ICIC Express Letters
Part B: Applications
Volume **16**, Number **3**, March **2025**

ICIC International ©2025 ISSN 2185-2766
pp. 333–341

SENTIMENT ANALYSIS OF THAI LABORERS' PERCEPTIONS OF WORKING ABROAD: A MACHINE LEARNING APPROACH USING YOUTUBE COMMENTS

Artitaya Rabruen[1], Pimpat Pokkate[1], Olarik Surinta[2] and Sanya Khruahong[3,*]

[1]Naresuan University Secondary Demonstration School
[3]Department of Computer Science and Information Technology
Faculty of Science
Naresuan University
99 Moo 9, Thapo Sub-District, Muang District, Phitsanulok 65000, Thailand
{ artitayar64; pimpatp64 }@nu.ac.th; *Corresponding author: sanyak@nu.ac.th

[2]Multi-agent Intelligent Simulation Laboratory (MISL) Research Unit
Faculty of Informatics
Mahasarakham University
Khamriang Sub-District, Kantarawichai District, Maha Sarakham 44150, Thailand
olarik.s@msu.ac.th

Received June 2024; accepted August 2024

Abstract. *The rising trend of Thai workers seeking employment overseas necessitates a nuanced understanding of their attitudes towards labor migration. This research paper delves into the perceptions of Thai workers about working overseas, focusing on five destinations: Australia, Japan, South Korea, Taiwan, and the United States. It applies machine learning methods to dissecting sentiments embedded in 37,077 comments from 400 YouTube videos, covering preparation and law, lifestyle, and work experience. The study uses Python for computational analysis to sort these comments into positive, negative, and neutral sentiments. The Naïve Bayes Support Vector Machine (NBSVM) algorithm emerged as the most effective model for classifying these sentiments. Our findings indicate that Australia elicited the most positive responses (32.31%) and the least negative perceptions, whereas Japan registered the highest proportion of negative sentiments (15.43%) across various aspects. The results, illustrated through quantitative percentages and visual representations like bar charts and word clouds, underscore the potential of machine learning in providing actionable insights for policymakers and market analysts in labor migration.*
Keywords: Sentiment analysis, Machine learning, Laborers' perceptions, Naïve Bayes support vector machine

1. **Introduction.** Thailand's labor market has long struggled with issues such as high unemployment rates and a shortage of skilled workers in key sectors. Consequently, a growing number of Thai workers have turned to international employment, both through legal and illegal channels, in pursuit of higher wages and better working conditions. However, this trend has raised significant concerns related to worker exploitation, human trafficking, and the long-term well-being of migrant laborers. These concerns have become central to labor policy discussions both within Thailand and on an international level. As such, there is a pressing need to better understand the perceptions and experiences of Thai laborers working abroad, which can inform policies aimed at improving labor migration conditions.

DOI: 10.24507/icicelb.16.03.333

The proliferation of social media platforms, especially YouTube, has provided a rich data source for sentiment analysis, a method widely used in Natural Language Processing (NLP) [1] to classify subjective opinions in user-generated content. Sentiment analysis has proven effective in identifying public attitudes across a wide range of topics. Sharma et al. [2] highlighted the use of sentiment analysis techniques in examining labor-related issues through social media data. In 2022, using the S-Sense framework [3] and the BiLSTM (Bidirectional Long Short-Term Memory) [4] yielded valuable insights into the intricate association between social media utilization and community-based tourism in Thailand. The findings have the potential to enhance our comprehension of the impact of social media on tourism practices and hold significant implications for tourism stakeholders and policymakers in Thailand. Additionally, Lu et al. [4] applied deep learning models to extracting sentiment trends from online platforms, further establishing the relevance of sentiment analysis in labor research. While sentiment analysis has been employed in broader contexts, there remains a notable gap in research specifically addressing the sentiments of Thai migrant laborers.

Recent advancements in machine learning, particularly the development of hybrid models like the Naïve Bayes Support Vector Machine (NBSVM), have significantly improved classification accuracy in sentiment analysis. Muhammad et al. [5] demonstrated the superior performance of NBSVM in classifying sentiments from social media data. However, the application of such models to the context of Thai labor migration has not been thoroughly explored, representing a critical gap in the existing literature.

This study aims to address this gap by utilizing machine learning techniques to analyze YouTube comments related to the experiences of Thai laborers in five key destination countries: Australia, Japan, South Korea, Taiwan, and the United States. Specifically, this research employs the NBSVM algorithm to classify sentiments from 37,077 comments, focusing on aspects such as preparation and law, lifestyle, and working conditions.

The key contributions of this study are 1) the introduction of a novel dataset centered on Thai labor migration, 2) the application of advanced machine learning techniques for more accurate sentiment classification, and 3) actionable insights for policymakers aimed at improving labor conditions for Thai workers abroad. The findings contribute to a more nuanced understanding of Thai labor migration, offering a data-driven foundation for future policy development.

The following sections detail sentiment analysis techniques and machine learning models used to analyze Thai labor sentiments from YouTube comments. We address challenges, and limitations, and provide recommendations for future research. Additionally, we explore the potential applications of sentiment analysis in improving worker welfare and guiding labor-related policy decisions.

2. **Sentiment Analysis.** Sentiment analysis [6,7] has been extensively studied in Natural Language Processing (NLP), particularly within the domain of machine learning. Various methods for sentiment analysis have been explored, including rule-based, machine learning-based, and hybrid approaches [8]. Rule-based methods rely on handcrafted rules to determine sentiment, while machine learning-based methods use labeled datasets to automatically learn and classify sentiment. Hybrid approaches combine both techniques to improve accuracy.

In the context of YouTube comments, sentiment analysis has been applied to assessing viewers' sentiments toward different types of content, such as movies, TV shows, and products. Researchers have utilized various techniques, including lexicon-based, machine learning-based, and deep learning-based approaches, to analyze sentiments from YouTube comments. Lexicon-based methods depend on pre-existing sentiment lexicons to classify text, whereas machine learning algorithms learn from labeled data to classify sentiments.

Deep learning methods [9], such as neural networks, offer another layer of sophistication in sentiment classification.

Overall, machine learning-based methods were selected for this study because of their exceptional ability to handle large datasets and automatically learn patterns from labeled data, making them well-suited for analyzing unstructured social media content like YouTube comments. Although rule-based approaches offer simplicity in sentiment classification, their reliance on handcrafted rules restricts their flexibility and scalability, particularly when dealing with large, diverse datasets. Considering the size and complexity of the dataset in this research, machine learning-based approaches offer a more efficient and practical solution.

3. **Research Methodology.** This study is based on a quantitative approach using sentiment analysis techniques and machine learning algorithms applied to YouTube comments. The research process can be divided into three main stages: data collection and preparation, analysis, and evaluation.

Figure 1. Three main stages of sentiment analysis

1) The first stage of the study involved the collection and preparation of the dataset. The researchers gathered comments made by Thai individuals on YouTube videos related to Thai workers abroad, focusing on five countries: Australia, Japan, South Korea, Taiwan, and the United States – countries where many Thai workers seek employment. The data was scraped from YouTube using the Instant Data Scraper application, and the resulting dataset was stored in Excel format. To ensure a diverse sample, 80 videos were randomly selected from each country, resulting in a total of 400 videos. Key variables in the dataset included the commenter's name, the comment text, the date of the comment, the number of likes, the total number of original comments, and the number of comments remaining after data cleaning. The cleaning process involved several steps: removing duplicate and irrelevant comments, such as spam or advertisements, and handling missing data by discarding incomplete entries. Special characters, punctuation, and URLs were removed, and tokenization and stop word removal were applied to filtering out common, non-informative words. Lemmatization was used to ensure that different forms of the same word were treated uniformly. After cleansing, the dataset was refined, leaving only relevant, high-quality comments for sentiment analysis, the number of original comments and the number of after-cleaning comments, as shown in Table 1. Table 2 outlines the different comment categories related to *preparation and law*, *lifestyle*, and *working*.

Table 1. The dataset of all comments

Countries	**Number of original comments**	**Number of after-cleaning comments**
Australia	14,464	8,634
Japan	13,425	7,457
South Korea	17,680	8,935
Taiwan	5,434	3,338
United States	14,901	8,713
Total	**65,904**	**37,077**

Table 2. The dataset of comments in each group

Countries	Preparation and law	Lifestyle	Working	All
Australia	3,837	5,652	7,872	8,634
Japan	2,041	5,996	6,287	7,457
South Korea	4,302	4,197	6,886	8,935
Taiwan	1,866	1,987	1,396	3,338
United States	4,952	6,606	6,641	8,713

Once the data was cleaned, the Naïve Bayes Support Vector Machine (NBSVM) model was used for sentiment classification. NBSVM combines two approaches: Naïve Bayes, which calculates the probabilities of word associations for sentiment categories, and Support Vector Machine (SVM), which classifies the comments based on the transformed feature space. By first transforming the text data with Naïve Bayes and then using SVM for classification, NBSVM effectively captures both the statistical relationships between words and the optimal decision boundaries for sentiment classification. This combination provided high accuracy and improved sentiment classification performance for the dataset.

2) The second stage of the research involved the analysis of the dataset using sentiment analysis techniques and machine learning algorithms. The sentiment analysis technique used was the Bag-of-Words (BoW) model [10], a typical text analysis approach. This technique involves separating a text into individual words or tokens and representing them as a bag of unordered words. The BoW model was used to determine the sentiment of the comments in the dataset, whether they were positive, negative, or neutral.

During this stage, the collected comments from step 1 underwent processing by the researchers using essential data analysis packages such as *pandas* and *matplotlib* in PyThaiNLP [11], a Thai Natural Language Processing library in Python. It was employed for word tokenization and removing the text's stop words and punctuation marks. These preprocessing steps ensured that the data maintained a high quality, accuracy, and appropriateness for further analysis [12]. Furthermore, the researchers added spaces between words in the text to properly prepare it for word cloud processing. This step was crucial to ensure that the text was formatted correctly and could be effectively visualized as a word cloud [13].

Machine learning algorithms were then applied to the dataset to classifying the comments into the three sentiment categories. The Naïve Bayes Support Vector Machine (NBSVM) algorithm [5,14,15] was used for this task. NBSVM is a supervised machine learning algorithm that can be used for classification tasks of sentiment analysis. It is a machine learning algorithm that combines the strengths of the Naïve Bayes (NB) [16] classifier and the Support Vector Machine (SVM) [17,18] classifier to improve performance in text classification tasks, especially sentiment analysis, as shown in Figure 2.

3) The final stage of the research involved performance evaluation. The accuracy of the three algorithms in classifying the comments was assessed using Precision, Recall,

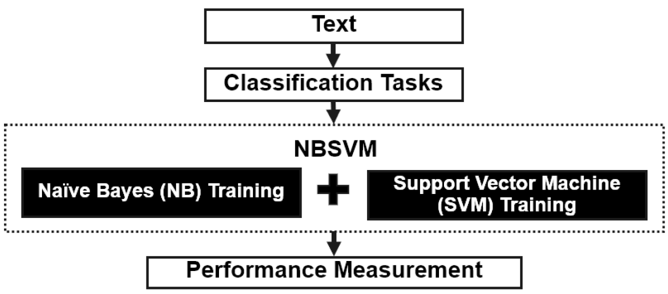

Figure 2. Process of machine learning algorithms

and F1-Score metrics, which are calculated based on True Positives (TP), False Positives (FP), and False Negatives (FN). The results were further analyzed to identify any trends or patterns in the sentiments expressed in comments about Thai workers overseas, as shown in Table 3.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{1}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

$$\text{F1-Score} = 2 * \frac{(\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{3}$$

TABLE 3. Results of the performance measurement of the classification models

Classification	Precision	Recall	F1-Score
NB	89.13%	89.55%	89.34%
SVM	90.45%	88.36%	89.39%
NBSVM	**92.21%**	**91.42%**	**91.81%**

Table 3 summarizes the performance metrics of three classification models: Naïve Bayes (NB), Support Vector Machine (SVM), and Naïve Bayes Support Vector Machine (NB-SVM). The metrics used are Precision, Recall, and F1-Score. According to the presented data, these results indicate that the NBSVM model has the best overall performance among the three models, with the highest Precision of 92.21%, Recall of 91.42%, and F1-Score of 91.81%.

4. **Results and Analysis.** Analyzing the comments in each data file's "comment" or "text_tokens" column revealed that a significant portion of the opinions fell within the neutral range. This was mainly because these comments consisted predominantly of question messages or narrative sentences that did not convey clear positive or negative sentiments. As a result, the overall opinion regarding working leaned toward neutrality.

4.1. **Analysis of labor data in Australia.** The sentiment analysis results were classified by the type of comments about working in Australia on YouTube, as shown in Table 4. The detail of Table 4 reveals that the majority of opinions expressed by Thai workers were classified as neutral, which indicated that most of the comments consisted of questions or informative statements without a clear positive or negative sentiment. However, when focusing specifically on the positive and negative opinions, it was observed that areas such as preparation and law, lifestyle, and working had a higher proportion of positive opinions than negative ones.

TABLE 4. Results of sentiment analysis in Australia

Comment type	Preparation and law	Lifestyle	Working	All
Positive	1,166 (30.39%)	**1,879 (33.25%)**	2,594 (32.95%)	2,790 (32.31%)
Negative	162 (4.22%)	267 (4.72%)	354 (4.50%)	391 (4.53%)
Neutral	**2,509 (65.39%)**	**3,506 (62.03%)**	**4,924 (62.55%)**	**5,453 (63.16%)**
Total	3,837 (100%)	5,652 (100%)	7,872 (100%)	8,634 (100%)

4.2. **Analysis of labor data in Japan.** Table 5 presents the sentiment analysis results categorizing opinions regarding working in Japan on YouTube. From Table 5, it was observed that more than 60% of the collected comments on YouTube expressed emotional or subjective feelings towards working in Japan. The overall sentiment was slightly positive (23.07%) compared to negative (15.43%). However, examining each sub-point made it apparent that most opinions leaned toward neutrality. It was attributed to the prevalence of question texts or narrative sentences that did not express explicit positive or negative views, which prominently featured phrases such as "อย่างไร" which means "How?" (commonly found at the end of Thai question sentences). Moreover, when focusing solely on positive or negative opinions, it was evident that opinions regarding preparation and law, lifestyle, and working leaned more toward the positive side.

Table 5. Results of sentiment analysis in Japan

Comment type	**Preparation and law**	**Lifestyle**	**Working**	**All**
Positive	386 (18.91%)	**1,384 (23.08%)**	1,395 (22.19%)	1,720 (23.07%)
Negative	288 (14.11%)	964 (16.08%)	1,016 (16.16%)	1,151 (15.43%)
Neutral	**1,367 (66.98%)**	**3,648 (60.84%)**	**3,876 (61.65%)**	**4,586 (61.50%)**
Total	2,041 (100%)	5,996 (100%)	6,287 (100%)	7,457 (100%)

4.3. **Analysis of labor data in South Korea.** According to the data presented in Table 6, the overall sentiment toward working in South Korea, as reflected in the comments on YouTube, was largely neutral, comprising 72.81% of the total comments. Positive opinions accounted for 18.27% of the comments, while negative opinions comprised 8.92%. When examining each sub-point individually, it was observed that the majority, more than 70%, of the opinions remained neutral. This pattern aligns with the findings from Japan, where many comments were in the form of questions. However, when focusing specifically on positive and negative opinions, it became apparent that there were more positive sentiments than negative ones across the categories of preparation and law, lifestyle, and working.

Table 6. Results of sentiment analysis in South Korea

Comment type	**Preparation and law**	**Lifestyle**	**Working**	**All**
Positive	678 (15.76%)	**880 (20.97%)**	1,257 (18.25%)	1,632 (18.27%)
Negative	516 (11.99%)	251 (5.98%)	691 (10.04%)	797 (8.92%)
Neutral	**3,108 (72.25%)**	**3,066 (73.05%)**	**4,938 (71.71%)**	**6,506 (72.81%)**
Total	4,302 (100%)	4,197 (100%)	6,886 (100%)	8,935 (100%)

4.4. **Analysis of labor data in Taiwan.** The results in Table 7 indicate that a significant majority, over 78%, of the comments collected from YouTube expressed emotional sentiments. Regarding the overall opinion on working in Taiwan, it leaned towards the neutral range, with a positive sentiment of 14.47% and a negative sentiment of 6.77%. When examining the sub-issues, it was observed that each aspect yielded diverse opinions, primarily falling within the neutral range, often represented through question texts. However, when focusing solely on positive and negative sentiments, it became apparent that all three aspects, namely preparation and law, lifestyle, and working, received predominantly positive opinions.

4.5. **Analysis of labor data in the USA.** The analysis of Thai people's opinions on working in the United States, as presented in Table 8, revealed that a significant majority, over 60 percent, expressed neutral sentiments. Positive opinions accounted for 27.33 percent, while negative opinions represented 12.59 percent. When focusing solely on

Table 7. Results of sentiment analysis in Taiwan

Comment type	Preparation and law	Lifestyle	Working	All
Positive	268 (14.36%)	**290 (14.60%)**	225 (16.12%)	483 (14.47%)
Negative	156 (8.36%)	132 (6.64%)	89 (6.37%)	226 (6.77%)
Neutral	**1,442 (77.28%)**	**1,565 (78.76%)**	**1,082 (77.51%)**	**2,629 (78.76%)**
Total	1,866 (100%)	1,987 (100%)	1,396 (100%)	3,338 (100%)

Table 8. Results of sentiment analysis in the USA

Comment type	Preparation and law	Lifestyle	Working	All
Positive	1,092 (22.05%)	**1,973 (29.87%)**	1,936 (29.15%)	**2,381 (27.33%)**
Negative	626 (12.64%)	810 (12.26%)	815 (12.27%)	1,097 (12.59%)
Neutral	**3,234 (65.31%)**	**3,823 (57.87%)**	**3,890 (58.58%)**	**5,235 (60.08%)**
Total	4,952 (100%)	6,606 (100%)	6,641 (100%)	8,713 (100%)

positive and negative sentiments, it became evident that opinions regarding preparation and law, lifestyle, and working leaned towards the positive rather than the negative. Lifestyle received the highest positive opinions among these areas.

5. **Discussion.** The examination of YouTube comments on videos about Thai individuals working overseas in Australia, Japan, South Korea, Taiwan, and the United States yielded noteworthy insights. Data illustrated in Figure 3 indicates that the majority of the opinions extracted from these video comments were of a neutral stance, exceeding 60 percent. These predominantly neutral comments were characterized by interrogative or questioning tones. This pattern suggests a prevalent feeling of uncertainty or indecision among the Thai audience concerning the prospect of working abroad.

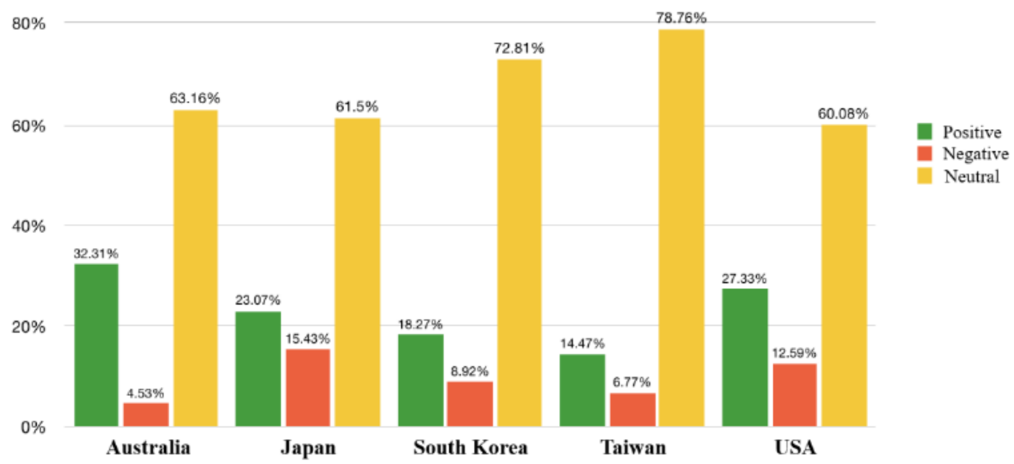

Figure 3. Comparison of the results of the overall analysis of comments

Figure 3 shows that views towards Thai people working abroad are most favorable in Australia (32.31%) and the USA (27.33%) when considering only the positive sentiment. Conversely, Japan (15.43%) has the highest level of negative sentiment, indicating that there may be more challenges or issues expressed in the videos or the viewer comments regarding Thai workers in that country. The lower negative sentiment in Australia (4.53%) and Taiwan (6.77%) suggests a more favorable or less problematic view of Thai workers in these regions.

The findings suggest that Thai viewers' opinions on working abroad vary depending on the destination country. The prevalence of interrogative texts indicates a need for more clarity and information regarding working conditions and experiences. These insights contribute to a better understanding of Thai workers' perspectives and can assist policymakers and stakeholders in addressing concerns and improving conditions related to working abroad.

6. **Conclusions.** In conclusion, this study applied machine learning techniques to analyzing the sentiments of Thai workers abroad in five countries: Australia, Japan, South Korea, Taiwan, and the United States. The dataset comprised 37,077 comments from 400 YouTube videos, focusing on three key aspects of working abroad: preparation and law, lifestyle, working, and overall experience. Various techniques in Python were used to classify these comments into positive, negative, and neutral categories. The Naïve Bayes Support Vector Machine (NBSVM) algorithm outperformed other models in sentiment classification, with results visualized through percentages, bar charts, and word clouds.

When compared with previous studies, such as Sharma et al. [2], which reviewed sentiment analysis techniques for labor-related issues using social media data, this study aligns with the existing literature in its use of machine learning for analyzing sentiment from unstructured data. However, this study differentiates itself by focusing specifically on Thai laborers and leveraging the NBSVM model, which has shown improved performance over traditional approaches like Naïve Bayes (NB) and Support Vector Machine (SVM) alone. While previous research has demonstrated the utility of both NB and SVM in various contexts, the NBSVM combination used in this study resulted in higher classification accuracy, reflecting the effectiveness of hybrid models in sentiment analysis tasks [5].

The analysis revealed that Australia received the highest proportion of positive opinions (32.31%), indicating a generally favorable sentiment towards working in the country, while also having the lowest negative sentiments. Conversely, Japan recorded the highest negative sentiment (15.43%) across multiple aspects compared to the other four countries.

This study underscores the value of machine learning methodologies for analyzing large social media datasets, providing critical insights into Thai workers' perceptions of overseas employment. The findings have significant implications for shaping labor migration policies and improving market predictions within this sector. A deeper understanding of Thai laborers' sentiments will better equip policymakers and stakeholders to improve working conditions and the overall well-being of labor migrants.

For future research on the sentiments of Thai workers abroad, several recommendations are suggested. First, it is crucial to clearly define the research parameters, including the specific types of data, the required volume, and potential data sources, to ensure efficient and targeted data collection. Second, for a more straightforward analysis, it may be beneficial to categorize sentiments into positive and negative only, allowing for more direct insights into workers' opinions. Finally, continuous updates to analytical models will be necessary to adapt to evolving language use, ensuring that the sentiment analysis remains both accurate and relevant over time. These steps will be essential in enhancing the assessment of Thai workers' experiences abroad and informing future policymaking and labor migration program development.

REFERENCES

[1] Y. Kang, Z. Cai, C.-W. Tan, Q. Huang and H. Liu, Natural language processing (NLP) in management research: A literature review, *Journal of Management Analytics*, vol.7, no.2, pp.139-172, 2020.

[2] D. Sharma, M. Sabharwal, V. Goyal and M. Vij, Sentiment analysis techniques for social media data: A review, *The 1st International Conference on Sustainable Technologies for Computational Intelligence*, pp.75-90, 2020.

[3] C. Haruechaiyasak, A. Kongthon, P. Palingoon and K. Trakultaweekoon, S-Sense: A sentiment analysis framework for social media sensing, *Proc. of the IJCNLP 2013 Workshop on Natural Language Processing for Social Media (SocialNLP)*, pp.6-13, 2013.
[4] W. Lu, J. Li, J. Wang and L. Qin, A CNN-BiLSTM-AM method for stock price prediction, *Neural Computing and Applications*, vol.33, no.10, pp.4741-4753, 2021.
[5] A. N. Muhammad, S. Bukhori and P. Pandunata, Sentiment analysis of positive and negative of YouTube comments using Naïve Bayes Support Vector Machine (NBSVM) classifier, *2019 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE)*, pp.199-205, 2019.
[6] A. Mittal and S. Patidar, Sentiment analysis on twitter data: A survey, *Proc. of the 7th International Conference on Computer and Communications Management*, pp.91-95, 2019.
[7] Q. Wang, L. Sun and Z. Chen, Sentiment analysis of reviews based on deep learning model, *2019 IEEE/ACIS 18th International Conference on Computer and Information Science (ICIS)*, pp.258-261, 2019.
[8] U. Mishra, S. Gautam, W. J. Riley and F. M. Hoffman, Ensemble machine learning approach improves predicted spatial variation of surface soil organic carbon stocks in data-limited northern circumpolar region, *Frontiers in Big Data*, vol.3, 528441, 2020.
[9] S. Kaur, S. Singh and S. Kaushal, Abusive content detection in online user-generated data: A survey, *International Conference on Arabic Computational Linguistics*, vol.189, pp.274-281, 2021.
[10] W. A. Qader, M. M. Ameen and B. I. Ahmed, An overview of bag of words; importance, implementation, applications, and challenges, *2019 International Engineering Conference (IEC)*, pp.200-204, 2019.
[11] P. Chormai, P. Prasertsom and A. Rutherford, AttaCut: A fast and accurate neural Thai word segmenter, *arXiv Preprint*, arXiv: 1911.07056, 2019.
[12] P. Bowornlertsutee and W. Paireekreng, The model of sentiment analysis for classifying the online shopping reviews, *Journal of Engineering and Digital Technology (JEDT)*, vol.10, no.1, pp.71-79, 2022.
[13] Y. Jin, Development of word cloud generator software based on Python, *Procedia Engineering*, vol.174, pp.788-792, 2017.
[14] R. Ardianto, T. Rivanie, Y. Alkhalifi, F. S. Nugraha and W. Gata, Sentiment analysis on E-sports for education curriculum using Naïve Bayes and support vector machine, *Jurnal Ilmu Komputer dan Informasi*, vol.13, no.2, pp.109-122, 2020.
[15] M. Ahmad, S. Aftab and I. Ali, Sentiment analysis of tweets using SVM, *International Journal of Computer Applications,* vol.177, no.5, pp.25-29, 2017.
[16] S. Xu, Bayesian Naïve Bayes classifiers to text classification, *Journal of Information Science,* vol.44, no.1, pp.48-59, 2018.
[17] W. S. Noble, What is a support vector machine?, *Nature Biotechnology,* vol.24, no.12, pp.1565-1567, 2006.
[18] D. A. Pisner and D. M. Schnyer, Support vector machine, in *Machine Learning*, Elsevier, 2020.