The Indonesian Journal of Computer Science
www.ijcs.net
Volume 14, Issue 5, October 2025
https://doi.org/10.33022/ijcs.v14i5.4956

Comparative Analysis of Naïve Bayes and K-NN Methods on Social Media Boycott Issue X Case Study: McDonald's

Morra Fatya Gisna Nourielda Azzahra[1], Anik Vega Vitianingsih[2], Dwi Cahyono[3], Anastasia Lidya Manukar[4], Fawaidul Badri[5]
morrafatya@gmail.com[1], vega@unitomo.ac.id[2*], dwikk@unitomo.ac.id[3], almaukar@president.ac.id[4], fawaidulbadri@unisma.ac.id[5]
[1,2,3]Informatic Department, Universitas Dr.Soetomo, Surabaya, Indonesia
[4] Industrial Engineering Department, President University, Bekasi, Indonesia
[5] Informatics Department, Universitas Islam Malang, Indonesia

Article Information

Received : 16 Jul 2025
Revised : 6 Aug 2025
Accepted : 23 Oct 2025

Keywords

Natural Language Processing, Machine Learning, Sentiment Analysis, McDonald's Boycott, Pro-Israel Boycott, Naïve Bayes, KNN

Abstract

The boycott movement against McDonald's, triggered by its alleged support for Israel during the conflict in Gaza, has generated significant public discourse, particularly on the social media platform X (formerly Twitter). This study investigates public sentiment regarding the boycott campaign by analyzing comments and reactions to related content. A total of 1,585 tweets were collected using techniques for web scraping and underwent a comprehensive pre-processing phase, encompassing cleaning, tokenization, filtering, and stemming. Sentiment categories, namely positive, neutral, and negative, are automatically assigned using a lexicon-based technique customized for the Indonesian language. Text data was transformed into numerical form through the Term Frequency-Inverse Document Frequency (TF-IDF) technique, followed by sentiment classification using two supervised machine learning algorithms: Naïve Bayes and K-Nearest Neighbor (K-NN). Evaluation of both models was conducted using a confusion matrix and classification metrics. The results show that the dataset is highly imbalanced, with 93.5% of the tweets labelled as negative, 6.1% as neutral, and only 0.3% as positive. The K-NN model achieved better performance than Naïve Bayes (NB), with an accuracy of 93%, a precision of 31%, a recall of 33%, and an F1-score of 32%. On the other hand, the Naïve Bayes algorithm reached 39% accuracy, 33% precision, 29% recall, and an F1-score of 22%. These findings highlight the dominance of negative sentiment toward McDonald's and demonstrate the efficacy of the K-NN algorithm in sentiment classification in unbalanced datasets. The insights from this study can inform public relations strategies and corporate reputation management in the face of socio-political controversies.

A. Introduction

The current digital era is marked by the rapid increase in the use of social media as a means of public expression. One of the popular platforms is X Apps, which has become the primary space for discussing various social issues [1]. The Israel–Palestine conflict has remained unresolved for decades and has become a controversy that began in 1917 and continues to this day. [2]. Since the outbreak of the conflict between Israel and Hamas on October 7, 2023 [3], many products are considered to support Israel, including McDonald's products, which have been affected [2]. Indonesia, through the MUI Fatwa Number 83 of 2023, urges the public to eschew transactions and the utilization of products associated with Israel, especially from countries that support colonialism and Zionism [4].

The problem in this research is the limited systematic understanding available regarding public views on the boycott calls that are developing on social media X, especially since the escalation of the conflict between Israel and Palestine on October 3, 2023 [5]. This boycott was triggered by the assumption that McDonald's Indonesia is one of the brands that support Israel. Until now, there has been no automated system capable of accurately and efficiently analyzing public opinion on social media X Apps regarding this issue [6]. This research is focused on the implementation and evaluation of sentiment analysis methods on a comment dataset from social media X Apps related to the issue of boycotting pro-Israel products, particularly those of McDonald's Indonesia. Therefore, this research focuses on developing an automated system that can classify public opinion sentiment into positive, negative, and neutral categories[7].

Based on a literature study of previous research that discussed the boycott of pro-Israel brand Starbucks products on X Apps[2], using the NB method, yielded results with an accuracy of 71%. Meanwhile, other researchers who discussed boycotting pro-Israel Unilever Indonesia products on X Apps [8] utilizing the Long Short-Term Memory (LSTM) algorithm method, obtained an 80% result. Based on the literature study, no researchers have specifically compared the performance of the NB and K-NN algorithms in sentiment analysis on the McDonald's boycott issue with the Lexicon-Based labelling method approach [9].

This study aims to implement and evaluate sentiment classification algorithms based on Lexicon-Based, NB, and K-NN approaches in the context of public opinion regarding boycott issues on social media [10]. A system capable of conducting sentiment analysis on public opinion in X Apps regarding the issue of boycotting McDonald's Indonesia by classifying opinions into positive, negative, and neutral classes related to the boycott of pro-Israel brands. The dataset used in this study was obtained from comments on X Apps with the hashtags *#BoikotMcDonaldIndonesia* and *#BoikotProIsrael* [11]. A lexicon-based method is employed to assign sentiment labels, while opinion classification is performed using the NB and K-NN algorithms. This study employs the NB and K-NN methods, which were chosen because they are well-known classifiers with a high level of accuracy. K-NN is used in research because this method employs lazy learner classification, where training data is stored continuously and processed only when test data appears, making it a relatively easy-to-understand method [12]. In contrast, the NB method is used because its performance remains superior when Testing is conducted on categorical data types [13]. Meanwhile, the Lexicon-Based method is utilized in research for opinion classification and sentiment analysis [14].

B. Research Method

This study employs a structured methodology comprising multiple stages to perform sentiment analysis, as illustrated in Figure 1. The process begins with users providing links from social media platform X (formerly known as Twitter), which is then used for crawling to obtain raw data. The collected data then goes through a pre-processing stage to be cleaned and prepared for further analysis. The pre-processing process includes several steps. Initial data preparation involves several stages, including cleaning, converting text to lowercase (case folding), breaking text into tokens, normalizing language, filtering out unnecessary elements, applying stemming, and labelling the sentiment. Afterwards, TF-IDF is used to assign weight to each term for further analysis. This stage aims to extract important features from documents based on the frequency and significance of words in the entire dataset, thereby identifying the most relevant keywords. The results of this weighting process then serve as input for the classification stage, which utilizes two machine learning methods: NB and K-NN. This classification process will yield information regarding model accuracy and the distribution of positive, negative, and neutral sentiments in the analyzed dataset. To enhance clarity, research methods may be supported with visual elements such as tables, figures, or charts. Tables should avoid the use of vertical lines, and horizontal lines are permitted only at the top and bottom of the table.

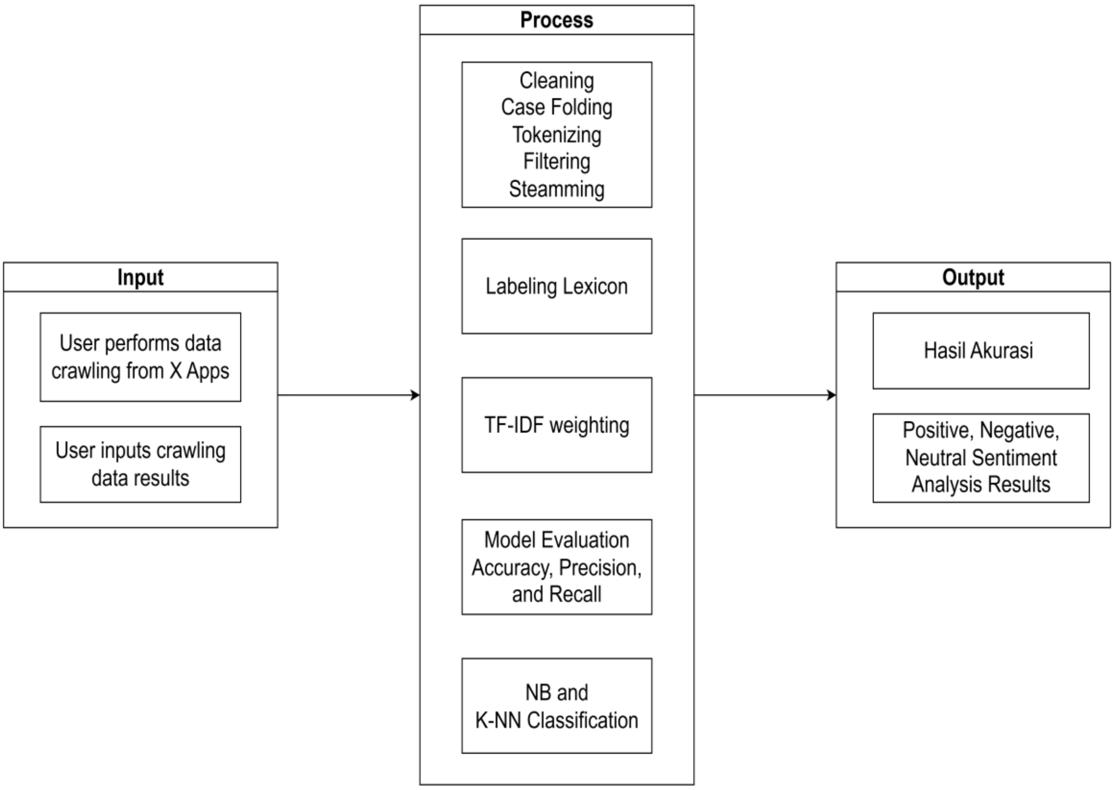

Figure 1. Research Method

1. Dataset

The data for this study were sourced from the social media platform Twitter, which served as the primary source of information. A total of 1,586 tweets were collected using the Twitter API, specifically targeting content relevant to the research topic of the McDonald's boycott. All collected tweets were written in Indonesian and served as the primary corpus for sentiment analysis.

2. Pre – Processing

The pre-processing phase of text analysis involves a set of crucial steps, including cleaning, case folding, tokenization, filtering, stemming, and sentiment labelling. These steps aim to prepare text data for further analysis by making it more structured and ready for use [15]. This stage serves to remove noise, standardize writing formats, identify important information units, simplify word variations, remove irrelevant words, normalize word forms, reduce feature complexity, and assign labels appropriate to the analysis's needs [10].

Step 1: The data cleansing process, which primarily aims to filter data from invalid, incomplete, or irrelevant elements. This stage also addresses duplication and outliers that can affect the quality of interpretation. Thorough cleansing yields a clean and consistent dataset, enhancing analysis accuracy and minimizing errors, as demonstrated in Table 1.

Table 1. Cleaning Data

Dataset	Cleaning Data
Boikot McDonald's nggak bakal ngaruh sih, ini cuma sebatas tren aja.	*Boikot McDonald's tidak akan berpengaruh, ini hanya tren sementara.*

Step 2: Case folding is the process of standardizing text by converting all alphabetic characters to lowercase. This process does not affect non-alphabetic characters and aims to create consistency in text representation. Furthermore, this stage also involves removing elements such as URLs, numbers, and punctuation that are unnecessary for the analysis. Thus, case folding helps simplify text data and ensure a uniform format, as shown in Table 2.

Table 2. Case Folding

Cleaning	Case Folding
Boikot McDonald's nggak bakal ngaruh sih, ini cuma sebatas tren aja.	*Boikot McDonald's ngga bakal ngaruh sih, ini cuma sebatas tren aja.*

Step 3: Tokenization refers to the technique of dividing text into smaller components, known as tokens. Depending on the goals of the text analysis, these tokens may consist of words, phrases, or even individual characters, varying in complexity. As shown in Table 3, the tokenization process also plays a role in word normalization and the removal of irrelevant characters, thereby simplifying text processing and enhancing the accuracy of the analysis results.

Table 3. Tokenizing

Case Folding	Tokenizing
Boikot McDonald's nggak bakal ngaruh sih, ini cuma sebatas tren aja.	*["boikot", "mcdonald's", "nggak", "bakal", "ngaruh", "sih", "ini", "cuma", "sebatas", "tren", "aja"]*

Step 4: Language filtering refers to the process of removing irrelevant or non-target language content from the dataset to ensure consistency and accuracy in text analysis. It standardizes word representation in text to ensure consistency in processing by computing systems. This process involves adjusting the text format and removing irrelevant elements, such as punctuation and common words (filtering), as illustrated in the results presented in Table 4.

Table 4. Filtering

Tokenizing	Filtering
["boikot", "mcdonald's", "nggak", "bakal", "ngaruh", "sih", "ini", "cuma", "sebatas", "tren", "aja"]	*["boikot", "mcdonald", "nggak", "bakal", "ngaruh", "cuma", "sebatas", "tren"]*

Step 5: The stemming process focuses on stripping words of prefixes and suffixes to retrieve their base form. This process helps identify words that share the same root, thereby increasing the efficiency of text analysis. The primary purpose of stemming is to simplify comparisons between words and support more effective meaning extraction. For example, a word like "running" can be simplified to "run" without compromising its meaning. Furthermore, stemming enhances consistency and reduces complexity in language models, as demonstrated in the results presented in Table 5.

Table 5. Steamming

Filtering	Steamming
["boikot", "mcdonald", "nggak", "bakal", "ngaruh", "cuma", "sebatas", "tren"]	["boikot", "mcdonald", "nggak", "bakal", "ngaruh", "cuma", "sebatas", "tren"]

Step 6: Sentiment analysis refers to the task of labelling sentences or words based on the emotions or opinions they convey. The main goal is to classify the text as positive, negative, or neutral, typically using lexicon-based approaches, as shown in Table 6.

Table 6. Labeling

Dataset	Score	Labeling
Boikot McDonald's nggak bakal ngaruh sih, ini cuma sebatas tren aja.	-5	Negative

3. TF-IDF

TF-IDF serves as a crucial method for representing text as numerical vectors, contributing significantly to the performance and reliability of various natural language processing operations [11]. This technique is recognized for its reliability, simplicity in application, and the precision it offers. The TF-IDF approach calculates both Term Frequency (TF) and Inverse Document Frequency (IDF) for each term across a set of documents, enabling the determination of word importance based on its occurrence both locally and globally within the corpus. Subsequently, the labelled text is transformed into a numerical format using the TF-IDF scheme [16]. The TF-IDF is presented in Equations (1), (2), and (3).

$$w\, tf_{t,d} = \{^{1}_{0} + log_{10} tf_{t,d}, if\ tf_{t,d} > 0 \qquad (1)$$

$$idf_{\boldsymbol{t}} = log_{10}(N/df_t) \quad (2)$$

$$idf_{\boldsymbol{t}} = log_{10}(N/df_t) \quad (3)$$

Tf(t,d) signifies how often term t appears in document d, while Idf(t) measures how rare or common the term is across all documents. TF-IDF(t) expresses the computed weight of term t in document d, and n represents the total number of documents analyzed.

4. Naïve Bayes

Naïve Bayes is a widely used classification algorithm in the fields of data analysis and statistics. It is based on Bayes' theorem and assumes that all features used in classification are mutually independent. The model calculates the likelihood of each class for a given instance and assigns the class with the highest probability as the predicted result [17]. The probabilities used to generate the final estimate are calculated based on the total frequency in the "main" decision table.

5. K-Nearest Neighbors

K-NN, a machine learning approach, leverages existing datasets to carry out classification or regression by clustering data points according to their proximity to others in the feature space, based on their closeness to the data points closest to the object [18]. The algorithm functions by identifying the K closest data points to the target instance and assigning a class or regression value based on the predominant label or average value among those neighbors.

The K-NN algorithm operates by measuring the distance between a test instance and each data point within the training dataset. The parameter K defines how many of the closest neighbours will be considered in making a prediction. Data points with the smallest distances to the test instance are identified as its nearest neighbors. Once the K value is set, the algorithm proceeds to compute the distance between the new (unlabeled) data and the existing training data. The K most similar instances are then selected. For classification tasks, the predicted class is determined by the majority vote among these neighbors. To assess the model's performance, evaluation is conducted using the test dataset, employing metrics such as accuracy, precision, recall, and the F1-score. If these evaluation results meet the expected thresholds, the model is considered reliable for making future predictions [19].

6. Evaluation Model

The evaluation model in this study utilizes a confusion matrix, a visual tool employed to evaluate the effectiveness of machine learning algorithms. Within a confusion matrix, the columns display the predicted classifications, whereas the rows indicate the actual classes. With this structure, all possible cases in a classification problem can be calculated and analyzed [20]. The confusion matrix includes several evaluation metrics. The evaluation metrics used include precision, recall, and the F1-score. Accuracy is determined by dividing the number of correct predictions by the total number of instances. Precision refers to the ratio of correctly identified positive instances to the total instances predicted as positive. Recall, on the other hand, represents the proportion of true positive predictions relative to all actual positive cases. The F1-score is the harmonic mean of precision and recall, balancing both metrics to provide a single performance measure that accounts for their contributions. [20]. Accuracy and metrics in the confusion matrix can be calculated using Equations (4), (5), (6), and (7).

$$Akurasi = \frac{True\ Positive + True\ Negatif}{True\ Positive + True\ Negative + True\ Positive + True\ Negative} \quad (4)$$

$$Presisi = \frac{True\ Positive}{True\ positive + False\ Positive} \quad (5)$$

$$Recall \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (6)$$

$$Score\ F1 \frac{2 \times Recall \times Presisi}{Recall + Presisi} \quad (7)$$

C. Result and Discussion

This research employs the NB and K-NN classification methods, with sentiment labelling performed using a lexicon-based approach. The data collection process involved examining content accessible through social media channels, particularly X (Twitter). At this stage, the researcher collected relevant data regarding the McDonald's boycott by crawling content from various user accounts. As a result, a dataset comprising 1,586 user-generated comments was compiled, as illustrated in Table 7. Notably, all data employed in this study were written in Indonesian. The dataset reflects a wide range of user opinions and sentiments regarding the boycott, from which 1,586 entries were selected for further analysis.

Table 7. Dataset

Created_at	Full_text
Mon Dec 25 09:46:48 +0000 2023	*Boikot McDonalds dan produk yang pro-sirewel #McDonalds #boikot #unyucomic*
Mon Dec 25 10:27:56 +0000 2023	*Boikot McDonald s Makin Gencar Fans K-Pop Minta Idol Tak Terlibat karena Diduga Dukung Israel*
Wed November 1 06:58:13 +0000 2023	*McDonald's Indonesia akhirnya buka suara terkait ramainya suara boikot McDonald's karena membantu tentara zionis Israel*
Thu Nov 09 00:46:07 +0000 2023	*Dapat Seruan Boikot McDonald's Indonesia Kini Salurkan Bantuan untuk Warga Gaza Palestina*

To ensure more accurate sentiment analysis outcomes, researchers typically implement several data processing stages. The first stage is pre-processing, which aims to clean and organize text data for easier analysis and interpretation. This process includes data cleaning, case folding, tokenization, filtering, stemming, and sentiment data labelling.

The next step is to apply term weighting through the TF-IDF approach to the cleaned dataset, thereby obtaining a numerical representation of the words in the documents. With TF-IDF, researchers can identify important words that differentiate one document from another, thus improving the accuracy of sentiment classification.

After the weighting process, feature extraction and classification were performed using the NB and K-NN methods. The performance of both models was evaluated using a confusion matrix. Based on the evaluation results, both K-NN and NB demonstrated their respective classification capabilities. This is evident from the confusion matrices shown in Figure 2 and the detailed performance metrics presented in Table 8 for both models.

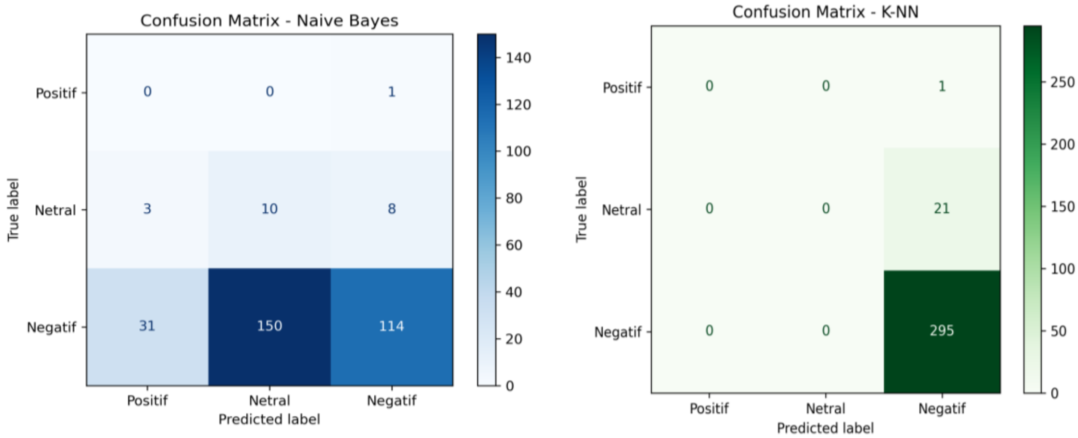

Figure 2. The Confusion Matrix

Table 8. Performance results of NB and K-NN

Metric	Naïve Bayes	K-NN
Accuracy	0.3912	0.9306
Precision	0.3298	0.3102
Recall	0.2875	0.3333
F1 - Score	0.2187	0.3214

The bar chart shown in Figure 3 illustrates the performance comparison between the NB and K-NN algorithms on sentiment classification related to the boycott of McDonald's. Evaluation of the model involves metrics such as Accuracy, Precision, Recall, and F1-Score.

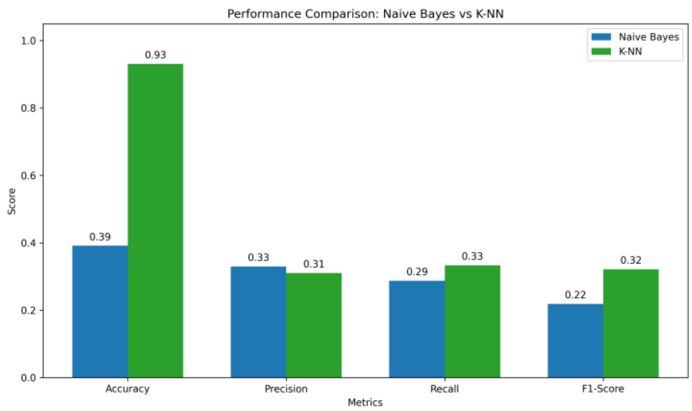

Figure 3. Performance Comparison NB vs K-NN

The K-NN algorithm demonstrates superior performance across all metrics. It achieves the highest accuracy at 93%, compared to NB, which scores only 39%. Although both models have relatively low Precision, Recall, and F1-scores due to class imbalance (with the Negative class dominating the dataset), K-NN still outperforms NB, achieving a Precision of 0.31, a Recall of 0.33, and an F1-score of 0.32. In contrast, NB only reaches 0.33, 0.29, and 0.22, respectively.

The findings indicate that the K-NN algorithm demonstrates superior performance in sentiment classification, especially in identifying the most prevalent class (Negative sentiment), as also reflected in the confusion matrix. The imbalance in sentiment distribution is evident, with the values Negative, Neutral, and Positive being 295, 21, and 1, respectively. It significantly influences the model's performance, making class-sensitive performance indicators, such as Recall and the F1-Score, essential for evaluation beyond accuracy.

D. Conclusion

This research employed Twitter as the primary source for conducting sentiment analysis data related to the boycott of McDonald's, resulting in a dataset comprising 1,586 comments. The analytical process followed a systematic pipeline including pre-processing (cleaning, tokenizing, filtering, and stemming), automatic sentiment labelling using a lexicon-based approach, TF-IDF vectorization, and sentiment classification performed through the use of NB and K-NN machine learning techniques.

The dataset was split into training and Testing sets using an 80:20 per cent ratio. The K-NN algorithm demonstrated superior performance compared to NB, K-NN demonstrated 93% accuracy, along with 31% precision, 33% recall, and a 32% F1-score. On the other hand, Naïve Bayes yielded 39% accuracy, 33% precision, 29% recall, and an F1-score of 22%.

The dataset exhibited the following sentiment distribution dominated by negative sentiment, comprising 93.5% of the total data. The neutral sentiment accounted for 6.1%, while positive sentiment was extremely rare, with only 0.3%. This imbalance significantly influenced the model's predictive capability, especially for underrepresented classes.

The findings suggest that the K-NN algorithm is more reliable in handling sentiment classification on imbalanced datasets, particularly in identifying negative sentiments, which are the majority. These insights can be utilized to understand public reactions to controversial corporate actions and to enhance decision-making strategies in brand management and public relations.

E. Acknowledgment

The author would like to express his deepest gratitude to the Informatics Department at Universitas Dr. Soetomo for the valuable support and guidance provided throughout the preparation and implementation of this research. He also sincerely thanks his supervisors for their invaluable advice and feedback, which greatly contributed to the refinement of this final project. In addition, he extends heartfelt appreciation to his family and friends for their constant encouragement, motivation, and prayers, which played a vital role in the successful completion of this study.

F. References

[1] N. A. R. Putri, "Analisis Jaringan pada Media Sosial X dengan# Boikot Menggunakan Social Network Analysis," *IJITECH: Indonesian Journal of Information …*, vol. 01, pp. 1–5, 2024, [Online].

[2] N. Fadhilah Az-Haari, D. Juardi, and A. Jamaludin, "Analisis Sentimen Terhadap Boikot Brand Pro-Israel Pada Twitter Menggunakan Metode Naïve Bayes (Studi Kasus: Starbucks)," 2024.

[3] I. Y. P. Salfa Aqila1, "Dampak Konflik Israel-Palestina Terhadap Harga Saham Perusahaan Makanan Dan Minuman Di Bursa Efek Indonesia," *Journal Of Social Science Research*, vol. Volume 4, pp. Page14246-14256, 2024.

[4] M. Yasir, Marissa Grace Haque, Robertus Suraji, and Istianingsih, "Analisis Sentimen Terhadap Kontroversi Fatwa MUI Nomor 83 Tahun 2023 Tentang Pemboikotan Produk yang Terafiliasi Israel," *Jurnal Ekonomi Manajemen Sistem Informasi*, vol. 5, no. 4, pp. 409–422, 2024, doi: 10.31933/jemsi.v5i4.1845.

[5] M. Buheji, D. Ahmed, and A. Ahmed, "Keeping the Boycott Momentum-from 'WAR on GAZA' Till 'Free-Palestine," *International Journal of Management (IJM)*, vol. 14, no. 7, pp. 205–229, doi: 10.17605/OSF.IO/NC75D.

[6] Z. A. Sriyanti, D. S. Y. Kartika, and A. R. E. Najaf, "Implementasi Model Bert Pada Analisis Sentimen Pengguna Twitter Terhadap Aksi Boikot Produk Israel," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, pp. 2335–2342, 2024, doi: 10.23960/jitet.v12i3.4743.

[7] R. Mujahidah and M. Sufyan Abdurrahman, "Pengaruh Terpaan Twit Boikot Mcd Terkait Konflik Israel-Palestina Terhadap Reputasi Mc Donald's Indonesia," 2025.

[8] A. M. Aditya *et al.*, "Analisis Sentimen Pengguna Twitter Terhadap Boikot Produk Israel Pada Merek Dagang Unilever Indonesia Menggunakan Algoritma Long Short Term Memory," *JATI (jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 5, pp. 9883–9890, 2024.

[9] J. Khatib Sulaiman Dalam No, A. Ibrahim Hamid, and A. Mohsin Abdulazeez, "Sentiment Analysis Based on Machine Learning Techniques: A Comprehensive Review," *Indonesian Journal of Computer Science*.

[10] Syahril Dwi Prasetyo, Shofa Shofiah Hilabi, and Fitri Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *Jurnal KomtekInfo*, vol. 10, pp. 1–7, 2023, doi: 10.35134/komtekinfo.v10i1.330.

[11] A. Darmawan and M. Ficky Aulia, "Desain Dan Komunikasi Merek McDonald's: Pengaruh Dukungan Terhadap Israel Dan Reaksi Konsumen."

[12] S. Adi and A. Wintarti, "Komparasi Metode Support Vector Machine (Svm), K-Nearest Neighbors (Knn), Dan Random Forest (Rf) Untuk Prediksi Penyakit Gagal Jantung," *MATHunesa: Jurnal Ilmiah Matematika*, vol. 10, no. 2, pp. 258–268, 2022, doi: 10.26740/mathunesa.v10n2.p258-268.

[13] A. P. Wibawa, M. Guntur, A. Purnama, M. Fathony Akbar, and F. A. Dwiyanto, "Metode-metode Klasifikasi," *Prosiding Seminar Ilmu Komputer dan Teknologi Informasi*, vol. 3, no. 1, pp. 134–138, 2018.

[14] F. Amaliah and I. K. Dwi Nuryana, "Perbandingan Akurasi Metode Lexicon Based Dan Naive Bayes Classifier Pada Analisis Sentimen Pendapat

Masyarakat Terhadap Aplikasi Investasi Pada Media Twitter," *Journal of Informatics and Computer Science (JINACS)*, vol. 3, no. 03, pp. 384–393, 2022, doi: 10.26740/jinacs.v3n03.p384-393.
[15] N. Muchammad Shiddieqy Hadna, P. Insap Santosa, and W. Wahyu Winarno, "Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen Di Twitter," 2016.
[16] U. Rofiqoh, R. S. Perdana, and M. A. Fauzi, "Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexion Based Feature," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, vol. 1, no. 12, pp. 1725–1732, 2017, [Online].
[17] R. I. Putra Selian, A. V. Vitianingsih, S. Kacung, A. Lidya Maukar, and J. Febrian Rusdi, "Sentiment Analysis of Public Responses on Social Media to Satire Joke Using Naïve Bayes and KNN," *Sinkron*, vol. 8, no. 3, pp. 1443–1451, 2024, doi: 10.33395/sinkron.v8i3.13721.
[18] N. K. Putri, A. V. Vitianingsih, S. Kacung, A. L. Maukar, and V. Yasin, "Sentiment Analysis of Brand Ambassador Influence on Product Buyer Interest Using KNN and SVM," *Indonesian Journal of Artificial Intelligence and Data Mining*, vol. 7, no. 2, p. 327, 2024, doi: 10.24014/ijaidm.v7i2.29469.
[19] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN Model-Based Approach in Classification," 2003.
[20] A. Ridhovan and A. Suharso, "Penerapan Metode Residual Network (Resnet) Dalam Klasifikasi Penyakit Pada Daun Gandum," *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 7, no. 1, pp. 58–65, 2022, doi: 10.29100/jipi.v7i1.2410.