International Journal of Advanced Culture Technology Vol.10 No.3 272-280 (2022)
DOI https://doi.org/10.17703/IJACT.2022.10.3.272

IJACT 22-9-32

Sentiment Analysis on Global Events under Pandemic of COVID-19

[1]Zhang Junjun, [2]Giseop Noh

[1]PhD Candidate, Dept. of Computer Information Engineering, Cheongju Univ., Korea
[2]Assistant Prof., Division of Software Convergence, Cheongju Univ., Korea
370622322@qq.com, kafa46@cju.ac.kr

Abstract

During last few years, pandemic of COVID-19 has been a global issue. Under the COVID-19, global events have been restricted or canceled to secure public hygiene and safety. Since one of the largest global events is Olympic Games, we selected recent Olympic Games as our case of analysis. Tokyo Olympic Games (TOG) was held in 2021, but it encountered a millennium disaster, the pandemic of COVID-19. In such a special period, it is of great significance to explore the emotional tendency of global views before and TOG via artificial intelligence. This paper vastly collects the TOG comment data of mainstream websites in South Korea, China, and the United States by implementing crawler program for sentiment analysis (SA). And we use a variety of sentiment analysis models to compare the accuracy of the experimental results, to obtain more reliable SA results. In addition, in the prediction results, to reduce the distortion of opinion by a minority, we introduce an algorithm called "Removing Biased Minority Opinions (RBMO)" and provide how to apply this method to the interpretation domain. Through our method, more authoritative SA results were obtained, which in turn provided a basis for predicting the sentiment tendency of countries around the world in TOG during the COVID-19 epidemic.

Keywords: *Sentiment Analysis, Deep Learning Model, Tokyo Olympic Games, RBMO*

1. INTRODUCTION

The goal of sentiment analysis (SA) is to conduct opinion mining, which belongs to the category of text classification in natural language processing (NLP). In recent years, with the rapid advancement of NLP technology, SA tasks have also been increasingly used on social media, e-commerce, movie reviews, news reviews and other network platforms to dig out people's views and opinions from different angles. There are various global events around world. To review on the events is crucial in that organization of the events need to evaluate their successiveness.

In 2020, a sudden COVID-19 wreaked havoc all over the world, and the Olympic Games (OG) to be held in Tokyo had to be postponed. Even so, the Tokyo Olympic Games (TOG) was held in Tokyo in July 2021 under the voices of opposition and support all over the world. Facing a major disaster that was rare in a millennium, it has brought huge challenges to the organizers the Olympic games (OG). Especially one month before and after the opening of TOG, the TOG in a special period has become the focus of global attention. In this case, the analysis of the emotional would help the International Olympic Committee (IOC) to accurately

Manuscript received: June 02, 2022 / revised: August 01, 2022 / accepted: September 01, 2022
Corresponding Author: kafa46@cju.ac.kr
Tel: +82-43-229-8496, Fax: +82-43-229-8110
Professor, Dept. of Software Convergence, Cheongju Univ., Korea

predict public opinions, effectively guide the public trend, and make efficient responses in the face of major emergencies. To complete these objectives, sentiment analysis (SA) is the one of promising options. This paper facilitates the crawler implemented to gather Internet commentary data regarding Olympic news topics in the three countries (South Korea, China, and the United States). The data collection period is of the 30 days before and after the opening of the TOG.

Through the analysis of the comment data, we can explore the opinions and tendencies of different countries under the GOVID-19 epidemic. However, since the comment data obtained through web crawlers often have grammatical spelling errors, Internet-specific languages, special characters, obscure semantics, and many other data irregularities, as well as large data deviations for different numbers of news online comments. This brings great difficulties to SA.

Moreover, due to the low accuracy of natural language machine translation (NMT) technology before the appearance of transformer architecture [1], the quantitative SA research among multi-countries have been limited. As far we know, there are few research works related to SA in multi-countries. The authors [2] conducted SA research using sentiment word dictionary implementation of [3]. However, SA dictionary-based approach has limitations since the limited number of sentiment words and domain specific characteristics. Like our work, [4] conducted global SA of COVID-19 using Long Short-Term Memory (LSTM) [5] and artificial neural networks (ANN) in Twitter dataset. However, the results of [4] is limited in that they did not test SA in SOTA such as BERT [6] and just uses Twitter dataset which is unofficial information from individual sources (we use more official opinion using news media from multi nations).

In this paper, we gather data and experiment under our approach from TOC news commends in online which could be more official and relevant to target topics than Twitter dataset. Our contributions are summarized as follow:

1. We extensively conduct experiments to find comparative analysis with variety of emotional text classification models (TextCNN, FastText, BERT). After implementation and experiments of those models, this paper validates the accuracy of polarity prediction with the models.
2. To relief the bias from mean score from minority voting, we designed an algorithm named 'Removing Biased Minority Opinions (RBMO)' approach and provide how the approach can be applied in opinion interpretation domain.
3. We gathered data comes from the mainstream news websites of the three countries, which is helpful to obtain more authoritative emotional analysis results, and then understand the emotional tendency of countries around the world to hold TOG in COVID-19 epidemic.

2. RELATED WORK

2.1 Research status of SA methods

At present, there are mainly three common SA methods, namely, the method based on the sentiment dictionary, the machine learning method based on a large-scale corpus, and the currently popular SA method based on deep learning. Method based on emotional dictionary is to construct an emotional dictionary through the analysis of grammar and syntactic rules. The sentiment value is calculated by using the relevant sentiment analysis algorithm to strengthen the judgment of sentiment tendency. This method is relatively simple to implement, but the effectiveness of the analysis depends to a large extent on the chosen sentiment lexicon. If the coverage of sentiment words is not broad enough and the fields involved are narrow, it will seriously affect the quality of the analysis. For example, Keshavarz et al. [9] combined corpus with dictionary, manually added common words as emotional words, and generated a dynamic adaptive emotional dictionary, so as to achieve better emotional polarity classification.SA methods based on machine learning use supervised learning to obtain a model. It was first used by Pang et al. [10] and others to analyze the emotional tendency of film reviews by using machine learning algorithms. The experimental results show that SVM has the best effect in emotional classification. The research work of sentiment classification based on machine learning technology relies heavily on manual design, and at the same time, most machine learning algorithms use a bag-of-words model to realize the vector representation of text. This method will lead to sparse data and cannot extract the emotional information implicit in the text very well.

In recent years, with the rise of deep learning technology, NLP has made up for the shortcomings of machine learning algorithms. The early CNN [12] model and RNN [13-16] series models cannot solve the long-distance dependency problem or are prone to gradient disappearance or gradient explosion and the limitation of information encoding, which makes the SA task unable to achieve ideal results. Attention mechanism introduced by [20] have given great attention from NLP areas and keep developing so far [21]. And the attention mechanism has gradually emerged in text classification. These models have achieved promising results in the task of text emotion analysis in 2010s and made a breakthrough in performance improvement of deep learning-based architecture. With the emergence of transformer [1], the BERT language model was born [6]. It uses mask language model (MLM) and next sentence prediction (NSP) for pre-training, while considering word order and contextual semantic relevance. This enables word vectors to be dynamically updated according to contextual semantic information. BERT achieves the state-of-the-art performance on all NLP tasks.

2.2 Our approach in text classification

In this work, we choose the BERT model to predict the emotional analysis task of multinational users in TOG. At the same time, to verify its effect, we compare the FastText [23] model which is the subword segmented version of Word2Vec [24] provided by Facebook’s AI Research (FAIR). We also implement directly from TextCNN [12] for extensive experiments, and finally facilitate BERT as our final evaluation model for text classification tasks in 30 days before and after the opening of TOG.

In addition, in the prediction results of the model, due to the large difference in the number of comments per day, when calculating the sentiment polarity score, we pay more attention to the average score deviation caused by minority votes. Therefore, in our approach, inspired by Bayesian theory and [25-26], we design an algorithm called "Removing Biased Minority Opinions (RBMO)" and provide a derivation of this algorithm process and how to apply the algorithm to the field of opinion interpretation.

3. SA ARCHITECTURE

In this section, we elaborate how we build our SA Architecture for multi-country language domain. After subsection of data collection, we briefly explain deep learning-based SA architectures used in this paper. In the consecutive subsection, we elaborate how we apply RBMO from rating system to SA domain.

3.1 Data Collection

This paper selects 3 datasets for experiments:

(1) Chinese dataset

Sina News, a more mainstream Chinese website, is used as the target website for crawling TOG data. We crawled two time periods of network comment data on Sina News by implementing a crawler program.

Training dataset: There are a total of 50,000 sports news comment data from January 2020 to June 2020. We select 30,000 sports news comment data to mark as the training data set, with 15,000 positive and negative comments each. After manual labeling, we get 15,000 positives and 15,000 negatives, respectively. Test dataset: Crawling a total of 50,000 online comment data on Olympic theme news content for a total of 60 days from June 21 to August 22, 2021.

(2) Korean dataset

Using a crawler program we implemented, a total of 1.39 million online comment data on TOG news content was crawled from June 21, 2021, to August 22, 2021, on the 'Naver news network' in South Korea. We randomly select 30,000 comments of them training dataset and conducted manual labeling for model training. After labeling we get including 24,000 positive comments and 6,000 negative comments. Then use the rest dataset (1.36 million comments) as the forecast (test) for SA.

(3) English dataset

We implemented a crawler program using APIs provided by New York Times (NYT). A total of 19,000 comments data on TOG content for a total of 60 days from June 21 to August 22, 2021, were crawled on NYT in the United States. We select 2,000 comments among them to label them as the training dataset and make test on the remaining 17,000 dataset. After manual labeling on training dataset, we finally get English dataset containing 1,000 positives and 1000 negatives, respectively.

3.2 Generating the Final Polarity Scores

In this paper, the SA polarity score is generated by statistical analysis of the prediction results. Therefore, it is particularly important to get accurate prediction results. Our method is to predict the positive/negative sentiment of the collected TOG comment data by constructing a stable and reliable SA model, and then use the algorithm formula to calculate the final sentiment polarity scores specific steps are as follows:

- Step1: Training SA model
- Step2: Use SA model to predict all unlabeled TOG comment data.
- Step3 : Compute averaging score for each day. The average positive score on day i : $V_i = C_{pos} / C_i$,The average negative score on day i : $1 - V_i$. where C_{pos} is the number of comments with a predicted value of 1 on day i, and C_i is the number of all comments on the day.
- Step4: Calculation of the final polarity scores
- Step5 : Generate results & visualization

3.3 Removing biased minority opinions (RBMO)

3.3.1 Building RBMO using Bayesian Average

One of our concerns in emotional polarity is distortion of mean value since we calculate the daily polarity of emotion by averaging the number of positive (or negative) emotion for each date. For instance, we may have a question to which date is more positive (or negative) polarity level of emotion in the following cases:

- Emotion polarity in day p: 0.8 with 4 positive emotions out of 5 emotion values.
- Emotion polarity in day q: 0.7 with 70 positive emotions out of 100 emotion values.

If we only consider just averaged score, we may decide such that day p is judged more positively by 0.1 (10%) than day q. However, we know the day q is more confident and higher positive polarity since the value of day q is yielded by 20 times more values. To resolve this phenomenon, inspired by Bayesian theory and [25-26], we elaborate how we derive RBMO equation used in this paper (for brevity, we explain with the case of positive emotion only.) To overcome this phenomenon, we apply our equation (1) using Table 1.

Table 1. Summary of terms regarding RBMO

Terms	Descriptions	Remarks
$\boldsymbol{t}$	Monitoring (Crawling) period	$t = 30$ in this paper
$\boldsymbol{d_i}$	The date of $\mathrm{i-th}$ day, where $-t \leq i \leq t$	d_0: target date If $i > 0$, d_i is after target date, else: d_i is before target date
$\boldsymbol{c_i}$	The number of daily online comments in a day	$\#ObservedEotions$
$\boldsymbol{c_{pos}}$	The number of positive comments in c_i	$\#ObservedPositiveEmotions$
$\boldsymbol{m_G}$	The average number of comments (global mean) during data collection period ($-t \leq i \leq t$)	$\#MinRequiredEmotions$ (hyper-parameter)
$\boldsymbol{m_{pos}}$	The average number of positive comments during data collection period ($-t \leq i \leq t$)	$\#\mathrm{MinRequiredPositiveEmotions}$ (hyper-parameter)

The final form of RBMO is stated in Eq. (12) using our notation in Table 1.

$$RBMO(d_i) = \frac{c_{pos} + m_{pos}}{c_i + m_G}\ ,$$
$$where\ \ m_G = \frac{1}{2t+1}\sum_{i=-t}^{t} c_i \quad and \quad m_{pos} = \frac{1}{2t+1}\sum_{i=-t}^{t} c_{pos} \tag{1}$$

Eq. (1) is interpreted intuitively as follow.

$$RBMO = \frac{\#ObservedPositiveEmotionsInDay + \#MinRequiredPositiveEmotions}{\#ObservedEotionsInDay + \#MinRequiredEmotions}$$

3.3.2 Building RBMO using Bayesian Average

Using our RBMO, now we design the algorithm for yielding emotional polarity scores. We named this algorithm as ‘Balanced Binary Emotional Polarity (BBEP)’. We summarize the BBEP procedures in Table 2.

Table 2. Balanced Binary Emotional Polarity (BBEP) algorithm using RBMO

Algorithm 1. BBEP
Input: $D = \{(date, comment)\}_{i=0}^{2t}$
Output: Two emotional polarity score lists (original polarity, BBEP polarity)

Initialize required variables

$c_i,\ c_{pos} = [\ ], [\ ]$ # Empty lists comments count and positive comment count for each date

$E_o = [\ \ ]$ # Empty list for emotional polarity for each date

$E_b = [\ \ ]$ # Empty list for balanced emotional polarity for each date

$c_i, c_{pos} \leftarrow$ count daily comments & daily positive comments in D

$m_G \leftarrow$ compute average number of comments using Eq. (12)

$m_{pos} \leftarrow$ computer average number of positive comments using Eq. (12)

for each date in D

: $E_o \xleftarrow{\text{append}}$ Compute average emotional polarity of the date

: $E_b \xleftarrow{\text{append}}$ Compute RBMO score of the date using Eq. (12)

end for

Return E_o and E_b

4. EXPERIMENT AND RESULT ANALYSIS

4.1 Experimental setup

For Korean data, Chinese data, and English data, when training, we selected BERT models in different languages. For the Chinese BERT model, choose the full-word coverage Chinese BERT pre-training model Chinese-BERT-wwm jointly released by Harbin Institute of Technology IFLYTEK. Korean BERT chooses the BERT-kor-base model, and English BERT chooses the BERT-base-uncased model. We applied batch size 64, learning rate 1e-5, epoch 10 with AdamW optimizer.

4.2 Analysis of sentiment classification results

Through the above experiment, we got the result of data prediction. To get an intuitive visualization effect, we plot the distribution of the 60-day Final Polarity Scores. In the graphs drawn, we observe and compare the results of SA through two sets of graphs. The first group of graphs: Compare the distribution changes of the polarity scores before and after applying RBMO (Take the positive polarity value as an example. Figure 1).

In Figure 1, the plots whose legend name is suffixed with '_weighted 'is the result of adding the RBMO algorithm. At the same time, to understand the overall situation of positive and negative emotional polarity, we calculated the average value of positive and negative emotional polarity and plotted it in the figure. As can be seen from the figure, after the RBMO algorithm is added, the shape has been fine-tuned, and the original polarity values of the lower/higher ranks are ranked in the front/back positions, which is in line with the RBMO algorithm. Target, the emotional polarity value obtained is more scientific and reasonable.

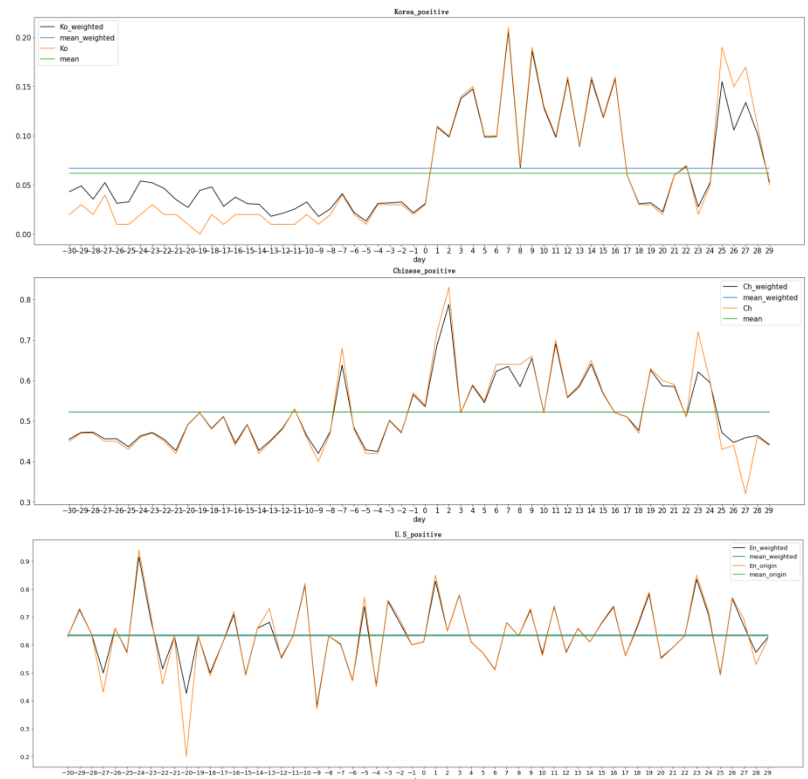

Figure 1. SA plots in Korea, China and U.S.

5. CONCLUSIONS

In this paper, according to the characteristics of online comment data about Olympic news topics in South Korea, China, and the United States obtained through web crawlers, a classification model that can adapt to the task of sentiment analysis is constructed. Through multiple model comparison experiments, the BERT model with stronger feature extraction ability on text classification tasks is selected. It can comprehensively judge the expected sentiment tendency from the word and sentence level, which can significantly improve the performance of sentiment classification. We then derived a so-called “Balanced Binary Emotional Polarity (BBEP) Algorithm” from RBMO and used this algorithm to calculate the final sentimental polarity score. The emotional polarity results obtained by our method are more authoritative and can provide a reliable basis for predicting the emotional tendency of countries around the world to hold TOG during the COVID-19 epidemic. Meanwhile, our proposed RBMO algorithm can be generalized to other opinion interpretation domains. Finally, by comparing the emotional tendencies of netizens in the three countries towards hosting the OG during the GOVID-19 epidemic, people had a negative emotion under pandemic. In future work, we will study how to further improve the model on this basis to improve the accuracy of sentiment analysis.

REFERENCES

[1] A. Vaswani et al., "Attention is all you need," in Advances in neural information processing systems, pp. 5998-6008, 2017.

[2] S.-J. Yea, S. Kim, T. John-Michaël, and J.-G. Lee, "SentiWorld: Understanding Emotions between Countries Based on Tweets," in Tenth International AAAI Conference on Web and Social Media, 2016.

[3] P. S. Dodds, K. D. Harris, I. M. Kloumann, C. A. Bliss, and C. M. Danforth, "Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter," PloS one, vol. 6, no. 12, p. e26752, 2011.

[4] M. Mansoor, K. Gurumurthy, and V. Prasad, "Global Sentiment Analysis Of COVID-19 Tweets Over Time," arXiv preprint arXiv:2010.14234, 2020.

[5] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735-1780, 1997.

[6] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," arXiv preprint arXiv:1810.04805, 2018.

[7] V. S. Subrahmanian and D. Reforgiato, "AVA: Adjective-Verb-Adverb Combinations for Sentiment Analysis," IEEE Intelligent Systems, vol. 23, no. 4, pp. 43-50, 2008.

[8] Z. Gong-rang, B. Chao, W. Xiao-yu, G. Dong-xiao, Y. Xue-jie, and L. Kang, "Sentiment Analysis and Text Data Mining Based on Reviewing Data," Information Science, vol. 39, no. 5, pp. 53-61, May 2021.

[9] H. Keshavarz and M. S. Abadeh, "ALGA: Adaptive lexicon learning using genetic algorithm for sentiment analysis of microblogs," Knowledge-Based Systems, vol. 122, pp. 1-16, April 2017.

[10] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," arXiv preprint cs/0205070, 2002.

[11] J. Zhang, C. Zhao, F. Xu, and P. Zhang, "SVM-Based Sentiment Analysis Algorithm of Chinese Microblog Under Complex Sentence Pattern," International Conference in Communications, Signal Processing, and Systems Springer, pp. 801-809, Singapore, 2018.

[12] Y. Kim, "Convolutional Neural Networks for Sentence Classification," Doha. Association for Computational Linguistics, in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1746-1751, Oct 2014.

[13] P. Liu, X. Qiu, and X. Huang, "Recurrent neural network for text classification with multi-task learning," arXiv preprint arXiv:1605.05101, 2016.

[14] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," arXiv preprint arXiv:1406.1078, 2014.

[15] G. Xu, Y. Meng, X. Qiu, Z. Yu, and X. Wu, "Sentiment Analysis of Comment Texts Based on BiLSTM," IEEE Access, vol. 7, pp. 51522-51532, 2019.

[16] Z. Cui, J. Zhang, G.Noh, HJ.Park, "MFDGCN: Multi-Stage Spatio-Temporal Fusion Diffusion Graph Convolutional Network for Traffic Prediction," Applied Sciences, Vol. 12, No. 2, PP. 2688,March 2022.

[17] L. Guo, D. Zhang, L. Wang, H. Wang, and B. Cui, "CRAN: a hybrid CNN-RNN attention-based model for text classification," in International Conference on Conceptual Modeling, Springer, pp. 571-585, 2018.

[18] X. She and D. Zhang, "Text classification based on hybrid CNN-LSTM hybrid model," 2018 11th International Symposium on Computational Intelligence and Design (ISCID), IEEE, vol. 2, pp. 185-189, 2018.

[19] J. Zhang, Y. Li, J. Tian, and T. Li, "LSTM-CNN hybrid model for text classification," 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), IEEE, pp. 1675-1680, 2018.

[20] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," arXiv preprint arXiv:1409.0473, 2014.

[21] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," Neurocomputing, vol. 452, pp. 48-62, September 2021.

[22] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. R. Salakhutdinov, and Q. V. Le, "Xlnet: Generalized

autoregressive pretraining for language understanding," Advances in neural information processing systems, vol. 32, 2019.

[23] I. Santos, N. Nedjah, and L. d. M. Mourelle, "Sentiment analysis using convolutional neural network with fastText embeddings," Arequipa, Peru, 2017 IEEE Latin American Conference on Computational Intelligence (LA-CCI), pp. 1-5, Nov. 8-10, 2017.

[24] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," arXiv preprint arXiv:1301.3781, 2013.

[25] P. Masurel, Of bayesian average and star ratings. https://fulmicoton.com/posts/bayesian_rating/ (accessed Dec. 16, 2021).

[26] C. Study. "What algorithm does IMDB use for ranking the movies on its site?" Quora. https://www.quora.com/What-algorithm-does-IMDB-use-for-ranking-the-movies-on-its-site (accessed Dec. 14, 2021).