

# Examining Cultural Differences in Korean and English Speakers: A Study in Sentiment Analysis

Anthony Sainez<sup>1</sup>, Jingyi Jennie Wu<sup>2</sup>

<sup>1</sup>University of California, Merced – CA, U.S.

<sup>2</sup>University of California, Santa Barbara – CA, U.S.

e-mail: asainez@ucmerced.edu, jingyi\_wu@umail.ucsb.edu

**Abstract** – While the influence of the cultural differences between the East and West on human behaviors has been studied for years, there has been less attention paid to its influence on online behaviors, no need to say about opinion expression regarding Korean idols involved in scandals. By analyzing Korean and English tweets from Twitter, one of the most active Korean popular music fandom spaces, through the lens of sentiment analysis, the paper aims to identify a relationship between the cultural differences and online opinion expressions. Possibly undergoing parasocial breakup, do people from different cultural backgrounds react differently toward the same event? Analyzing the recent school bullying scandal with Kim Garam, a member of the newly debuted girl group Le Sserafim, the paper aims to answer the question if there is a difference in sentiment expression regarding this school bullying event between Korean and English users. The study shows a trivial fact that both the Korean and English data sets follow a Zipf's law, but a significant enough difference between the data sets that support our hypothesis. However, due to several limitations, such as restricted access to the Twitter database or little research done dealing with Korean sentiment analysis, the results of this paper should be dealt with caution and require further validation.

**Keywords** – English, Korean, Natural Language Processing, Sentiment Analysis

## I. INTRODUCTION

This paper introduces a data set from Twitter that seeks to study the cultural difference entailed by the controversial event involving Kim Garam (김가람), a member of Hybe Entertainment's new girl group, Le Sserafim (르세라핌). Hybe is an internationally recognizable South Korean entertainment company, known mainly for their most influential boy band BTS (Bangtan Boys 방탄소년단), which is the best-selling artist in South Korean history among other accolades. Due to their success with BTS, Hybe received high attention for their formation of the new girl group. Especially after the news that former Iz\*One members Sakura Miyawaki (宮脇 咲良) and Kim Chaewon (김채원) would join the group, the expectation for this girl group peaked. However, as soon as Kim was unveiled to be the second member of the group, Korean internet forums were filled with her past middle-school bullying histories and subsequent petitions to exclude her from the group. Kim went on indefinite hiatus on May 20th, 2022, eighteen days after the group's debut,

due to bullying allegations in middle school. Different voices emerged on Twitter to either support or defy the decision. Even though there has been previous research done regarding cultural influence on social networking service (SNS) usage [1], there have been limited studies on if people behave differently online regarding controversial events according to their cultural context. Consequently, by studying Kim's case through analyzing the sentiments of English and Korean tweets on Twitter, this paper aims to explore the influence of cultural differences between the East and the West on online tweeting behaviors.

K-pop (케이팝 Korean Popular Music) has gained international recognition in the past decades. The high-quality production of music videos and unique style of this music genre attracted millions of fans worldwide. In the middle of this K-pop wave are the K-pop idols (아이돌), who usually went through auditions and hard-core trainee periods to finally debut. However, a debut does not promise immediate success for their career. With numerous South Korean idol groups debuting annually, the competition to win over fans is more than intense. Usually, only those who debut from the three biggest entertainment companies in South Korea, SM Entertainment, YG Entertainment, and JYP Entertainment, could attain global popularity.

Nevertheless, because of the seemingly easy success of most K-pop idols, people are critical of who could become the glorious celebrities. First of all, when the wannabes attend auditions, they are selected mostly based on several factors: appearance, performing skills (singing, dancing, and/or rapping), personality, etc [2]. Among all these factors, physical appearance might be the most determinant facet of a wannabe's possibility to become successful. After the debut, most idols also undergo constant cyberbullying and only those who are psychologically mindful of these reckless comments could continue their career. They could be requested by fans to lose weight or to attain plastic surgery to improve their presence on camera. If someone is not good at singing, dancing, or rapping, they will be criticized for their unprofessional skills as well. After overcoming all these obstacles to finally become an idol, however, the efforts could be eliminated within seconds if he/ she got reported on any wrong-doings, including but not limited to sexual assault, controversial behaviors/ comments, or bullying. Instead of being passive pursuers after idols, the K-pop fandom, therefore, exerts huge power over the idols as well.

In South Korea, particularly, people who have had histories of bullying others are likely to be boycotted. Celebrities are highly influential people whose behaviors could bring a huge impact on their fans (Shabahang et al., 2021). To prevent the

possibility that the young adults could follow their idols to bully their classmates, which could possibly lead to incidents of suicides [3], Korean people are usually strict on idols with this kind of scandals. If the scandal got verified to be true, the idol's career could terminate forever. For example, Seo Soo-jin (서수진), a former member of the girl group (G)I-dle under Cube Entertainment, stopped her idol career and withdrew from (G)I-dle in 2021 due to her school bullying accusations. Her withdrawal from the group is a reasonable practice in South Korea to avoid bringing more harm to the whole group's reputation.

The seriousness of Korean people regarding K-pop idols with scandals then explains the reason why people keep debating about Kim's case with Le Sserafim. As most Korean fans are worried about the group's reputation being harmed by Kim's scandal, they petition on Twitter to Hybe to terminate Kim's activities with the group. However, there have also been other voices saying why people care about someone's bullying history back in the date, which does not imply anything about this person today. Therefore, this study aims to explore if the opposing opinions on the scandal reveal whether the post is posted by someone from the Eastern or Western country. Through analyzing people's reactions to the scandal on Twitter, one of the biggest active K-pop fandom [4], this paper aims to answer the question: "if the online tweeting behaviors unveil the differences between the Eastern and Western cultures. More specifically, the differences between collectivism and individualism."

## II. LITERATURE REVIEW

Previous research has been done analyzing the relationships built between celebrities and fans, which is referred to as the "parasocial relationship." Horton and Wohl were the first ones who coined the word to refer to the particular relation between the celebrity and viewer built directly/ indirectly through in-person interaction/ media [5]. Under usual circumstances, the relationship is one-sided and illusive, as the only times the fans could interact with the celebrities are during official events arranged by the management companies. Even though the parasocial relations only exist in the viewer's illusion, they still regard it as a reciprocal interaction between the celebrities and themselves [6]. In this case, the celebrities are positioned as the active provider of illusion, while the viewers are the passive receivers with no significant power [2]. However, this situation, as time goes by, changes gradually.

Most of the time, people view celebrities as distant exemplars. People follow their fashion, behaviors and consume the brands they like. As public figures, celebrities usually receive high attention from the general public, which also results in high retention of information because celebrities "stimulate people's collective or cultural consciousness" [7]. For example, the case when "Magic" Johnson disclosed his human immuno- deficiency virus (HIV) infection raised the public's attention to HIV indicates the positive effects brought by celebrities. Thus, the general public is more interested in celebrity-related scandals, news, or advertisements, and could be easily affected by their behaviors.

Instead of being the mere receiver of information from their idols, K-pop fans enhance the sustainability of the

idol-fan relationship by transforming "fans' engagement from a parasocial illusion to an economic activity" [2]. As mentioned before, numerous wannabes enter the K-pop industry annually despite the high competition. On the other hand, this competition also transformed the relationship between idols and fans from distant to dependent. Idols became more dependent on fan support to become popular or even take action against unfair treatment from companies. For example, the famous case when Junsu, Yoochun, and Jaejoong, previous members of TVXQ (Tong Vfang Xien Qi), left the group because of the unfairness of the profit share on their contract with the SM Entertainment manifested the power of K-pop fan club. The fans actively supported these three members' decision to leave and even filed petitions to different institutions, which finally led to a revised standard contract for the idols from the Fair Trade Commission. The new contract promises more reasonable conditions for idols, such as a reduced penalty for breach of contract [2]. This change, in a sense, brought the fans' power to be in a more dominant position in this idol-fan relationship.

Rather than being permanent, the parasocial relationship is vulnerable to parasocial breakup. Parasocial breakup is defined as the termination of a parasocial relationship because of a negative reaction to the celebrity. [8] Among all the factors that cause the parasocial breakup, celebrity scandals are one of the contributors. There are mainly two reasons for this: 1) increased uncertainty 2) contradiction with their expectation. Favorability towards a celebrity is positively related to the certainty viewers have towards the latter. As scandals add more uncertainty to the parasocial relationship, the feeling of being deceived could possibly lead to a parasocial breakup. In addition, the repeated interactions between the celebrity and fans could build up expectations the fans hold. Whenever a scandal of a celebrity breaks out, this expectation is violated and results in negative reactions [9]. Consequently, a parasocial breakup could terminate parasocial relations due to scandals.

The fact that the parasocial relationship with celebrities are more intense among adolescent fans compared to the middle-aged groups has also been verified through different research [10]. Usually, young adults are not mature enough to differentiate what kind of actions carried out by their loved celebrities are good or bad to follow. If they blindly follow what seems "cool" conducted by the celebrities, these behaviors could bring harm to the young generation. In particular, in South Korea where K-pop idols are the dominant influencers to young adults, the general public is critical of any idols with scandals that could serve as bad exemplars for the young adults. Plus, the young adults could react aggressively after a parasocial breakup happens that terminates their once intense parasocial relationship [9]. In the case of bullying scandals by K-pop idols, the South Korean people are afraid that the fans would follow their idols to bully their classmates, which, evidently, could lead to depression, decreased self-worth and even suicides [11]. The zero-tolerance of scandals that any Korean celebrities might be involved in explains the South Korean people's petition to stop Kim from participating in Le Sserafim's activities.

The question now is: "Do westerners, who are raised in a

different cultural background, hold different opinions towards celebrities with scandals?" The cultural differences between the East, the collectivism, and the West, the individualism, have been studied for a long time. This difference is demonstrated in various contexts. For instance, as shown by research, the motivations to utilize social media are distinct between American and Korean college students. While American students tend to use social network sites (SNSs) for casual relations, Korean students seek more serious and deeper connections. This difference is related to the cultural difference between individualism and collectivism, while the former highlights more casual and instrumental relationships, the latter emphasizes deep social bonds and commitments [1].

Not only in motivations to use SNSs, but also in the bullying behaviors, can people see the cultural differences. Compared to most individualistic nations, i.e. western countries, the eastern countries with an emphasis on collectivism are prone to conduct bullying behaviors as a group. Two Korean words are used frequently to describe the bullying incidents happened in school contexts: wang-ta (왕따) and jun-ta (전따). The former refers to the situation when the whole class shuns a pupil while the latter means the whole-school distancing from a pupil. In both cases, they refer to the group bullying behaviors [9]. In contrast, the bullying behaviors in western countries could be more personal.

Hence, to study the online tweeting behaviors, this paper propose the hypothesis:

*H*: The online expressions towards Kim's school bullying event on Twitter are influenced by cultural differences between Korean and English-speaking countries.

### III. METHOD

We looked at the problem of identifying global perspectives on K-Pop through the lens of natural language processing—specifically sentiment analysis, wherein the goal is to determine attitudes and opinions (sentiments) expressed around a topic. We began our investigation by attempting to create a dataset that might give us more information about international sentiments around K-Pop. The guiding principle was to curate data from English and Korean sources in order to examine the differences between these two sociolinguistic groups.

We considered multiple platforms for data gathering, though we focused mainly on popular social media websites such as YouTube, Twitter, Reddit, and Instagram because we believed these would encapsulate a more global perspective than other more potentially niche or focused online communities. By the end of the process, we only selected data from Twitter because it is the most widely used platform by both linguistic groups. In addition, the K-pop fandom community on Twitter could be seen as the most active one as well. People from different parts of the world come on Twitter to establish interpersonal relationships and participate in mutual engagements. They follow influential fansites, interact with people who have similar interests or try to stay updated with their idols' schedules, etc [4]. Within an online platform, everyone is granted with the freedom to obtain and share knowledge with one another. More importantly, they

can express their opinions on controversial events freely as well. In that case, the tweets on Twitter could be the most dynamic and diverse data to analyze. It should be noted, however, that there are still interesting trends to be gleaned from other platforms—we suggest perhaps a multi-platform analysis might yield the most balanced results.

To further explain our reasoning for eliminating other platforms, we can start with Reddit, which appeared to not have very much Korean data, and so would not prove useful for comparing linguistic groups. Instagram data would have been multimodal—posts are both images and captions, which presents complications we are not yet prepared to deal with. There is not much information on multimodal sentiment analysis for the Korean language. YouTube data, which is gathered from comments under popular music videos, did not appear to contain much meaningful discourse: comments were usually simple compliments (or detriments) towards a group's performance, musical aptitude, or aesthetics.

After the target platform was established, we began the data collection process. The initial raw data set we generated included 2,295 Tweets, 177 of which contained no text information. According to standard data mining principles, the next step was to clean up and preprocess the data set [12]. We removed all null or empty data points, duplicates, and non-English or non-Korean texts. Additionally, we removed some spam or irrelevant data points, such as advertisements or giveaways. In the end, the data shrank from 2,295 entries to 1,566 entries, which is a 37.7622% decrease. When splitting it by language, there were 526 Korean entries and 1,040 English entries.

**TABLE I**  
**English Dataset**

Sample	
Text	Date
fuck you hybe	2022-05-26 05:29
Listen if you support Kim Garam but believe...	2022-05-26 06:55
anyway no center ever could outdo garam	2022-05-26 07:05
garam's face is top tier	2022-05-26 07:05

For the Korean data set, an additional column is provided wherein the text was batch translated using Google Cloud Translate's API from Korean into English. As noted by others who have utilized this technique, there is a good possibility though that these translations do not fully capture the nuances in sentiment for every text [13]. However, for simplicity's sake, we chose to use this translated text as predictors for sentiment in an actual model.

When it comes to predicting the sentiment of some text, there are multiple state-of-the-art methods. For example, it is common to use base classification techniques, such as naive Bayes classifiers, support vector machines, or even ensemble techniques such as random forests for natural language processing tasks in general [14]. However, there also exist other models that are tailored specifically for solving these tasks, such as VADER (Valence Aware Dictionary for sEntiment Reasoning) [15]. Unlike other supervised learning methods, VADER uses a hand-crafted lexicon and rule set to identify the polarity of a text—which also means it only exists

for English texts.

There are also deep learning methods that use pre-trained transformers to achieve highly accurate results; BERT, developed by Google, is a famously well-performing model that employs this method [16]. The Korean equivalent, KoBERT, uses mostly the same model with a different tokenizer [17]. Since Korean is an agglutinative language, there are many different ways to tokenize sentences—ranging from as tokenizing on basic consonant-vowels up to advanced and more linguistically-aware methods— and because tokenization is a component of data preprocessing before model training, it can affect the results (accuracy or precision) of a deep learning model [18]. While these models are impressive, they require millions of labeled data points and can take a long time to train.

Because our data set is unlabeled, we were put in a somewhat uncomfortable position when it comes to sentiment analysis, which is mostly performed on labeled data sets. In other words, this meant that we had to use a preconstructed model (possibly trained on another data set) in order to analyze our results. However, training and fine-tuning a deep learning model was prohibitively time-consuming, so we opted to use VADER, using the translations as input for the Korean data set.

#### IV. RESULT

The resultant data set we have created thus far is a rudimentary one, and it is our hope that later research efforts might be able to improve upon our existing process. Part of the purpose of this paper is to document the mistakes and corrections made during the data collection process. Since there is a research gap in the area of multilingual sentiment analysis (especially with respect to Korean), there is still value in this basic study.

We observed a few things about the data. A somewhat trivial observation was that this set follows Zipf’s law, which is a statistical phenomenon observed wherein the rank-frequency of a data set (often a corpus of text) follows an inverse relationship. When examining both our Korean and English data sets, we observed that it was true, as both follow a Zipfian distribution of keywords— as expected.

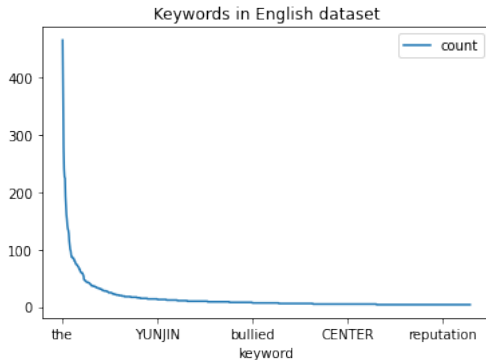


Fig. 1. An example of the Zipfian distribution of keywords in the English dataset.

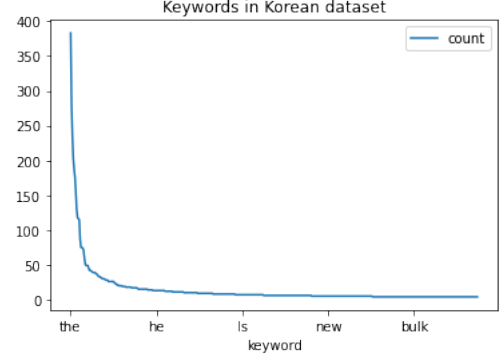


Fig. 2. An example of the Zipfian distribution of keywords in the Korean dataset.

As far as we are aware there aren’t many serious implications of this other than it being a somewhat neat statistical fact that this corpus also happens to obey, despite the issues with the data set.

If we want to look at the distribution of sentiments, we look to the VADER’s compound sentiment metric. The compound score is a metric that calculates the sum of all the ratings which have been normalized between -1 (most extreme negative) and +1 (most extreme positive). So, a positive sentiment has a compound score greater than 0.05, while a neutral sentiment has a compound score between -0.05 and 0.05. It follows then that negative sentiments are expressed by compound scores less than -0.05.

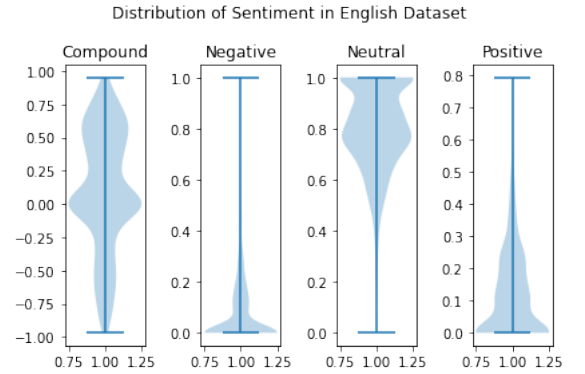


Fig. 3. Violin plots of the distribution of sentiment scoring in the English data set. Note that this plot visualizes both range and distribution.

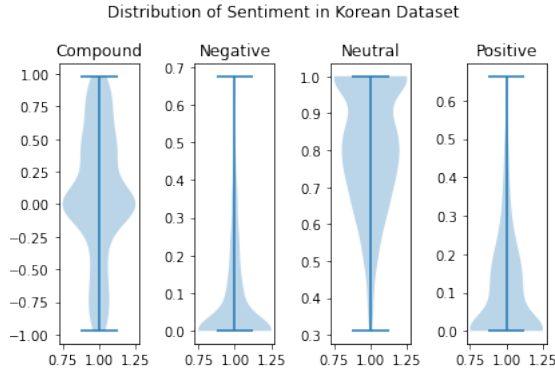


Fig. 4. Violin plots of the distribution of sentiment scoring in the Korean data set.

As you can see from statistical analysis, the distribution of sentiment in the Korean data set is very slightly different from the English one. This can be observed via qualitative analysis (from the violin plots shown above) as well as quantitative. For instance, for the English data set, the mean and standard deviation are 0.098602 and 0.452713 respectively. For the Korean data set, it's 0.072254 and 0.490215 respectively.

This means there is a 30.8423% difference in mean compound scores, though the standard deviations are similar. We find this to mean that the distributions of similar follow a similar pattern but are slightly offset, which is just statistically significant to be noted. This supports our hypothesis that sentiments expressed between Korean and English internet users might be different, though we urge further investigation.

## V. CONCLUSION

### A. Discussion

This research's initial task was to examine differing opinions about K-pop from a global perspective; that is to say, we wanted to examine how different sociocultural groups perceived this huge and emerging industry. It is by looking at this that we might be able to glean some sort of information about global perspectives possibly about South Korea or the music industry in general.

This insight could be useful for many reasons. It potentially points to new areas for research in the fields of sociolinguistics, machine learning and statistics, and cultural studies on international relations between South Korea (or Asia) and the rest of the world. Regardless, the industry of K-pop is an emerging giant in the global cultural sphere, and it would be important to understand this new phenomena well.

### B. Limitations

We applied and were denied access to Twitter's Academic Research API v2, which would have simplified the process greatly. It should be noted that we were also denied the possibility of reapplying for access. Because of this limitation, we created a program that scrapes data only up to a week old and continually polls for newly generated tweets that match a search keyword.

There were other issues with the data set. For example, there were multiple hangups we encountered during the data cleanup process. We were not able to find adequate libraries in Python that would have aided us in the process of cleaning

up Korean Twitter text,<sup>1</sup> and so that data had to be managed differently than our English text— that is to say, in a more manual manner. This means that some of our Korean data still has anomalies such as broken links (which are difficult to find even via regular expressions), and other artifacts that were a result of previous attempts to manually preprocess the text. Because our Korean data set was so small and because we were denied access to gathering additional data in a more comprehensive manner, we ultimately made the decision to keep this data despite its problems. We believe it would be beneficial for a manual overview of every point in the data set, such as via Amazon's Mechanical Turk service. However, such a process can be time-consuming and expensive.

Additionally, we believe that although the use of Google Cloud Translate for sentiment analysis of another language is not unprecedented, we do acknowledge that it is possible that some nuance may be left out due to this process. It would likely be better to run individualized, tailored or fine-tuned sentiment analyzers on each individual language.

### C. Implication

Because we did find some interesting statistical data we have shown that this is an area that might require further research instead of being a trivial question. The statistical difference we found in our data set might be verified or invalidated by further research, and we cannot be totally certain of our results because of the limitations expressed earlier.

Additionally, all of the code associated with this work can be found on GitHub, and is well documented for reproducibility purposes. We hope that possibly by learning from the mistakes we made that other research groups might be able to make faster progress.

## VI. ACKNOWLEDGEMENTS

The authors thank Seungyeon Won for the instruction for this paper and Barun ICT Research Center at Yonsei University for providing the research opportunity.

## REFERENCES

- [1] Y. Kim, D. Sohn, S. M. Choi, "Cultural difference in motivations for using social network sites: A comparative study of American and Korean college students", vol. 27, no. 1, pp. 365–372, doi:10.1016/j.chb.2010.08.015, URL: <https://linkinghub.elsevier.com/retrieve/pii/S0747563210002736>.
- [2] S. Barbara, "Gender, Labor, and the Commodification of Intimacy in K-pop", p. 270.
- [3] Y. S. Kim, B. L. Leventhal, Y.-J. Koh, W. T. Boyce, "Bullying Increased Suicide Risk: Prospective Study of Korean Adolescents", vol. 13, no. 1, pp. 15–30, doi:10.1080/13811110802572098, URL: <http://www.tandfonline.com/doi/abs/10.1080/13811110802572098>.

<sup>1</sup>The library we used to preprocess English tweets could not handle non-ASCII encoded characters, which resulted in the removal of all non-English characters. This extended to other Latin characters with diacritics, such as ü in German or é in Spanish [19].

- [4] Z. Malik, S. Haidar, "Online community development through social interaction — K-Pop stan twitter as a community of practice", pp. 1–19, doi:10.1080/10494820.2020.1805773, URL: <https://www.tandfonline.com/doi/full/10.1080/10494820.2020.1805773>.
- [5] T. Hartmann, C. Goldhoorn, "Horton and Wohl Revisited: Exploring Viewers' Experience of Parasocial Interaction", *Journal of Communication*, vol. 61, pp. 1104 – 1121, 12 2011, doi: 10.1111/j.1460-2466.2011.01595.x.
- [6] M. Hu, "The influence of a scandal on parasocial relationship, parasocial interaction, and parasocial breakup.", vol. 5, no. 3, pp. 217–231, doi:10.1037/ppm0000068, URL: <http://doi.apa.org/getdoi.cfm?doi=10.1037/ppm0000068>.
- [7] W. Yoo, "The influence of celebrity exemplars on college students' smoking", vol. 64, no. 1, pp. 48–60, doi:10.1080/07448481.2015.1074238, URL: <http://www.tandfonline.com/doi/full/10.1080/07448481.2015.1074238>.
- [8] J. Cohen, "Parasocial Breakups: Measuring Individual Differences in Responses to the Dissolution of Parasocial Relationships", *Mass Communication and Society*, vol. 6, no. 2, pp. 191–202, 05 2003, doi: 10.1207/S15327825MCS0602{\\\_}5, URL: [https://doi.org/10.1207/S15327825MCS0602\\_5](https://doi.org/10.1207/S15327825MCS0602_5).
- [9] R. Sittichai, P. K. Smith, "Bullying in South-East Asian Countries: A review", vol. 23, pp. 22–35, doi:10.1016/j.avb.2015.06.002, URL: <https://linkinghub.elsevier.com/retrieve/pii/S135917891500083X>.
- [10] C. Klimmt, T. Hartmann, H. Schramm, "Parasocial Interactions and Relationships", p. 13.
- [11] S. Hinduja, J. W. Patchin, "Bullying, Cyberbullying, and Suicide", vol. 14, no. 3, pp. 206–221, doi:10.1080/13811118.2010.494133, URL: <http://www.tandfonline.com/doi/abs/10.1080/13811118.2010.494133>.
- [12] B. P. C. G. P. P. N. R. Shmueli, Galit, *Data mining for business analytics : concepts, techniques and applications in Python*, 2020.
- [13] M. Bautin, L. Vijayarenu, S. Skiena, "International Sentiment Analysis for News and Blogs", *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 2, no. 1, pp. 19–26, Sep. 2021, URL: <https://ojs.aaai.org/index.php/ICWSM/article/view/18606>.
- [14] Ankit, N. Saleena, "An Ensemble Classification System for Twitter Sentiment Analysis", *Procedia Computer Science*, vol. 132, pp. 937–946, 2018, doi:<https://doi.org/10.1016/j.procs.2018.05.109>, URL: <https://www.sciencedirect.com/science/article/pii/S187705091830841X>, international Conference on Computational Intelligence and Data Science.
- [15] C. Hutto, E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text", *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, no. 1, pp. 216–225, May 2014, URL: <https://ojs.aaai.org/index.php/ICWSM/article/view/14550>.
- [16] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", , 2018, doi:10.48550/ARXIV.1810.04805, URL: <https://arxiv.org/abs/1810.04805>.
- [17] K. Yang, "Transformer-based Korean Pretrained Language Models: A Survey on Three Years of Progress", , 2021, doi:10.48550/ARXIV.2112.03014, URL: <https://arxiv.org/abs/2112.03014>.
- [18] K. Park, J. Lee, S. Jang, D. Jung, "An Empirical Study of Tokenization Strategies for Various Korean NLP Tasks", , 2020, doi:10.48550/ARXIV.2010.02534, URL: <https://arxiv.org/abs/2010.02534>.
- [19] "Python Package Index - PyPI", URL: <https://pypi.org/>.