Comparison Study on Sentiment Analysis Using Lexicon for Airlines Using Supervised Methods

NURUL IZZA MOHD JOHARI[1], SOFIANITA MUTALIB[2*], NURUL NADZIRAH MOHD HASRI[2], MUHAMMAD ARDIANSYAH SEMBIRING[3]

[1] Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450, Shah Alam, Selangor, MALAYSIA

[2] School of Computing Sciences, College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, 40450, Shah Alam, Selangor, MALAYSIA

[3] STMIK Royal, Jl. Prof.H.M. Yamin No.173, Kisaran Naga, Kec. Kota Kisaran Timur, Kabupaten Asahan, Sumatera Utara, INDONESIA

*Corresponding author

Abstract: Nowadays, sentiment analysis usually uses social media websites such as Twitter to analyse the public's opinion on a particular topic. Users have unrestricted access to this website and can express their opinions freely without any restrictions, and it is well-known that opinions influence readers. Therefore, the main objective of this research is to identify the public's positive, negative, and neutral attitudes towards airlines such as Malaysian Airlines, Air Asia, and Malindo Air. Two approaches are adopted: the lexicon-based approach to label the tweets and the machine learning approach such as Naïve Bayes, SVM, and Deep Learning to predict and compare the performance. A total of 35,005 tweets from airlines with all three keywords were evaluated. Deep Learning achieved the highest accuracy and f1 score with 74.10% and 73.49%, respectively. The results show that Deep Learning outperforms the other classifiers by having the highest precision and f1 score. Finally, the sentiment analysis results are visualized in a dashboard to enable a more accurate research analysis. For future work, the dashboard could be integrated into a web-based dashboard to be published for the public and not only for airlines.

Keywords: Airlines, Lexicon-based, Machine Learning, Sentiment Analysis, Twitter.

Received: April 24, 2023. Revised: April 29, 2024. Accepted: May 25, 2024. Published: June 25, 2024.

1. Introduction

Microblogging websites have become a source of a wide range of information due to the nature of microblogs, where individuals post their thoughts on various topics in real time, discuss current issues, offer criticism, and express positive emotions about things they use every day. Millions of travellers share their opinions and viewpoints on airlines, facilities, and services provided on social media platforms such as Twitter, Facebook, and blogs due to the advancement of microblogging [1]. Twitter is another of the world's most successful microblogging services and is seen as a timely source of information and newsfeed rather than an online networking forum. The vast amount of material created and shared by people on Twitter, from individuals to organizations, creates new research opportunities in various fields, including media and communication studies, sociology, psychology, linguistics, political science, and computer science [2][3]. Customer satisfaction is the most crucial component that every airline analyses to ensure that its services are pleasant for passengers. According to data from Pandya and his colleagues, in 2020, about 81% of internet users researched online at least once about the product they are interested in [4]. This shows that public opinion and reviews are very important for companies today, as they strongly impact their sales. This is where the role of sentiment analysis comes into play. With the explosion of the internet and microblogging, millions of reviews and opinions are generated by internet users daily. Companies, therefore, need means to monitor the performance of their products or services in the market. Consequently, they use sentiment analysis to track how the community reacts to their products and services [5]. Sentiment analysis is a natural language pre-processing (NLP) technique that enables researchers to determine the general public's opinions through textualization. It has proven to be a captivating area of research as it is much more effective in capturing the general public's sentiments, especially in the airline industry, where people are easily satisfied or dissatisfied and often express their feelings on Twitter [6]. It consists of many approaches: abstract mining, fine-grained sentiment analysis, emotion recognition, and multilingual sentiment analysis [5].

Air transport is one of the fastest ways to travel worldwide. Malaysia has renowned airlines such as Air Asia, Malaysian Airlines, and Malindo Air. With each airline trying to provide its customers with the best possible facilities and services, competition in the airline industry is fierce. Each airline strives to make more profit and avoid any possibility of losses. The most crucial factor that contributes to making profits is customer satisfaction. In [7], Onat stated that the airline industry is confronted with various aspects, including passengers' needs, wants, and comfort.

The rest of this paper is structured as follows: Section 2 provides the related works, and Section 3 presents the literature about predictive models. Section 4 introduces the methods and dataset used in this study. The results and findings are given in Section 5. Lastly, Section 6 concludes the study.

2. Related Studies

Before the emergence of the World Wide Web, which is now widespread, many of us asked for information and recommendations regarding almost anything from our friends and acquaintances before really proceeding to buy or commit to something. People usually have to gain experience using the stuff and can only evaluate whether it is worth committing to or vice versa. Moreover, the reviews they acquired from friends are most likely biased, so they need to determine whether the reviews are authentic, work well for them, and are justifiable. However, with the emergence of the Internet and the Web, people can easily scrape information and data from strangers and gather that information and reviews to evaluate whether they are worth having. According to Liu in 2012 [8], about 81% of internet users do online research on the product they are interested in at least once.

This shows that public opinion and reviews are crucial to business companies nowadays as they affect their sales heavily. This is where the sentiment analysis role comes in. With the explosion of the Internet and microblogging, millions of reviews and opinions are generated daily by Internet users. Businesses need mediums to monitor the performance of their products or services in the market. Thus, they use sentiment analysis to track how the community responds to the products and services [9]. Sentiment analysis is a part of NLP that determines texts and classifies them as positive, negative, and neutral. Sentiment analysis, also known as mining of opinion, examines people's feelings and sentiments about a particular individual, service, topic, or product [8]. Liu Hu's method is simpler, generating a single sentiment integer output [10]. However, the sentiment integer label is not represented as sentiment classification.

Due to the rapid expansion of social interaction on the Internet, sentiment analysis can be used for decision-making, people and businesses. There are several discussions and evaluations regarding the products and services; people do not have to seek recommendations and reviews from friends and acquaintances, which can be biased. Companies do not need to develop needless surveys because of the existence of information that can be scraped online [7][8]. The proliferation of multiple sites has been observing opinions, and detecting the sentiments on the Internet, and screening the information in the comments and reviews.

Several unsupervised learning techniques first attempt to create an unsupervised sentiment lexicon and then identify a text unit's degree of positivity or subjectivity through some functions based on the positive and negative indicators within, as determined by the lexicon. The approach entirely relies on lexical resources concerned with mapping the words to a polarity of categorical sentiments. Furthermore, the lexicon-based method requires no training data and relies simply on dictionaries. Nevertheless, the lexical dictionary’s boundary is that not all words in the sentiments could be assigned a value [11].

On the other hand, in [12], Drus & Khalid stated that the use of a lexicon has its advantages, such as the ability to classify positive and negative terms more straightforwardly, the flexibility to deal with multiple languages, and the faster speed with which the analysis may be completed. They did a comparative study on the techniques used in sentiment analysis of social media by other researchers. Another advantage of the learning-based method is its capacity to modify and create trained models for a specific purpose, even though this method can be costly and time-consuming for some tasks [13]. In order to produce the sentiment score, Hu Lui and Vader's methods are mostly applied in sentiment analysis [14][15].

In [16], Gulati et al. did a comparative analysis of tweets by implementing machine learning algorithms. They acknowledged that machine learning is beneficial in NLP work and was extensively employed in this study. This is because, with ML, computers can learn cognitive behaviors such as forecasting decision-making. On the contrary, Chandra and Jana claim that deep learning has shown outstanding performance compared to machine learning algorithms [17]. However, Dhaoui et al. found that both ML approaches produced accuracy results almost as good as deep learning methods [18]. Nonetheless, the classification ensembles between them differed significantly. Other than that, Amin et al. developed an intelligent model to identify the COVID-19 pandemic in Twitter posts using standard machine learning-based techniques such as SVM, Naive Bayes, Logistic Regression, and others, with the aid of the term frequency-inverse document frequency (TF-IDF), [19]. The results of the experiments show that the proposed 24 approach is promising in detecting the COVID-19 pandemic in Twitter messages, with overall accuracy, precision, recall, and F1 score between 70% and 80% and the confusion matrix for machine learning approaches with the TF-IDF feature extraction technique.

Twitter becomes a common platform for business organizations which offer services, for getting customers' feedback, reviews and comments. Though, processing the textual data needs to be done with appropriate machine learning methods due its complexity. Some studies had performed the sentiment analysis in airlines data set, such as Gupta and Bhargav [20] and Li et al, uses the Kaggle datasets with BERT and variants [21]. The success of these studies had given rise to the motivation to proceed with local airlines and come out with meaningful analysis. Table 1 shows the summarization of machine learning techniques.

Table 1. Machine learning algorithms used in building classification models.

Reference	Description	Techniques and results
[11]	An improved lexicon-based analysis that aggregates the sentiment values of positive and negative words within a message.	Techniques: Lexicon sentiment analysis algorithms labelled as L, LN, LNS, LNW, and LNWS. Results: LNW achieved highest accuracy on Stanford Twitter dataset (77.3%) and LNSW achieved highest accuracy on Stanford IMDB dataset (74.2%).
[12]	This paper is a report of a review on sentiment analysis in social media that explored the methods, social media platform used, and its application.	Techniques: Lexicon-based and machine learning. Results: Drus and Khalid found that researchers argue that both lexicon-based and machine learning techniques has similar performance in terms of accuracy.
[16]	This research conducts sentiment analysis by using seven popular machine learning techniques.	Techniques: Passive-aggressive classifier, Linear SVC, Multi-Nomial Naïve Bayes, Bernoulli Naïve Bayes, Logistic Regression, Ada Boost Classifier, and Perceptron.

		Results: Linear SVC obtained the highest accuracy (98.7%).
[17]	This research aims to improve percentage accuracy of classifier models while comparing the performance of both machine learning and deep learning models.	Techniques: Machine learning classifiers (Naïve Bayes, MNB Classifier, Bernoulli Classifier, Logistic Classifier, Stochastic Gradient Descent, Linear SVC, and NuSVC Classifier) and deep learning models (CNN-LSTM and LSTM) Results: Deep learning models obtained better classification results compared to machine learning models.
[18]	The research conducts sentiment analysis with both machine learning through separate and combining both approaches.	Techniques: Various machine learning approaches including maximum entropy, random forests, SVM, bagging, and decision tree, and lexicon-based approaches. Results: F-score is at 0.77 for lexicon-based approach and 0.78 for machine learning for positive valence classification.
[19]	The study evaluated COVID-19 related tweets with five different machine learning approaches.	Techniques: Machine learning (SVM, Naïve Bayes, Logistic Regression, Decision Tree, Random Forest) Result: SVM obtained the best precision (80%), recall (81%), and F1-score (81%) values.

3. Methodology

3.1 Data Collection

The first phase of this research is to collect data on the area selected for analysis. The preferred source of data for this research is the Twitter website. The total length of characters allowed is up to 240 characters. They use a combination of emoticons, acronyms, and sarcasm while expressing themselves in the messages. This research uses the Python library snscrape to extract the data from Twitter. The keywords used in the data collection are "malaysian airlines", "air asia" and "malindo air". The timeline ranges from January 2018 to March 2022 and includes about 13,000 tweets for each keyword. A total of 39,909 tweets were evaluated.

3.2 Pre-processing

The second phase is called the pre-processing phase of the data. This step is important because the raw data contains noise, such as emoticons, RTs, and hashtags, which are irrelevant to the analysis. The pre-processing phase includes six steps, which are shown in Figure 1.

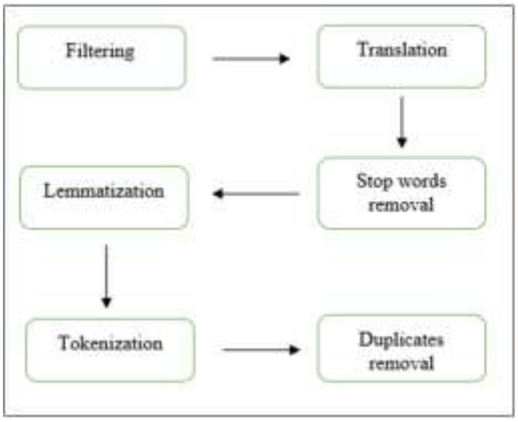

Figure 1. Flow of data pre-processing techniques

3.2.1 Filtering

Filtering is a form of cleaning noise from raw data. Data such as URL links, retweet counts, usernames, hashtags, etc., are removed from the dataset. Letters are also standardized by converting them to lowercase. The final steps in filtering the tweets are the removal of punctuation marks and special characters.

3.2.2 Translation

Before further pre-processing techniques are carried out, the scraped tweets are only partially in English. Instead of removing the non-English tweets, they are translated into English using a Python library called Googletrans, which implements the Google Translate API.

3.2.3 Stop Words Removal

Stop words are words that contribute nothing to the meaning of a sentence. Therefore, they can be safely omitted without jeopardizing the sense of the sentence. Examples are the, a, he, she, has, have, etc. We will improve the model's performance by removing the meaningless term from the model's evaluation.

3.2.4 Lemmatisation

Lemmatisation is a method of converting words into their basic forms, considering the context of thought. Lemmatisation allows an accurate calculation and analysis of the frequency of the root word used in the data set. For example, the word "playing" is converted into its root word "play" after lemmatisation.

3.2.5 Tokenization

Using tokenisation, the tweets were divided into words, known as tokens. Word tokenisation, character tokenisation and subword tokenisation (n-gramme characters) are the broad categories into which tokenisation can be broadly divided. In this study, word tokenisation is the main focus. Each word was tokenised after a space.

3.2.6 Duplicates Removal

After the above processes are completed, the data duplication is removed. This is because some tweets contain repetitions of the original message, which would confuse the sentiment analysis algorithm. However, RapidMiner is used to handle this process separately.

3.3 Data Labelling

The target variable is described as a data label. The polarity score of the tokens is determined based on the three classes: positive, negative, and neutral. A lexicon-based approach is used in this procedure. Since this research implements modelling using supervised learning classifiers, a labelled dataset is crucial to obtain more detailed results. RapidMiner implements this process. Labelling thousands of data with RapidMiner reduces the workload instead of labelling them manually. The chosen lexicon dictionaries are VADER and SentiWordNet. The sentiment polarity for negative words is labelled as -1, and the polarity score for positive words is labelled as +1. Meanwhile, the polarity score for neutral words is labelled as 0.

3.4 Modelling

In the modelling phase, three supervised learning classifiers are used to compare the models' performance and the predicted sentiments' accuracy. The classifiers chosen are Support Vector Machine (SVM), Naïve Bayes, and Deep Learning. The dataset is split into a training dataset and a test dataset with three ratios to be compared, namely 70:30, 80:20, and 90:10. The performance of these models is compared to determine which combination achieves the highest accuracy.

In our study, we used to observe the accuracy to evaluate the classifiers. When applied to data, accuracy is seen as a model's correctness. A confusion matrix is required to choose the best model between SVM, Naïve Bayes, and Deep Learning. True Positive (TP) are positive subjects that have been correctly labelled as positives, False Positive (FP) are negative subjects that have been incorrectly labelled as positives, True Negative (TN) are negative subjects that have been correctly labelled as negative, and False Negative (FN) are positive subjects that have been incorrectly labelled as negative. Next, the accuracy of the classifiers is calculated based on the formula:

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

3.5 Dashboard Development

The development of a dashboard follows the modelling phase to visualise the sentiment analysis results by selecting appropriate charts and graphs to make the results more meaningful and easier to interpret. The dashboard is developed using a business analytics service application from Microsoft called Power BI.

4. Results and Discussion

After lemmatisation and tokenisation are completed in the preprocessing phase, data duplicates are removed separately with RapidMiner. After duplicate removal, 28,891 tweets remain out of 35,003 tweets. After the dataset has been cleaned, RapidMiner is used to determine the polarity score of the sentiments. This process identifies whether the tweets are positive, negative, or neutral. This labelled dataset is needed for later data modelling. VADER and SentiWordNet are the models used to determine the sentiment polarity.

Table 2 shows the total number of sentiments categorised into positive, negative, and neutral classes. Again, the VADER method provides more stability, and each label is almost balanced, while the SentiWordNet approach produces unbalanced total sentiments across the classes. Therefore, both approaches are used to compare the modelling performance later.

Table 2. Count of sentiment labels using VADER and SentiWordNet

Sentiment	Total count (VADER)	Total count (SentiWordNet)
Positive	10, 286	17, 902
Negative	9, 615	8, 824
Neutral	8, 990	2,165

After labelling the sentiments, the next step is to perform the classification task using machine learning classifiers through RapidMiner. The classifiers selected are SVM, Naïve Bayes, and Deep Learning. The models are compared based on their performance using different parameters. The parameters to be changed are the percentage split of the training and test dataset, which is 70:30, 80:20, and 90:10. Each classifier is tested with two different datasets with different lexicon dictionaries: VADER and SentiWordNet approach. All results are shown in Table 3.

Table 3. Summary of classification performance

ML Classifier	Lexicon Approach	Percentage Split	Accuracy	Precision	Recall	F1-Score
Support Vector Machine	VADER	70:30	62,13	69,62	58,08	63,33
		80:20	64,10	72,48	59,96	65,63
		90:10	63,18	73,32	58,87	65,31
		Avg	**63,14**	**71,81**	**58,97**	**64,76**
	SentiWord	70:30	61,76	68,08	37,55	48,4
		80:20	62,01	73,32	37,98	50,04
		90:10	61,48	67,55	37,61	48,32
		Avg	**61,75**	**69,65**	**37,71**	**48,92**
Naïve Bayes	VADER	70:30	69,05	68,53	67,53	68,03
		80:20	70,07	69,77	68,29	69,02
		90:10	69,34	69,08	67,43	68,25
		Avg	**69,49**	**69,13**	**67,75**	**68,43**
	SentiWord	70:30	51,89	53,13	60,99	56,79
		80:20	51,28	53,1	61,69	57,07
		90:10	51,74	52,65	60,85	56,45
		Avg	**51,64**	**52,96**	**61,18**	**56,77**
Deep Learning	VADER	70:30	71,64	71,55	69,99	70,76
		80:20	73,03	73,49	70,92	72,18
		90:10	74,10	74,02	72,96	73,49
		Avg	**72,92**	**73,02**	**71,29**	**72,14**
	SentiWord	70:30	62,49	53,72	44,08	48,42
		80:20	57,89	49,88	45,72	47,71
		90:10	60,32	55,32	43,57	48,75
		Avg	**60,23**	**52,97**	**44,46**	**48,29**

Eighteen models were compared based on their performance. The box emitting yellow represents the highest accuracy achieved by each classifier, while the box emitting green represents the highest F1 score achieved by each classifier. From the above summary, VADER is the best lexicon approach for this project. This is because VADER has the highest average accuracy and F1 score for each classifier. On the other hand, Deep Learning is the best classifier for this project according to the table above. The Deep Learning classifier achieved the highest accuracy among all classifiers with an accuracy of 74.10% with an average of 72.92% and an F1 score of 73.49% with an average of 72.14% using the VADER approach and a split of 90:10 percent. Figure 3 below shows an example of the dashboard that was created. Figure 3 below shows an example of the dashboard created.

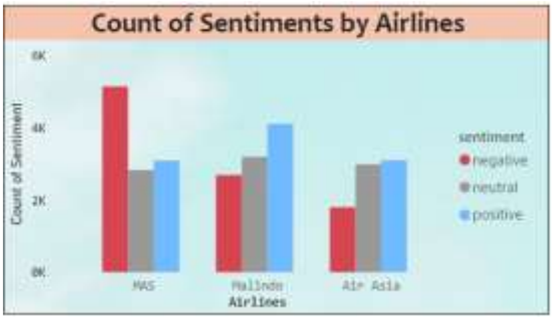

Figure 2. Count of sentiment labels by airlines

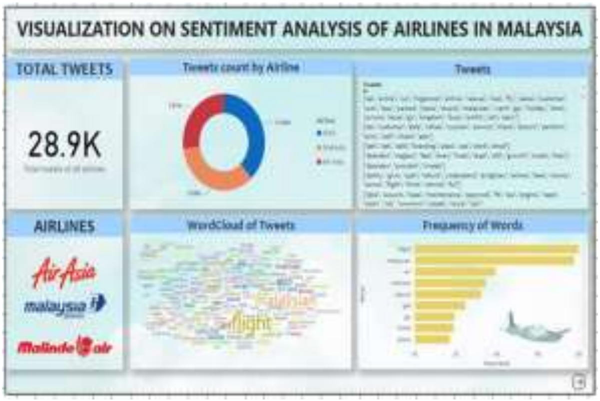

Figure 3. Dashboard

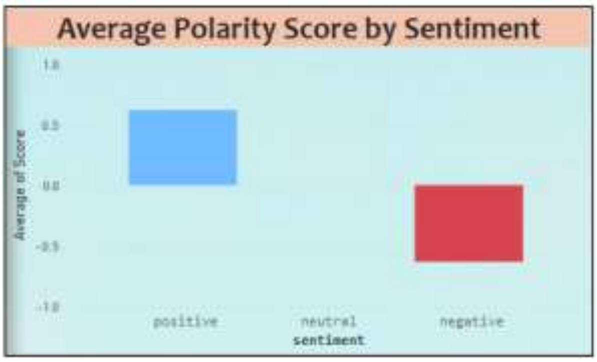

Figure 4. Average Polarity Score by Sentiment

Figure 4 shows the number of sentiment labels for each airline. The display changes depending on the filtering done on the left side of the second section of the dashboard. In this way, users can compare the performance of each airline based on the number of sentiments. From the above chart, it can be seen that MAS has far more negative tweets than positive ones, while Malindo and AirAsia have received slightly more positive tweets than the other two sentiments. It can be assumed that MAS has more negative tweets due to the infamous tragedy of the disappearance of MH370, with some closely linking the case to a political conspiracy and so on. Despite Malindo having a slightly higher number of positive tweets, AirAsia has the best reputation as it has the lowest number of negative tweets.

5. Conclusion

In summary, the model using VADER and the Deep Learning classifier with a 90:10 percentage split has the highest accuracy of 74.10 percent and an f1 score of 73.49 percent. The sentiment analysis results are visualised in a dashboard developed with Power BI. The visualisation consists of interactive filtering that changes the result based on the selected filters. This project's results are presented to be easily analysed using the visualisation. From the visualization result, AirAsia has the best reputation as it has the second highest positive reviews (3,095) and the lowest negative reviews (1,795).

Moreover, the developed models in this study can be expanded to build an application with the following advantages:

- for individuals to view the airline analysis based on public opinions.
- for organizations to identify satisfaction factors and extract the sentiments based on the factors.

The research scope can be further improved by including more than three airline companies. The data can be enhanced by scraping through a more reliable site like TripAdvisor, which has more meaningful reviews. This project is also recommended to implement aspect-based sentiment analysis, which extracts topics from the tweets and determines their polarity. This way, the project would have deeper understanding about the sentiments and can produce various visualisation results for the users to get more valuable insights.

Acknowledgement:

The authors would like to express their gratitude to the College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia for the research support. The authors also thank Muhammad Umarul Aiman Bin Mohamad Zulhilmi for their assistance.

References:

[1] Sreeja, I., Sunny, J. V., & Jatian, L, *Twitter sentiment analysis on airline tweets in India using R language. In Journal of Physics: Conference Series* (Vol. 1427, No. 1, p. 012003). IOP Publishing., 2020.

[2] Edirisinghe, S., (2020). *Effectiveness of Twitter as a Social Media Platform Used by Starbucks - United States,* ResearchGate, 2020. [Online]. Available: https://www.researchgate.net/publication/338487623_EFFECTIVENESS_OF_TWITTER_AS_A_SOCIAL_MEDIA_PLATFORM_USED_BY_STARBUCKS_-UNITED_STATES

[3] Dikiyanti, T. D., Rukmi, A. M., & Irawan, M. I. (2021, March). *Sentiment analysis and topic modeling of BPJS Kesehatan based on twitter crawling data using Indonesian Sentiment Lexicon and Latent Dirichlet Allocation algorithm. In Journal of Physics: Conference Series*, (Vol. 1821, No. 1, p. 012054). IOP Publishing, 2021, March.

[4] Pandya, Sharnil & Mehta, Pooja. (2020). A Review on Sentiment Analysis Methodologies, Practices and Applications.

[5] Alsaeedi, A., & Khan, M. Z. (2019). *A study on sentiment analysis techniques of Twitter data. International Journal of Advanced Computer Science and Applications*, 10(2), 362-374, 2019.

[6] de Melo, T., & Figueiredo, C. M, *Comparing news articles and tweets about COVID-19 in Brazil: sentiment analysis and topic modeling approach. JMIR Public Health and Surveillance*, 7(2), e24585, 2021.

[7] Onat Kocabiyik, O. (2021). *Social Media Usage Experiences of Young Adults during the COVID-19 Pandemic through Social Cognitive Approach to Uses and Gratifications.* International Journal of Technology in Education and Science, 5(3), 447-462.

[8] Liu, B., *Sentiment Analysis and Opinion Mining. Morgan & Claypool*, 2012. https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf

[9] *Pang, B., & Lee, L., Opinion mining and sentiment analysis. Computational Linguistics*, 35(2), pp. 311–312, 2009. https://doi.org/10.1162/coli.2009.35.2.311

[10] M. Hu and B. Liu, *"Mining opinion features in customer reviews," in AAAI'04: Proceedings of the 19th national conference on Artifical intelligence*, San Jose California 25 - 29 July 2004: AAAI Press.

[11] Jurek-Loughrey, A., Mulvenna, M., & Bi, Y., *Improved lexicon-based sentiment analysis for social media analytics. Security Informatics*, 4. https://doi.org/10.1186/s13388-015-0024-x

[12] Drus, Z., & Khalid, H., *Sentiment Analysis in Social Media and Its Application: Systematic Literature Review. Procedia Computer Science*, 161, pp. 707–714, 2019. https://doi.org/10.1016/j.procs.2019.11.174

[13] Sarlan, A., Nadam, C., & Basri, S., *Twitter sentiment analysis. Proceedings of the 6th International Conference on Information Technology and Multimedia*, pp. 212–216, 2014. https://doi.org/doi:10.1109/ICIMU.2014.7066632

[14] Mahmud, Y., Shaeeali, N.S., & Mutalib, S., *Comparison of Machine Learning Algorithms for Sentiment Classification on Fake News Detection. International Journal of Advanced Computer Science and Applications*, pp. 665-658, 2021. https://doi.org/10.14569/ijacsa.2021.0121072

[15] C. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," in Proceedings of the International AAAI Conference on Web and Social Media, 2014, vol. 8, no. 1, pp. 216-225.

[16] K. Gulati, S. Saravana Kumar, R. Sarath Kumar Boddu et al., *Comparative analysis of machine learning-based classification models using sentiment classification of tweets related to COVID-19 pandemic*, Materials Today: Proceedings, 2021, https://doi.org/10.1016/j.matpr.2021.04.364

[17] Chandra, Y., & Jana, A., *Sentiment Analysis using Machine Learning and Deep Learning.* 2020 7th International Conference on Computing for Sustainable Global Development (INDIACom), pp. 1–4, 2020.

https://doi.org/10.23919/INDIACom49435.2020.9083703

[18] Dhaoui, C., Webster, C., & Tan, L. (2017). *Social media sentiment analysis: lexicon versus machine learning.* Journal of Consumer Marketing, Vol. 34 No. 6, pp. 480-488. https://doi.org/10.1108/JCM-03-2017-2141

[19] Amin, S., Uddin, M. I., Al-Baity, H. H., Zeb, M. A., & Khan, M. A. (2021). *Machine learning approach for COVID-19 detection on twitter*. Computers, Materials and Continua, 68(2), pp. 2231–2247. https://doi.org/10.32604/cmc.2021.016896

[20] Gupta, N., Bhargav, R. (2023). Sentiment Analysis in Airlines Industry Using Machine Learning Techniques. In: Dutta, P., Chakrabarti, S., Bhattacharya, A., Dutta, S., Shahnaz, C. (eds) Emerging Technologies in Data Mining and Information Security. Lecture Notes in Networks and Systems, vol 490. Springer, Singapore. https://doi.org/10.1007/978-981-19-4052-1_12

[21] Li, Zehong, Chuyang Yang, and Chenyu Huang. 2024. "A Comparative Sentiment Analysis of Airline Customer Reviews Using Bidirectional Encoder Representations from Transformers (BERT) and Its Variants" Mathematics 12, no. 1: 53. https://doi.org/10.3390/math12010053

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.