Implementation of Deep Learning Models in Conducting Aspect-Based Sentiment Analysis

Chintika Yusni Adhelina[1], Muhammad Halmi Dar[2*], Mila Nirmala Sari Hasibuan[3]

[1,2,3] Faculty of Science and Technology, Universitas Labuhanbatu, Sumatera Utara Indonesia
*Corresponding Author:
Email: mhd.halmidar@gmail.com

Abstract.

The increasing volume of consumer reviews on e-commerce platforms has highlighted the need for sentiment analysis methods capable of capturing user opinions more specifically concerning particular aspects of products or services. Aspect-Based Sentiment Analysis (ABSA) addresses this need by identifying the aspects discussed in a review and determining the polarity of sentiment expressed toward each aspect. This study aims to explore and compare the effectiveness of two deep learning models, namely Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) in performing ABSA on Indonesian-language e-commerce user reviews. The research methodology comprises several stages: data exploration and cleaning, text preprocessing, aspect and sentiment annotation, training of CNN and LSTM models, and performance evaluation using metrics such as accuracy, precision, recall, and F1-score. The dataset is divided into training, validation, and testing subsets. The analyzed aspects include delivery, product, price, application, and service. Results show that the LSTM model outperforms CNN across all evaluation metrics. LSTM achieved an accuracy of 86.10%, precision of 85.70%, recall of 85.90%, and an F1-score of 85.80%, while CNN reported slightly lower values. Based on these findings, LSTM proves to be more effective in understanding the contextual and linguistic structure of the Indonesian language in ABSA tasks. This study provides a valuable contribution to the development of automatic sentiment analysis systems in the e-commerce sector. Future research can expand this approach by incorporating transformer-based models such as IndoBERT or integrating attention mechanisms to further improve predictive accuracy. These findings offer practical insights for industry stakeholders seeking to enhance customer experience through a deeper understanding of user sentiment.

Keywords: *Aspect-Based Sentiment Analysis (ABSA); CNN; Customer Reviews and Deep Learning, LSTM.*

1. INTRODUCTION

The development of information and communication technology has revolutionized the way consumers shop, making e-commerce an integral part of everyday life. Today, almost any product is easily accessible, both in terms of quantity and category variety, and even with unlimited availability. In Indonesia, e-commerce not only increases public access to goods and services but also reaches consumers in remote areas.

Most e-commerce platforms offer customer reviews that reflect users' experiences and satisfaction with products and services. These reviews cover a wide range of dimensions, including price, product quality, customer service, and availability [1]. For potential buyers, these reviews are a valuable source of information in the decision-making process [2], [3], [4], [5]. For manufacturers and service providers, reviews can be used to evaluate and improve product quality and customer experience [6]. However, as the number of e-commerce users increases, so too does the volume of unstructured review data. This large volume of data presents a challenge, as potential consumers must manually review numerous reviews before making a purchasing decision [7], [8]. In the era of Big Data and Artificial Intelligence (AI), the ability to understand and extract relevant information from unstructured data has become crucial [9]. On the other hand, the human ability to summarize and extract insights from massive user reviews is challenging, making it difficult to identify relevant and useful information [10]. In this context, sentiment analysis comes as a solution to transform unstructured data into organized and meaningful information [11], [12], [13], [14].

The main goal of sentiment analysis is to identify whether customers are satisfied or dissatisfied based on the opinions they express in reviews [15]. However, most sentiment analysis approaches are still conducted at the document or sentence level, which tends to provide a general overview without considering the differences in sentiment towards specific aspects within a single text [16], [17], [18], [19]. This approach becomes inadequate when a single review covers opinions on several different aspects, for example, a customer likes the product quality but is dissatisfied with the delivery service. Therefore, a more in-depth approach is needed, namely Aspect-Based Sentiment Analysis (ABSA). ABSA focuses on identifying opinions towards a specific entity and its aspects, such as performance, price, or service. This approach allows the system to extract and classify sentiment towards each aspect separately, thus providing richer and more specific insights [20], [21].

In this regard, various methods have been applied to sentiment analysis tasks, ranging from traditional machine learning algorithms such as K-Nearest Neighbors, Logistic Regression, Naïve Bayes, Random Forest, and Support Vector Machine [22], [23]. However, these traditional methods have limitations because they rely on manual feature engineering processes, produce sparse and high-dimensional feature vectors, and are less able to capture the syntactic and semantic structure of text [24], [25]. These challenges can be overcome by implementing deep learning models that are able to automatically extract features, understand the context of words in a sequence of text, and capture complex patterns in data [26].

Several previous studies have explored the application of deep learning in sentiment analysis. Research by [27] demonstrated the success of a CNN-LSTM model combination in identifying emotions and attitudes in reviews. Meanwhile, [28] compared the performance of various supervised machine learning algorithms on an Amazon product review dataset and found the superiority of deep learning models in sentiment classification. Other studies, such as by [29], developed an LSTM-based Enhanced Golden Jackal Optimizer algorithm, and [30], [31] showed that LSTM is effective in capturing context and long sequences in text. However, the results of a systematic literature review by [22] covering 20 international scientific publications from 2018–2024 indicate that research contributions from Indonesia in this field are still very limited. Only two publications originated from Indonesia, namely by [32] which compared Naïve Bayes and SVM, and by [33] which developed a Naïve Bayes model for predicting customer satisfaction scores.

Based on these facts, it can be concluded that research utilizing deep learning models for Aspect-Based Sentiment Analysis (ABSA) in the context of e-commerce reviews in Indonesia is still very limited. Therefore, this study aims to explore and compare the effectiveness of deep learning models, namely Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM), in conducting ABSA on e-commerce user reviews in Indonesia. This study will also measure model performance using evaluation metrics such as accuracy, precision, recall, and F1-score, to determine which model is most appropriate in expressing customer opinions based on specific aspects of the reviewed product or service.

II. METHODS

This study uses an experimental quantitative approach to evaluate and compare the performance of several deep learning models in conducting aspect-based sentiment analysis (ABSA) on customer review data on e-commerce platforms in Indonesia. The main objective of this study is to determine the effectiveness of deep learning models—specifically Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) models—in identifying specific aspects of reviews and determining the sentiment polarity associated with each aspect.

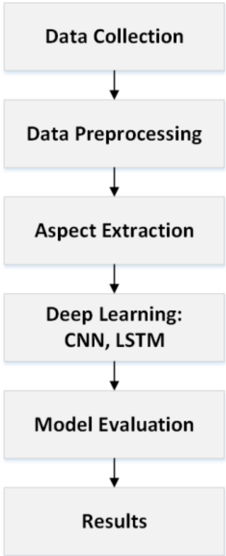

Fig. 1. Research Stage

The research process was conducted in a structured manner and consisted of several main stages as shown in Figure 1, starting with review data collection, followed by data pre-processing, aspect extraction, deep learning model implementation, and evaluation and analysis of model results. Each stage was systematically designed to ensure that the data used had undergone appropriate cleaning, normalization, and transformation processes so that it could be optimally used in model training. To support the generalizability and reliability of the research results, the dataset used was divided into training data and test data.

The dataset collected is the E-commerce Classification Review data (Shopee & Tokopedia) from Hugging Face [34]. This dataset contains review texts in Indonesian along with sentiment polarity labels (0: Negative, 1: Positive). The data consists of three main parts: train.csv, validation.csv, and test.csv. Data preprocessing includes: converting all text to lowercase, removing special characters, emojis, and irrelevant punctuation, removing Indonesian stopwords, tokenizing sentences using an Indonesian tokenizer, and, if needed for ABSA, aspect extraction or annotation (e.g., price, delivery, product quality). The Deep Learning models used are: CNN and LSTM. Inputs are review tokens and labels (0/1). The models are trained on the training data, validated on the validation set, and tested on the test set. The models are evaluated using several performance metrics such as accuracy, precision, recall, and F1-score. In addition, comparisons between models will also be conducted to determine which approach is most optimal in the context of aspect-based sentiment analysis for Indonesian in the e-commerce domain.

III. RESULT AND DISCUSSION

This section presents the results of the implementation of the research procedure in applying a deep learning model for Aspect-Based Sentiment Analysis (ABSA) to Indonesian-language e-commerce user reviews, specifically from the Shopee and Tokopedia platforms. The research was conducted through the stages of data exploration, preprocessing, aspect annotation, CNN and LSTM model training, and model performance evaluation using accuracy, precision, recall, and F1-score metrics.

The data used in this study consists of three parts: training data, validation data, and test data, all in CSV format and containing e-commerce user reviews. Each data set contains a review column (text) and a sentiment label (positive, neutral, negative). The number of data sets is shown in Table 1. The distribution of sentiment labels in the data is quite balanced, ensuring that the model training results are not biased toward any particular class.

Table 1. Dataset

Dataset	Amount of Data
Train	12,000
Validation	3,000
Test	3,000

Preprocessing steps are carried out to ensure the data is clean and ready for analysis. Some of the main processes include: Tokenization, which is useful for breaking review sentences into tokens or words; Case folding, which converts all text to lowercase; Text cleaning, which removes punctuation, numbers, URLs, emoticons, and special characters; Stopword removal, which removes common words that do not carry important meaning (such as yang, dan, di); Stemming, which converts words to their base form. The preprocessing results are saved in a .csv file and used in the training stage.

Table 2. Data Preprocessing

Data Mentah	Setelah Preprocessing
Pengiriman barangnya cepat banget, tapi kualitas produknya jelek dan tidak sesuai gambar :(	pengiriman barang cepat tapi kualitas produk jelek tidak sesuai gambar

Table 2 illustrates the process of transforming text data from its original (raw) form to a preprocessed form before being used in the model training phase. The "Raw Data" column contains the original review text from e-commerce users. This text reflects complex opinions, as it encompasses two important aspects: delivery (positive) and product quality (negative). The raw text also contains informal elements such as the word “banget”, the emoticon ":(", and suffixes such as “-nya”. The “After Preprocessing” column shows the results after the review has gone through several stages of text cleaning and standardization, namely: The steps taken in preprocessing include: Case folding: Changing all letters to lowercase, Normalization: Removing non-standard words such as “banget” which is replaced with “cepat”, Stopword removal: Removing meaningless words such as “nya”, Symbol/emoticon cleaning: The sad emoticon “:(” is removed because it is not informative in text form, Tokenization and stemming: Although not displayed explicitly, word structures such as “produknya” become “produk”. The purpose of this preprocessing is to simplify data representation and improve the performance of machine learning or deep learning models by presenting data in a clean and structured format. With preprocessed data such as in this example, the ABSA (Aspect-Based Sentiment Analysis) model can identify and learn the relationship between aspect keywords (“pengiriman”, “kualitas produk”) and the sentiments that accompany them (“cepat” → positif, “jelek” dan “tidak sesuai” → negatif).

Next, aspect annotation was performed on each review. The annotated aspects covered several dimensions of e-commerce services, such as: Product, Shipping, Customer Service, Price, and Application/Platform. Annotation was performed both manually and semi-automatically to generate final labels for model training. The aspect annotation results are shown in Table 3.

Table 3. Aspect Annotation

Ulasan (Setelah Preprocessing)	Aspect	Sentiment
pengiriman cepat tapi produk jelek tidak sesuai gambar	Shipping	Positif
	Products	Negatif
harga murah kualitas oke sesuai harapan	Prices	Positif
	Products	Positif
aplikasi error loading lama bikin kesel	Applications	Negatif
respon penjual cepat ramah dan membantu	Services	Positif

Table 3 shows how a single customer review can contain more than one aspect, and how sentiment for each aspect is analyzed separately. The most frequently occurring aspects in the data are: Product, Delivery, Price, App, and Service. The ABSA approach is particularly useful because it can distinguish between good and bad experiences within a single review. Sentiment labels are adjusted based on the context of the sentence and keywords, not just the overall tone of the review. By performing this type of annotation, we can more precisely understand which parts of the customer experience are rated as good or bad. Implementing aspect-based annotations, as shown in Table 3, is crucial for developing more precise and informative machine learning models (such as CNNs and LSTMs). This allows the system to provide not only general conclusions about sentiment, but also specific insights into the strengths and weaknesses of e-commerce services based on key aspects.

The models used in this study are Convolutional Neural Network (CNN), which relies on convolution of word features to understand patterns in reviews, and Long Short-Term Memory (LSTM), which is better at capturing sequential context in sentences. The main parameters used for both models are: Epoch: 10, Batch size: 32, Optimizer: Adam, Loss Function: Categorical Crossentropy, and Pre-trained embedding: FastText Bahasa Indonesia. The CNN and LSTM models were trained using training data, validated with validation data, and tested on test data.

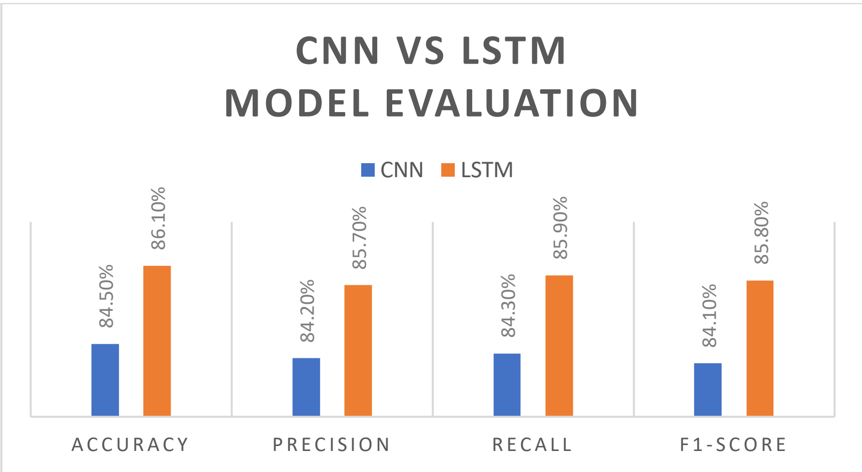

Fig. 2. Model Evaluation

Figure 2 compares the performance of two deep learning models, namely Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM), in performing aspect-based sentiment analysis (ABSA) on Indonesian-language e-commerce user reviews. This graph consists of four main evaluation metrics: Accuracy, Precision, Recall, and F1-Score.

Accuracy measures the proportion of total correct predictions across all tested data. The graph shows that LSTM has a higher accuracy than CNN, at 86.10% compared to 84.50%. This indicates that overall, LSTM is better able to predict sentiment labels that match the actual data, for both positive and negative classes. Precision measures the extent to which the model produces relevant predictions, or in the context of sentiment, how many of the positive predictions are actually positive. LSTM again excels with a precision of 85.70%, indicating that this model has a better ability to avoid false positives than CNN. This means that LSTM's positive sentiment predictions are more accurate than CNN's. Recall measures the model's ability to capture all relevant instances, namely how many data with a true positive label are successfully recognized as positive. LSTM's performance in this metric is also higher (85.90%) than CNN's (84.30%), indicating that LSTM is more reliable in capturing reviews that do contain a certain sentiment, without missing many. The F1-Score is the harmonic mean of precision and recall, providing a comprehensive overview of the model's balance in terms of prediction accuracy and completeness. The higher F1-Score for LSTM (85.80%) compared to CNN (84.10%) indicates that LSTM provides more stable performance and a balance between avoiding false positives and missing important data (false negatives).

Based on these four metrics, LSTM consistently outperforms CNN in all evaluation aspects. This can be explained technically: LSTM is designed to capture long-term dependencies in sequential data such as review text. With its ability to retain context from previous words through long-term memory, LSTM is more effective at understanding complex sentence structures and interpreting sentence context in customer reviews. In contrast, CNN, while known to be effective at capturing local patterns and spatial features, tends to be less effective at retaining long-term context, which is crucial in aspect-based sentiment analysis. Therefore, LSTM models excel in this task because customer reviews often have a narrative structure containing context between phrases or paragraphs.

The evaluation results presented in Figure 2 show that the LSTM model consistently outperformed the CNN model across all four evaluation metrics. In the accuracy metric, the LSTM achieved 86.10%, while the CNN achieved only 84.50%. Furthermore, the precision value for the LSTM was 85.70%, higher than the CNN's 84.20%. The LSTM's recall metric was also superior, with a value of 85.90% compared to the CNN's 84.30%. Finally, the LSTM's F1-score was 85.80%, while the CNN achieved 84.10%. This performance difference indicates that the LSTM has an advantage in understanding the context of word order within a sentence, which is crucial for handling text-based data such as customer reviews.

The superiority of LSTM in ABSA can be explained by its internal architecture, which is specifically designed for processing sequential data. LSTMs have long-term memory capabilities that enable them to understand the relationships between distant words in a sentence. This is very useful in detecting specific aspects within a review sentence and associating these aspects with relevant sentiment expressions. For example, in a review like "fast delivery but bad product, not as pictured," LSTMs are able to distinguish that positive sentiment is directed towards the "delivery" aspect, while negative sentiment is directed towards the "product" aspect. In contrast, CNNs tend to focus on local patterns through the use of convolutional filters, which, while effective in some text classification tasks, are less optimal at understanding long-term dependencies within sentences. CNN models are also more sensitive to local context, so they may have difficulty capturing the sentiment hidden in longer or more complex sentences.

Furthermore, the aspect annotation results shown in Table 3 demonstrate the diversity of aspects contained in e-commerce user reviews, such as "Shipping," "Product," "Price," "Application," and "Service." This diversity requires the model to be able to distinguish opinions on each aspect independently, even within a single sentence. LSTM has been shown to handle this complexity more accurately than CNN. This finding aligns with various previous studies showing that LSTM excels in tasks involving semantic and syntactic analysis of long or complex texts. Therefore, based on the experimental results in this study, it can be concluded that the LSTM model is more feasible and effective for implementation in an aspect-based sentiment analysis system, especially for Indonesian text data with diverse sentence structures and rich meaning.

IV. CONCLUSION

This study was conducted to explore and compare the effectiveness of two deep learning model architectures, namely Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM), in performing Aspect-Based Sentiment Analysis (ABSA) on Indonesian-language e-commerce user reviews. The research process includes a series of systematic stages, starting from data exploration, text preprocessing, aspect and sentiment annotation, model training, to performance evaluation using standard evaluation metrics: accuracy, precision, recall, and F1-score.

Based on the experimental results, it was found that the LSTM model consistently demonstrated superior performance compared to the CNN model in all evaluation metrics used.

LSTM recorded an accuracy of 86.10%, a precision of 85.70%, a recall of 85.90%, and an F1-score of 85.80%. In contrast, the CNN model recorded an accuracy of 84.50%, a precision of 84.20%, a recall of 84.30%, and an F1-score of 84.10%. The superiority of LSTM is mainly due to its ability to understand the context and dependencies of words in longer or more complex sentences, which is very relevant in detecting sentiment towards certain aspects of a review.

Thus, the main objectives of this study have been achieved. First, the study successfully explored and compared two deep learning models in the ABSA context, and confirmed that LSTM is more effective in revealing customer opinions on specific aspects of products and services. Second, through performance measurements using accuracy, precision, recall, and F1-score metrics, this study demonstrated that LSTM is a more appropriate model to be implemented in an aspect-based sentiment analysis system for Indonesian-language e-commerce reviews.

The results of this study can form the basis for the development of a smarter and more targeted customer review analytics system, which can ultimately be used by businesses to understand and improve the quality of products and services based on the aspects that customers care most about. In the future, this research can be expanded by involving transformer-based models such as IndoBERT, or by integrating an attention mechanism into the LSTM architecture to further improve the model's accuracy and efficiency.

REFERENCES

[1] Y. Liu, J. Lu, J. Yang, and F. Mao, "Sentiment analysis for e-commerce product reviews by deep learning model of Bert-BiGRU-Softmax," *Mathematical Biosciences and Engineering*, vol. 17, no. 6, pp. 7819–7837, 2020, doi: 10.3934/mbe.2020398.

[2] M. A. Kassem, A. A. Abohany, A. A. A. El-Mageed, and K. M. Hosny, "A novel deep learning model for detection of inconsistency in e-commerce websites," *Neural Comput Appl*, vol. 36, no. 17, pp. 10339–10353, 2024, doi: 10.1007/s00521-024-09590-5.

[3] S. Safar, B. R. Jose, and T. Santhanakrishnan, "An Improved Recommendation System with Aspect-Based Sentiment Analysis," in *Responsible Data Science*, J. Mathew, G. Santhosh Kumar, D. P., and J. M. Jose, Eds., Singapore: Springer Nature Singapore, 2022, pp. 75–87.

[4] T. S. Rambe, M. N. S. Hasibuan, and M. H. Dar, "Sentiment Analysis of Beauty Product Applications using the Naïve Bayes Method," *SinkrOn*, vol. 8, no. 2, pp. 980–989, 2023, doi: 10.33395/sinkron.v8i2.12303.

[5] R. Thomas and J. R. Jeba, "A novel framework for an intelligent deep learning based product recommendation system using sentiment analysis (SA)," *Automatika*, vol. 65, no. 2, pp. 410–424, 2024, doi: 10.1080/00051144.2023.2295148.

[6] P. Sasikala and L. Mary Immaculate Sheela, "Sentiment analysis of online product reviews using DLMNN and future prediction of online product using IANFIS," *J Big Data*, vol. 7, no. 1, 2020, doi: 10.1186/s40537-020-00308-7.

[7] N. Chintalapudi, G. Battineni, M. Di Canio, G. G. Sagaro, and F. Amenta, "Text mining with sentiment analysis on seafarers' medical documents," *International Journal of Information Management Data Insights*, vol. 1, no. 1, p. 100005, 2021, doi: 10.1016/j.jjimei.2020.100005.

[8] A.-M. Iddrisu, S. Mensah, F. Boafo, G. R. Yeluripati, and P. Kudjo, "A sentiment analysis framework to classify instances of sarcastic sentiments within the aviation sector," *International Journal of Information Management Data Insights*, vol. 3, no. 2, p. 100180, 2023, doi: 10.1016/j.jjimei.2023.100180.

[9] L. Yang, Y. Li, J. Wang, and R. S. Sherratt, "Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning," *IEEE Access*, vol. 8, pp. 23522–23530, 2020, doi: 10.1109/access.2020.2969854.

[10] E. S. P. Krishna *et al.*, "Enhancing E-commerce recommendations with sentiment analysis using MLA-EDTCNet and collaborative filtering," *Sci Rep*, vol. 15, no. 1, p. 6739, Feb. 2025, doi: 10.1038/s41598-025-91275-7.

[11] P. Savci and B. Das, "Prediction of the customers' interests using sentiment analysis in e-commerce data for comparison of Arabic, English, and Turkish languages," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 3, pp. 227–237, 2023, doi: 10.1016/j.jksuci.2023.02.017.

[12] N. Pughazendi, P. V Rajaraman, and M. H. Mohammed, "Graph Sample and Aggregate Attention Network optimized with Barnacles Mating Algorithm based Sentiment Analysis for Online Product Recommendation," *Appl Soft Comput*, vol. 145, p. 110532, 2023, doi: 10.1016/j.asoc.2023.110532.

[13] N. R. Bhowmik, M. Arifuzzaman, M. R. H. Mondal, and M. S. Islam, "Bangla Text Sentiment Analysis Using Supervised Machine Learning with Extended Lexicon Dictionary," *Natural Language Processing Research*, vol. 1, no. 3–4, p. 34, 2021, doi: 10.2991/nlpr.d.210316.001.

[14] J.-J. Wu and S.-T. Chang, "Exploring customer sentiment regarding online retail services: A topic-based approach," *Journal of Retailing and Consumer Services*, vol. 55, p. 102145, 2020, doi: 10.1016/j.jretconser.2020.102145.

[15] M. P. Geetha and D. Karthika Renuka, "Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model," *International Journal of Intelligent Networks*, vol. 2, pp. 64–69, 2021, doi: 10.1016/j.ijin.2021.06.005.

[16] E. Sihombing, M. Halmi Dar, and F. Aini Nasution, "Comparison of Machine Learning Algorithms in Public Sentiment Analysis of TAPERA Policy," *International Journal of Science, Technology & Management*, vol. 5, no. 5, pp. 1089–1098, 2024, doi: 10.46729/ijstm.v5i5.1164.

[17] L. Zhu, M. Xu, Y. Bao, Y. Xu, and X. Kong, "Deep learning for aspect-based sentiment analysis: a review," *PeerJ Comput Sci*, vol. 8, pp. e1044–e1044, Jul. 2022, doi: 10.7717/peerj-cs.1044.

[18] H. H. Do, P. W. C. Prasad, A. Maag, and A. Alsadoon, "Deep Learning for Aspect-Based Sentiment Analysis: A Comparative Review," *Expert Syst Appl*, vol. 118, pp. 272–299, 2019, doi: 10.1016/j.eswa.2018.10.003.

[19] M. H. D. T. Ari Bangsa, S. Priyanta, and Y. Suyanto, "Aspect-Based Sentiment Analysis of Online Marketplace Reviews Using Convolutional Neural Network," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 14, no. 2, p. 123, 2020, doi: 10.22146/ijccs.51646.

[20] A. Karimi, L. Rossi, and A. Prati, "Adversarial Training for Aspect-Based Sentiment Analysis with BERT," in *2020 25th International Conference on Pattern Recognition (ICPR)*, 2021, pp. 8797–8803. doi: 10.1109/ICPR48806.2021.9412167.

[21] J. K. Prathi, P. K. Raparthi, and M. V. Gopalachari, "Real-Time Aspect-Based Sentiment Analysis on Consumer Reviews BT - Data Engineering and Communication Technology," K. S. Raju, R. Senkerik, S. P. Lanka, and V. Rajagopal, Eds., Singapore: Springer Nature Singapore, 2020, pp. 801–810.

[22] A. Daza, N. D. González Rueda, M. S. Aguilar Sánchez, W. F. Robles Espíritu, and M. E. Chauca Quiñones, "Sentiment Analysis on E-Commerce Product Reviews Using Machine Learning and Deep Learning Algorithms: A Bibliometric Analysis, Systematic Literature Review, Challenges and Future Works," *International Journal of Information Management Data Insights*, vol. 4, no. 2, p. 100267, 2024, doi: 10.1016/j.jjimei.2024.100267.

[23] L. Davoodi, J. Mezei, and M. Heikkilä, "Aspect-based sentiment classification of user reviews to understand customer satisfaction of e-commerce platforms," *Electronic Commerce Research*, 2025, doi: 10.1007/s10660-025-09948-4.

[24] R. Kusumaningrum, I. Z. Nisa, R. Jayanto, R. P. Nawangsari, and A. Wibowo, "Deep learning-based application for multilevel sentiment analysis of Indonesian hotel reviews," *Heliyon*, vol. 9, no. 6, pp. e17147–e17147, Jun. 2023, doi: 10.1016/j.heliyon.2023.e17147.

[25] S. Tabinda Kokab, S. Asghar, and S. Naz, "Transformer-based deep learning models for the sentiment analysis of social media data," *Array*, vol. 14, p. 100157, 2022, doi: 10.1016/j.array.2022.100157.

[26] O. Bellar, A. Baina, and M. Ballafkih, "Sentiment Analysis: Predicting Product Reviews for E-Commerce Recommendations Using Deep Learning and Transformers," *Mathematics*, vol. 12, no. 15, p. 2403, 2024, doi: 10.3390/math12152403.

[27] H. J. Alantari, I. S. Currim, Y. Deng, and S. Singh, "An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews," *International Journal of Research in Marketing*, vol. 39, no. 1, pp. 1–19, 2022, doi: 10.1016/j.ijresmar.2021.10.011.

[28] M. F. bin Harunasir, N. Palanichamy, S.-C. Haw, and K.-W. Ng, "Sentiment Analysis of Amazon Product Reviews by Supervised Machine Learning Models," *Journal of Advances in Information Technology*, vol. 14, no. 4, pp. 857–862, 2023, doi: 10.12720/jait.14.4.857-862.

[29] P. Rasappan, M. Premkumar, G. Sinha, and K. Chandrasekaran, "Transforming sentiment analysis for e-commerce product reviews: Hybrid deep learning model with an innovative term weighting and feature selection," *Information Processing & Management*, vol. 61, no. 3, p. 103654, 2024, doi: 10.1016/j.ipm.2024.103654.

[30] H. Fang, G. Jiang, and D. Li, "Sentiment analysis based on Chinese BERT and fused deep neural networks for sentence-level Chinese e-commerce product reviews," *Systems Science & Control Engineering*, vol. 10, no. 1, pp. 802–810, 2022, doi: 10.1080/21642583.2022.2123060.

[31] R. Liu, H. Wang, and Y. Li, "AgriMFLN: Mixing Features LSTM Networks for Sentiment Analysis of Agricultural Product Reviews," *Applied Sciences*, vol. 13, no. 10, p. 6262, 2023, doi: 10.3390/app13106262.

[32] Y. Basani, H. V Sibuea, S. Ida, P. Sianipar, and J. Presly, "Application of Sentiment Analysis on Product Review E-Commerce Application of Sentiment Analysis on Product Review E- Commerce," 2019, doi: 10.1088/1742-6596/1175/1/012103.

[33] A. A. Ilham, A. Bustamin, and E. Wahyudiarto, "Customer Satisfaction Assessment System on Transactions E-commerce Product Purchases Using Sentiment Analysis," *Int J Adv Sci Eng Inf Technol*, vol. 13, no. 3, pp. 1041–1051, 2023, doi: 10.18517/ijaseit.13.3.18273.

[34] S. Mulyani, " ulasan-ecommerce-classification (Revision 7f44449) ," 2024, *Hugging Face* . doi: 10.57967/hf/2418.