen the bond within fan communities[3], [4]. However, this text-centric communication method has several limitations [5], [6]. For example, information overload due to rapidly scrolling comments, restrictions on global communication due to language barriers, and the limitations of expressing subtle emotions through text alone [7]. These constraints highlight the need for more effective and inclusive real-time interaction methods [8], [9]. Many researchers have attempted various approaches to overcome the limitations of text-based communication. For instance, techniques have been proposed to automatically classify and visualize the sentiment of comments using sentiment analysis technology. While this approach effectively grasps overall emotional trends from large volumes of text data, it still has limitations in the immediate expression and sharing of individual users' emotions. Another approach involves using emoticons or stickers for emotional expression. Although this allows for more intuitive and rapid emotional expression than text, it is limited in real-time capability and automation as users must manually select them. This study proposes a method of emotional expression through emojis to address these issues [10], [11]. Emojis serve as tools to complement subtle emotions and non-verbal signals that are difficult to convey through text, helping viewers express their feelings more intuitively and quickly [12]. Emojis play a crucial role in overcoming the limitations of text-based communication as a means for global viewers to share emotions without language barriers [13].

Recent studies have proposed sentiment analysis systems for sports events using social media data, demonstrating the ability to track viewers' reactions [14] effectively. While these studies have shown the possibility of processing large-scale data, they still need to resolve the limitations of text-based communication entirely [15]. The development of text-based sentiment analysis technology is also noteworthy [16]. Studies leveraging BERT models have reported significant improvements in sentiment analysis performance [17]. These research results influenced the selection of the KoBERT model in this study[18]. Research on sentiment analysis using emojis is also actively progressing[19], [20]. Studies have evaluated online emoji description resources and analyzed emotional expression using text and emojis [21], [22]. These studies suggest that emojis can play an essential role in improving the accuracy of sentiment analysis [23], [24]. Significant progress has also been made in real-time data processing and visualization [25]. These studies have presented new frameworks for processing and visualizing large-scale real-time data streams, providing important references for designing the real-time processing capabilities of this system [26], [27].

Based on the advancements in these previous studies, this research proposes an innovative system that automatically analyzes real-time soccer match comments and converts them into appropriate emojis. This system performs text-based sentiment analysis using the KoBERT model optimized for Korean natural language processing and implements real-time emoji conversion and display functions using the React framework [28], [29]. Real-time data processing is realized through WebSocket technology[30], and the viewers' visual experience is enhanced through animation effects. The primary purpose of this study is to improve the experience for real-time soccer match viewers. The goal of utilizing emojis instead of text is to provide a more immersive viewing experience by allowing viewers to effectively express their emotions without obscuring the game screen [31].

Additionally, we aim to improve performance in Korean natural language processing by advancing real-time sentiment analysis technology using the KoBERT model optimized for Korean text. Furthermore, we propose a new interaction method that allows viewers to communicate and share emotions in real time during sports broadcasts, presenting an innovative communication method that goes beyond the limitations of traditional comment systems. Lastly, we aim to create an environment where viewers of various nationalities can easily share emotions in real time through a languageindependent communication method using emojis, promoting global communication.

#### II. MATERIALS AND METHOD

This section will detail the data, models, and technical specifications for implementing the proposed system.

#### A. Data Collection and Preprocessing

In this study, we used a Korean dataset provided by AIHUB. This dataset includes various Korean conversations in different contexts, making it suitable for learning diverse emotional expressions that may occur during real-time soccer matches. A total of 38,594 sentence-emotion pairs were

collected, with each data point consisting of a Korean sentence and its corresponding emotion label.

E

Neutral

Disgust

Happiness

TABLE I					
DISTRIBUTION OF EMOTION CLASSES IN THE DATASET					
Emotion	Count (Percentage)				
Fear	5468 (14.17%)				
Surprise	5898 (15.28%)				
Anger	5665 (14.68%)				
Sadness	5267 (13.65%)				

4830 (12.51%)

6037 (15.64%)

5429 (14.07%)

Fig. 1 shows the data preprocessing steps for training the KoBERT model. We excluded 'fear' and 'disgust' from the original dataset and used five emotion classes: 'happiness,' 'sadness,' 'anger,' 'surprise,' and 'neutral.' In the preprocessing stage, we cleansed the data by removing special characters, emojis, and duplicate sentences.



Fig. 1 Data preprocessing steps for training the KoBERT model.

Regular expressions were used to remove special characters that did not convey meaning in the sentences, such as excessive punctuation (e.g., repeated exclamation marks or question marks). All emojis included in the original data were removed to ensure that the model performed sentiment analysis solely based on the text. Duplicate sentences were

merged to maintain diversity in the dataset. The sentences were tokenized using KoBERT's tokenizer, which splits the sentences into meaningful units while accounting for the unique characteristics of the Korean language. Padding was applied to standardize the length of all input sentences, making them suitable for input into the model. A maximum sentence length was set during this process, and shorter sentences were padded with special tokens to match the required length. The entire dataset was randomly split into training data (80%), validation data (10%), and test data (10%). This splits method, commonly used for model training and evaluation, involves training the model with the training data, monitoring its performance with the validation data, and finally assessing its generalization performance using the test data.

## B. KoBERT Model

For sentiment analysis, we used KoBERT, a pre-trained model optimized for the Korean language in Fig. 2.



Fig. 2 Structure of the sentiment analysis model based on KoBERT.

KoBERT is a Korean version of BERT, pre-trained on a large-scale Korean corpus to effectively reflect the linguistic characteristics of Korean. KoBERT follows the basic structure of BERT, consisting of 12 transformer layers, a 768dimensional hidden layer, and 12 attention heads. KoBERT is designed to accurately reflect the unique features of the Korean language, such as its agglutinative nature, by performing morpheme-level tokenization and learning the grammatical structure and word order of Korean for better contextual understanding. It also handles various forms of Korean expression, such as honorifics and casual speech, leading to high performance in processing diverse Korean texts. We added a fully connected layer for emotion classification on top of KoBERT's output layer. This layer receives 768-dimensional input and outputs probabilities for the five emotion classes: 'happiness,' 'sadness,' 'anger,' 'surprise,' and 'neutral.' The fully connected layer maps the high-dimensional features extracted by KoBERT to the final emotion classification result. We used Categorical Cross-Entropy as the loss function, which is suitable for multi-class classification problems and effectively measures the difference between predicted and actual labels. The Adam optimizer, with a learning rate of 3e-5, was used for optimization. Adam automatically adjusts the learning rate during training, enabling efficient learning. The batch size was set to 64, and the model was trained over five epochs. Both batch size and epoch number were selected based on experiments that yielded optimal performance. During training, we applied early stopping to prevent overfitting. This technique halts training when performance on the validation dataset no longer improves, helping to enhance the model's generalization performance.

#### C. Real-Time Emoji Transformation System

In this study, we implemented a system that analyzes realtime input comments and converts them into emojis. WebSocket technology was used to implement real-time communication, and the system is composed of a front end and a back end. For the backend, we used Python Flask to build a RESTful API server and Flask-SocketIO to implement real-time bidirectional communication. The trained KoBERT model was integrated into the Flask server to perform sentiment analysis on real-time input comments. On the frontend, React was used to create a dynamic and responsive user interface. The Socket.IO client library was used to implement real-time communication with the server, and a logic was created to select and render appropriate emojis based on the analyzed emotions.

To map emotions to emojis, we defined mapping rules for converting analyzed emotions into suitable emojis. For example, 'happiness' is mapped to 🕲, 'anger' to 😟, 'surprise' to (), 'sadness' to (), and 'fear' to (). These mappings were selected considering the system's requirements and the cultural context of the target users. To enhance visual effects, we implemented an animation where emojis rise from the bottom to the top of the screen. CSS keyframe animations and styled components were used to implement this. The animation's starting point is set below the bottom of the screen, and the endpoint is set above the top of the screen, giving the effect of the emoji floating upward across the screen. The component displaying the emoji is set in an absolute position to ensure it moves freely without overlapping with other elements. The size and position of the emoji are dynamically determined, and responsive design is implemented using rem units and percentages. The duration of the animation is set to 6 seconds, creating a slow upward motion, with the ease-out timing function applied for a natural movement.

We applied various optimization techniques to ensure the system's stability and performance. Emoji lifespan management was implemented to automatically remove each emoji after 5 seconds, which helps optimize memory usage and manage screen complexity. Data buffering was used to control the amount of data transmitted from the server to the client, with the server set to send a maximum of 10 new

emojis every 100 milliseconds. Additionally, rendering optimizations were performed using React's optimization features to prevent unnecessary re-rendering.

## D. System Integration and Deployment

We used various technologies and tools to integrate the components of the developed system and deploy it in a real environment. First, we packaged the backend server and frontend application into Docker containers. Using Docker ensured consistency between the development and production environments and simplified the deployment process while enhancing scalability. Each container includes all necessary dependencies, allowing consistent execution across different environments. NGINX was used to set up a load-balancing system that distributes traffic to multiple backend instances, enhancing the system's stability and scalability. We utilized NGINX's reverse proxy feature to route client requests to the appropriate backend server and return the server's response to the client. This structure improves security by preventing direct exposure of backend servers. MongoDB was used to store and manage user data and logs. MongoDB's document-oriented nature is suitable for handling unstructured data, allowing it to efficiently store and search large volumes of real-time comment data. We used Prometheus and Grafana to monitor system performance and resource usage in real time. Prometheus, a time-series database, was configured to collect and store various system metrics, such as CPU usage, memory consumption, and network traffic, as well as application-level metrics, such as the number of comments processed per second and average response time. Grafana was used to visualize these metrics and create dashboards, enabling realtime monitoring of the system's status and allowing us to identify performance bottlenecks or anomalies quickly. Jenkins was used to set up a continuous integration and deployment (CI/CD) pipeline, automating the process of testing and deploying code changes. The Jenkins pipeline automatically starts the build process when it detects changes in the code repository. During this process, unit tests and integration tests are performed to verify the quality of the code, and only code that passes these tests is deployed to the production environment.

To ensure system scalability, we partially adopted a microservices architecture. Independent services were created for the sentiment analysis, emoji transformation, and user management modules. This architecture allows for each service's independent scaling and maintenance, enhancing the system's overall flexibility. The developed system was successfully deployed in a stable, scalable, and secure form in a real operating environment through this comprehensive approach. All system components are tightly integrated, providing the high performance and stability required for real-time soccer broadcast environments.

#### III. RESULTS AND DISCUSSION

This chapter presents the performance evaluation results of the implemented system and conducts an in-depth analysis and discussion.

#### A. Document Model Performance Evaluation

To evaluate the performance of the KoBERT model in emotion classification, we measured the accuracy, precision, recall, and F1 score. The evaluation was conducted using the test dataset, and the results are as follows:

TABLE II					
MODEL PERFORMANCE EVALUATION RESULTS					
Metric	Score (%)				
Accuracy	71.12%				
Precision	71.25%				
Recall	71.12%				
F1 Score	71.15%				

These results suggest that the system performs at a practical level considering the real-time processing requirements. Given the complexity of classifying various emotions, achieving an accuracy above 70% is significant. The performance analysis for each emotion class is as follows:

TABLE III CLASSIFICATION PERFORMANCE BY EMOTION CLASS					
Emotion	Precision	Recall	F1 Score		
Happiness	0.63	0.68	0.66		
Sadness	0.76	0.71	0.73		
Anger	0.78	0.77	0.78		
Surprise	0.59	0.58	0.58		
Neutral	0.82	0.83	0.82		

The analysis shows that the model performed best in classifying neutral emotions (F1 Score: 0.82), while the performance was lowest for surprise (F1 Score: 0.58). This may be because neutral expressions tend to have more distinguishable features, making them easier to identify. In contrast, surprise emotions may have characteristics that overlap with other emotions, making them harder to classify accurately. The changes in loss and accuracy during the training process are as follows:



Fig. 3 Changes in loss and accuracy during training.

As shown in Fig. 3, the loss continuously decreases as training progresses, and accuracy shows an upward trend. Performance sharply improved during the early epochs, followed by gradual improvements. The training accuracy reached 88.40% by the fifth epoch, indicating that the model achieved a high fit to the training data.

## B. Comparison with Previous Studies

To evaluate the performance of the proposed KoBERT model, we compared it with two key previous studies that used similar approaches. Both studies proposed sentiment analysis systems based on BERT models and performed sentiment analysis on Korean text, similar to this study. They also addressed the challenges of classifying multiple emotion categories rather than simple positive/negative binary classifications. The comparison of the KoBERT model with models from previous studies is as follows:

TABLE IV Emotion Classification Performance by Class					
Emotion	This Study	SMERT Model [32]	Portal comment sentiment analysis model [33]		
Happiness	0.66	0.675	N/A		
Sadness	0.73	0.696	0.79		
Anger	0.78	0.777	0.58		
Surprise	0.58	0.853	0.65		
Neutral	0.82	N/A	N/A		

In happiness classification, the KoBERT model performed similarly to the SMERT model. For sadness and anger, the KoBERT model performed similarly or better than the SMERT model, and it outperformed the BERT-based model in these categories. Particularly in detecting anger, the KoBERT model showed outstanding performance. However, the SMERT model significantly outperformed KoBERT in surprise classification, likely due to SMERT's multimodal approach, which better captures this emotion. For neutral emotions, only the KoBERT model achieved high performance.

## C. System Performance Evaluation

To evaluate the real-time processing capabilities of the system, we conducted experiments measuring throughput and latency. The system was able to process an average of 308.71 comments per second, which is sufficient to handle the maximum number of comments that may occur during a live soccer broadcast. The average latency was measured at 37.52ms, allowing users to perceive the system's response in real-time, thus demonstrating its suitability for real-time sentiment analysis and emoji transformation. These results show the system can operate effectively in real-time soccer broadcast environments, ensuring stable performance even during large-scale live events with high comment traffic.

The image below illustrates the system in action during a live soccer match. The system detects emotions from the comments on the left side of the screen and converts them into corresponding emojis displayed over the video in real-time. Each emoji represents the emotion detected in the comment and rises upwards in an animated manner, providing an intuitive and visually engaging way for users to share and perceive emotions without obstructing the view of the match.



Fig. 4 Real-Time Soccer Match Comment Sentiment Analysis and Emoji Transformation System in Action

This visual representation helps demonstrate how the system integrates real-time sentiment analysis with emoji

conversion to enhance user interaction during live broadcasts. The system's ability to analyze and transform comments into emojis in real-time, coupled with its smooth animation effects, creates a more immersive and emotionally engaging viewing experience for the audience. Furthermore, this real-time emoji transformation system enhances cross-language emotional communication, as users from different linguistic backgrounds can quickly understand the emotions conveyed by others without needing to interpret the language of the comment. The image shows how the system bridges the communication gap between viewers, providing a seamless and inclusive environment for sharing emotions during live sports events.

## D. Discussion and Limitations

The KoBERT-based system proposed in this study successfully analyzes the emotions of real-time soccer match comments and converts them into emojis. The comparison with previous studies highlights both the strengths and limitations of the system. First, the system excelled in detecting anger and neutral emotions, suggesting its effectiveness in identifying negative comments and classifying neutral opinions in a real-time soccer broadcast environment. Moreover, the system achieved competitive performance without relying on external emotion lexicons.

However, there are several limitations to the current system. First, the relatively low performance in detecting surprise emotions (F1 Score: 0.58) needs improvement. It may be necessary to improve the model's architecture to capture better features related to surprise emotions. Second, while the overall accuracy (71.12%) is practical, further optimization of the model structure and acquiring more training data may help achieve higher accuracy. Currently, the model analyzes emotions on a sentence-by-sentence basis, which limits its ability to capture emotions in context accurately. To address this, we plan to introduce LSTM or Transformer-based models that consider the context for more precise sentiment analysis. Verifying the system in real-world environments is another important task. As the current performance evaluation is based on a limited dataset, largescale user studies in real-time soccer broadcasts are needed to validate the system's effectiveness and identify potential issues. Finally, the current system only supports Korean, and multilingual support is necessary for application in global sports events. Future works include expanding the system to process comments in multiple languages using multilingual BERT models.

To overcome these limitations and improve the system, future research will involve applying the system to real soccer broadcasts and conducting large-scale user studies. In addition, various state-of-the-art natural language processing technologies will be applied to enhance the accuracy of sentiment analysis and further strengthen the system's scalability and stability.

#### IV. CONCLUSION

In this study, we proposed and implemented an innovative system that analyzes the emotions of real-time soccer match comments and converts them into emojis. The sentiment analysis using the KoBERT model and the real-time emoji transformation system based on React demonstrated high accuracy and processing speed, proving its practicality for use in real sports broadcasting environments. The main objectives of this research were to enhance the viewer experience, advance real-time sentiment analysis technology, introduce a new interaction paradigm, and promote global communication. Using emojis instead of text allows users to express emotions effectively without obstructing the game screen. It advances real-time sentiment analysis technology for Korean text, contributing to the field of natural language processing.

Moreover, we proposed a new way for viewers to easily communicate and share emotions during sports broadcasts, enabling people from different countries to share emotions using emojis, regardless of language barriers. This system presents a new interaction paradigm where viewers can be more immersed in the match while sharing their feelings effectively. Using emojis, the system transcends the limitations of text-based communication, promoting emotional exchanges between global viewers.

In future research, we plan to collect much comment data generated during live soccer broadcasts to build a new dataset that more accurately captures the unique language patterns and emotional expressions specific to soccer matches. Using this dataset, we will retrain the model to improve the accuracy of sentiment analysis and develop an emoji transformation system more suitable for soccer match contexts. This research demonstrates the potential for artificial intelligence and realtime data processing technologies to innovatively enhance the sports viewing experience, achieving technical success and presenting new possibilities for emotional exchanges and communication through sports. We hope this technology will be applied to sports broadcasting and various real-time events and online communication platforms, providing more affluent and more intuitive channels for emotional interaction.

#### ACKNOWLEDGMENT

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea(NRF-2021S1A5C2A04088646)

#### References

- S. Sharma and V. Gupta, "Rio Olympics 2016 on Twitter: A Descriptive Analysis," *Computational Methods and Data Engineering*, pp. 151–162, Nov. 2020, doi: 10.1007/978-981-15-7907-3\_12.
- [2] C. Guo, S. Lin, Z. Huang, and Y. Yao, "Analysis of sentiment changes in online messages of depression patients before and during the COVID-19 epidemic based on BERT+BiLSTM," *Health Information Science and Systems*, vol. 10, no. 1, Jul. 2022, doi: 10.1007/s13755-022-00184-w.
- [3] B. Hutchins and J. Sanderson, "The primacy of sports television: Olympic media, social networking services, and multi-screen viewing during the Rio 2016 games," *Media International Australia*, vol. 164, no. 1, pp. 32–43, May 2017, doi: 10.1177/1329878x17707065.
- [4] L. Vale and T. Fernandes, "Social media and sports: driving fan engagement with football clubs on Facebook," *Contemporary Issues* in Sports Marketing, pp. 37–55, Dec. 2024, doi:10.4324/9781003590910-4.
- [5] I. Rabbimov, S. Kobilov, and I. Mporas, "Opinion Classification via Word and Emoji Embedding Models with LSTM," *Speech and Computer*, pp. 589–601, 2021, doi: 10.1007/978-3-030-87802-3\_53.
- [6] P. V. Tijare and J. Rani P., "Detecting Trending Event Topics Using Sentiment Driven Derivatives Method on Twitter," *Indian Journal of Computer Science and Engineering*, vol. 12, no. 4, pp. 818–826, Aug. 2021, doi: 10.21817/indjcse/2021/v12i4/211204077.

- [7] K. Park and B. Sharif, "Assessing Perceived Sentiment in Pull Requests with Emoji: Evidence from Tools and Developer Eye Movements," 2021 IEEE/ACM Sixth International Workshop on Emotion Awareness in Software Engineering (SEmotion), pp. 1–6, May 2021, doi: 10.1109/semotion52567.2021.00009.
- [8] G. Alexandridis, J. Aliprantis, K. Michalakis, K. Korovesis, P. Tsantilas, and G. Caridakis, "A Knowledge-Based Deep Learning Architecture for Aspect-Based Sentiment Analysis," *International Journal of Neural Systems*, vol. 31, no. 10, p. 2150046, Aug. 2021, doi:10.1142/s0129065721500465.
- [9] S. P. Ardakani, C. Zhou, X. Wu, Y. Ma, and J. Che, "A Data-driven Affective Text Classification Analysis," 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 199–204, Dec. 2021, doi: 10.1109/icmla52953.2021.00038.
- [10] P. G, V. V. T, K. V. N, B. K, and K. N, "Sentiment Analysis and Emoji Mapping," 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 1973–1976, Mar. 2022, doi: 10.1109/icaccs54159.2022.9785281.
- [11] D. Gimadi, R. Evans, and K. Simov, "Web-sentiment Analysis Of Public Comments (Public Reviews) For Languages With Limited Resources Such As The Kazakh Language," *Proceedings of the Student Research Workshop Associated with RANLP 2021*, pp. 65–68, 2021, doi: 10.26615/issn.2603-2821.2021 010.
- [12] B. Maram, B. S. Kumar, and P. Swaroopni, "Sentimental Analysis with Emojis by Using Machine Learning," *Data, Engineering and Applications*, pp. 583–597, 2022, doi: 10.1007/978-981-19-4687-5\_45.
- [13] S. A. A. Hakami, R. Hendley, and P. Smith, "Arabic Emoji Sentiment Lexicon (Arab-ESL): A Comparison between Arabic and European Emoji Sentiment Lexicons," in *Proceedings of the 6th Arabic Natural Language Processing Workshop (WANLP 2021)*, 2021, pp. 120–125.
- [14] H. Hettiarachchi, D. Al-Turkey, M. Adedoyin-Olowe, J. Bhogal, and M. M. Gaber, "TED-S: Twitter Event Data in Sports and Politics with Aggregated Sentiments," *Data*, vol. 7, no. 7, p. 90, Jun. 2022, doi:10.3390/data7070090.
- [15] N. Alharbi, O. M. Al-Mohammadi, R. A. Al-Rehaili, S. A. Al-Rehaili, and H. A. Al-Harbi, "Sentiment Analysis of Restaurants Customer Reviews on Twitter," *Advances on Intelligent Informatics and Computing*, pp. 295–308, 2022, doi: 10.1007/978-3-030-98741-1 25.
- [16] X. Li, Y. Lei, and S. Ji, "BERT- and BiLSTM-Based Sentiment Analysis of Online Chinese Buzzwords," *Future Internet*, vol. 14, no. 11, p. 332, Nov. 2022, doi: 10.3390/fi14110332.
- [17] M. J. Althobaiti, "BERT-based Approach to Arabic Hate Speech and Offensive Language Detection in Twitter: Exploiting Emojis and Sentiment Analysis," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 5, 2022, doi:10.14569/ijacsa.2022.01305109.
- [18] E. Tavan, A. Rahmati, and M. Ali Keyvanrad, "Persian Emoji Prediction Using Deep Learning and Emoji Embedding," 2020 10th International Conference on Computer and Knowledge Engineering (ICCKE), pp. 350–355, Oct. 2020, doi:10.1109/iccke50421.2020.9303639.
- [19] M. Fernández-Gavilanes, E. Costa-Montenegro, S. García-Méndez, F. J. González-Castaño, and J. Juncal-Martínez, "Evaluation of online emoji description resources for sentiment analysis purposes," *Expert Systems with Applications*, vol. 184, p. 115279, Dec. 2021, doi:10.1016/j.eswa.2021.115279.
- [20] S. Gupta, A. Singh, and J. Ranjan, "Sentiment Analysis: Usage of Text and Emoji for Expressing Sentiments," *Advances in Data and Information Sciences*, pp. 477–486, 2020, doi: 10.1007/978-981-15-0694-9 45.
- [21] C. Liu, X. Tan, T. Zhou, W. Zhang, J. Liu, and X. Lu, "Emoji use in China: popularity patterns and changes due to COVID-19," *Applied Intelligence*, vol. 52, no. 14, pp. 16138–16148, Mar. 2022, doi:10.1007/s10489-022-03195-y.
- [22] T. M. de Barros, H. Pedrini, and Z. Dias, "Data-Augmented Emoji Approach to Sentiment Classification of Tweets," *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, pp. 67–76, 2021, doi: 10.1007/978-3-030-93420-0\_7.
- [23] S. A. A. Hakami, R. Hendley, and P. Smith, "Emoji Sentiment Roles for Sentiment Analysis: A Case Study in Arabic Texts," *Proceedings* of the The Seventh Arabic Natural Language Processing Workshop (WANLP), pp. 346–355, 2022, doi: 10.18653/v1/2022.wanlp-1.32.
- [24] R. K. Singh and H. K. Verma, "Redis-Based Messaging Queue and Cache-Enabled Parallel Processing Social Media Analytics Framework," *The Computer Journal*, vol. 65, no. 4, pp. 843–857, Nov. 2020, doi: 10.1093/comjnl/bxaa114.
- [25] X. Tian, S. Tang, H. Zhu, and D. Xia, "Real-time sentiment analysis of students based on mini-Xception architecture for wisdom

classroom," Concurrency and Computation: Practice and Experience, vol. 34, no. 21, May 2022, doi: 10.1002/cpe.7059.

- [26] R. K. Singh and H. K. Verma, "Redis-Based Messaging Queue and Cache-Enabled Parallel Processing Social Media Analytics Framework," *The Computer Journal*, vol. 65, no. 4, pp. 843–857, Nov. 2020, doi: 10.1093/comjnl/bxaa114.
- [27] H. Yun, and J. Jo, "Research of Media Evaluation Index Generation Model based on Emotion," *The Journal of Korean Association of Computer Education*, vol. 23, no. 5, pp. 71-79, 2020.
- [28] M. Marreddy, S. R. Oota, L. S. Vakada, V. C. Chinni, and R. Mamidi, "Am I a Resource-Poor Language? Data Sets, Embeddings, Models and Analysis for four different NLP Tasks in Telugu Language," ACM Transactions on Asian and Low-Resource Language Information Processing, vol. 22, no. 1, pp. 1–34, Nov. 2022, doi: 10.1145/3531535.
- [29] M. Wook et al., "Opinion mining technique for developing student feedback analysis system using lexicon-based approach (OMFeedback)," *Education and Information Technologies*, vol. 25, no. 4, pp. 2549–2560, Dec. 2019, doi: 10.1007/s10639-019-10073-7.

- [30] S. M. Sarsam, H. Al-Samarraie, N. Bahar, A. S. Shibghatullah, A. Eldenfria, and A. Al-Sa'Di, "Detecting Real-Time Correlated Simultaneous Events in Microblogs: The Case of Men's Olympic Football," *HCI in Games: Experience Design and Game Mechanics*, pp. 368–377, 2021, doi: 10.1007/978-3-030-77277-2\_28.
- [31] G. Aditi, U. Sharma, S. Kumar, and J. S. Jadon, "A Quantitative Performance Evaluation Of Machine Learning Algorithms For Analysing Sentiments Of Emoticons," 2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence), pp. 606–611, Jan. 2022, doi:10.1109/confluence52989.2022.9734201.
- [32] K. Kim, J. Park, J. Lee, and S. Park, "SMERT: Single-stream Multimodal BERT for Sentiment Analysis and Emotion Detection," *Journal of KIISE*, vol. 48, no. 10, pp. 1122–1131, Oct. 2021, doi:10.5626/jok.2021.48.10.1122.
- [33] J. J. Ho, "Emotion trends analysis of user comment on Internet portal portal articles using BERT model," M.S. thesis, Dept. of Computer Science, Sogang Univ., Seoul, Korea, 2023.