Institute of Research and Publication Indonesia (IRPI)
Public Research Journal of Engineering, Data Technology and Computer Science
Journal Homepage: https://journal.irpi.or.id/index.php/predatecs
Vol. 2 Iss. 1 July 2024, pp: 52-59
ISSN(P): 3024-921X | ISSN(E): 3024-8043

Evaluation of the Effectiveness of Neural Network Models for Analyzing Customer Review Sentiments on Marketplace

Kana Karunia[1], Aprilya Eka Putri[2*], May Dila Fachriani[3], Muhammad Hilman Rois[4]

[1,2,3]Department of information System, Faculty of Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
[4]Al-Qasimia University, United Arab Emirates

E-Mail: [1]12050317289@students.uin-suska.ac.id, [2]12050323097@students.uin-suska.ac.id, [3]12050327198@students.uin-suska.ac.id [4]q20216124@alqasimia.ac.ae

Received Dec 25th 2023; Revised Feb 3rd 2024; Accepted Apr 25th 2024
Corresponding Author: Aprilya Eka Putri

Abstract

According to the 2019 report, Tokopedia is the most visited marketplace with 140,000,000 visitors per month, making it one of the most popular marketplaces in Indonesia. Customers have the opportunity to write reviews about the products they purchase at the end of the transaction process on Tokopedia. The aim of this research is to conduct sentiment analysis on product reviews on Tokopedia. Three neural networks that will be used for text classification are Bi-GRU, GRU, and LSTM. The data processing technique is divided into training and testing samples, split into 80%:20% using the holdout technique. The BI-GRU algorithm has an accuracy of 0.93% and precision of 0.96, better than the other two methods LSTM and GRU, which each have an accuracy of 0.92 and recall of 0.91.

Keyword: Bidirectional Gated Recurrent Unit, Customer Reviews, Gated Recurrent Unit, Long-ShortTerm Memory.

1. INTRODUCTION

Tokopedia, one of the marketplaces founded by William Tanuwijaya and Leontinus Alpha Edison in 2009. In 2019 Tokopedia is a marketplace that has the most visitors with more than 140,000,000 visitors / month. Even tokopedia itself is one of the companies that holds the title of unicorn or a company that has a valuation above $1 billion [1]. The existence of features on Tokopedia with the use of Online Customer Rating and Review as tools on the Tokopedia display, is expected to increase buying interest by potential consumers [2]. Customer review or customer review is one of the important elements in online-to-offline (O2O) commerce [3]. Customer reviews have become a rich source of information for assessing customer satisfaction with goods or services [4]. Therefore, improving the understanding of customer reviews on Tokopedia is an urgent need. The problem arises when the ever-increasing volume of customer reviews makes it difficult for sellers or platform managers to efficiently analyze and respond to each review manually. Customer reviews are divided into three sentiments: positive, negative, and neutral. Companies are more likely to respond to reviews with positive and negative comments covering more topics than comments with fewer topics [5]. Sentiment Analysis is an effective method to extract information from big data. Such as Sentiment Analysis of product reviews Marketplace platform, namely Tokopedia [6] Performing Sentiment Analysis on customer reviews and feedback is one way to ensure that the customer experience will be good, from product ordering to product delivery [7]. Sentiment Analysis of Tokopedia product reviews can be done with Deep Learning approaches, such as the Gated Recurrent Unit (GRU) model, Bidirectional Gated Recurrent Unit (Bi-GRU) model, and Long Short Term Memory (LSTM) model [8].

Deep learning adapts a multilayer approach to neural network layers. The modeling is extracted automatically, thus achieving better accuracy and performance. Sentiment analysis using deep learning approaches is a promising research area [9]. Deep learning methods work by extracting data using a neural network. From here the model will then learn from errors. The neural network itself consists of a set of layers mapped by an activity function [10]. Deep learning is becoming increasingly popular, one of the reasons is because when using large and complex dataset, machine learning methods that were previously used have several weaknesses, including training processing times that tend to be long, and poor accuracy[11].

LSTM is a deformed Recurrent Neural Network (RNN) structure that handles memory information by adding memory cells to the Hidden Layer. Information is passed between cells in the Hidden Layer through a series of programmable gates (Input, Output, and Forget gates) [12]. GRU is a variant of LSTM that has the same performance as LSTM, but produces more results. GRU improves the configuration of LSTM units and conjugates three LSTM gating units into two LSTM gating units as Update Gate and Reset Gate [13]. Bi-GRU is a sequence processing paradigm consisting of two GRUs working together. One GRU provides feedback in the forward direction, and the other GRU provides feedback in the backward direction [14].

The research conducted in 2017 by Hario Laskito Ardi et al. found that Unigram had the highest accuracy value, with 80.87%, in sentiment analysis of product reviews in the market using the Support Vector Machine (SVM) algorithm. This study found three elements of sentiment analysis: products, services, and delivery [15]. In 2020, based on research conducted by Arif Nur Rohman et al. on sentiment analysis of product reviews in the marketplace using Natural Language Processing and a machine learning approach with Naive Bayes and K-NN algorithms. The testing scenario found an average Naive Bayes accuracy of 52.4% on the Unigram dataset and an average KNN accuracy of 79.4% on the Bigram dataset [16]. Previous research on sentiment analysis using LSTM by Nitish Ranjan Bhowmik et al. in 2022 showed that the LSTM model can predict the sentiment of social media users very accurately with an accuracy value of 74.16%. [17] According to research conducted by Mohammed M. Abdelgwad et al., the best model for predicting the sentiment of a sentence in 2022 found that Bi-GRU has an accuracy value of 83.98% compared to GRU and LSTM. [18] In 2022, Jyothis Joseph et al. compared CNN, RNN, LSTM, and GRU models to analyze people's opinions or feelings towards entities such as products, services, and individuals. It was found that LSTMs and GRUs are better than simple RNNs because they can capture long-term dependencies [19]. Research by Wajdi Aljedaani et al. in 2022 found that LSTM-GRU outperformed all models and previous research with the highest accuracy of 0.97 and F1 score of 0.96 in aviation industry Twitter user sentiment to find customer feedback [20].

Unlike previous studies, this research examines product reviews in the market to conduct sentiment analysis by comparing the best algorithms from previous research, namely LSTM, GRU, and Bi-GRU algorithms. This research aims to address the challenge of efficiently analyzing and responding to customer reviews on Tokopedia. By applying LSTM, GRU, and Bi-GRU models, it is expected that the developed model can accurately classify customer reviews, allowing sellers or platform managers to respond more effectively to customer feedback.

2. MATERIAL AND METHOD

The implementation of the suggested technique can be seen in Figure 1, and the achievement to be aimed for has 3 stages. The first stage is Pre-processing. This procedure is used to clean and remove complexity or unnecessary text in a review. The second stage is to perform sentiment classification using 3 Deep Learning Models, namely Bi-GRU, GRU, and LSTM. Reviews are categorized into negative and positive at this stage. Measuring the ability of the model to perform classification is the last stage. The ability of the model is measured in terms of F1-Score, Accuracy, Recall, F1-Precision, and Computational Time.

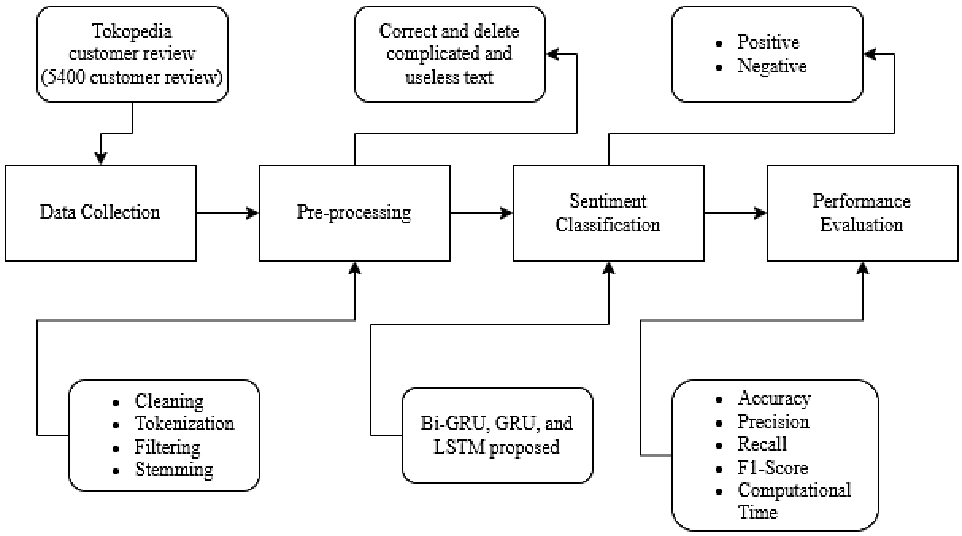

Figure 1. Research Methodology

2.1. Data Collection

The dataset used for this research is Customer Reviews from one of the giant e-commerce in Indonesia named Tokopedia, downloaded from Kaggle. This dataset contain 5400 Customer Reviews from 29 products on Tokopedia that use the Indonesian Language. Each Review has eleven attributes, namely Category, Product Name, Location, Price, Overall Rating, Number Sold, Total Review, Customer Rating, Customer Review, Sentiment, and Emotion. In this study only needs two attributes, namely Customer Review and Sentiment, which contains review and label from each product. The data collected is shown in Table 1.

Table 1. Raw Data

No	Review
1	Alhamdulillah berfungsi dengan baik. Packaging aman. Respon cepat dan ramah. Seller dan kurir amanah
41	Sangat kecewa. Baru 4 bulan scroll sudah rusak.
81	Sangat mengecewakan, barang yg dikirim tidak sesuai konfirmasi, sampai harus bolak-balik 2 kali, saya harus nanggung 6 X ongkos kirim....SANGAT MENGECEWAKAN
…	…
5400	Produk sesuai deskripsi, packing aman terlindung, pengiriman cepat, sangat direkomendasikan.

Each review is annotated with single emotion, i.e., love, happy, anger, fear, and sadness. The group of annotators does the annotation process to provide sentiment labels by following emotion annotation criteria created by expert in clinical psychology. Representation based on emotions is shown in Figure 2.

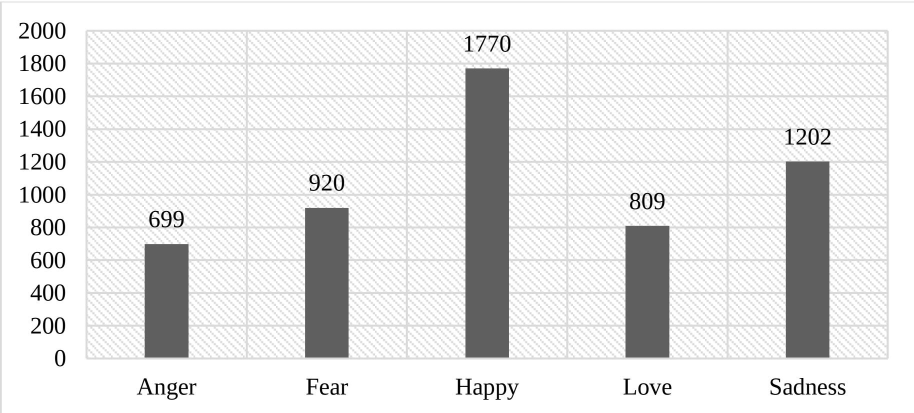

Figure 2. Review Division

Based on Figure 2, it can be categorized that reviews with happy and love emotions are positive sentiments, while reviews with anger, fear, and sadness emotions are negative sentiments. The percentage of each sentiment is shown in Figure 3.

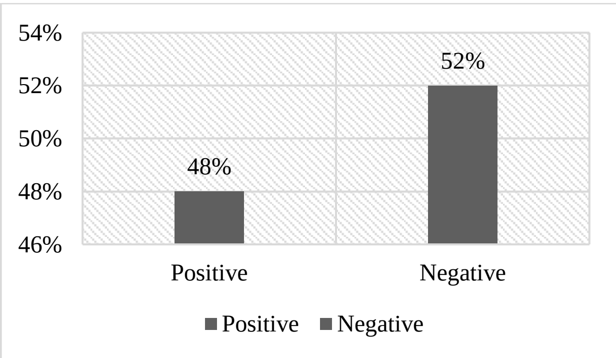

Figure 3. Sentiment Percentage

2.2. Pre-processing

For Machine Learning related research, once the data collection has been done, the next step is to perform pre-processing to create useful inputs for the proposed model. Preprocessing is an important task in the Machine Learning. Data Pre-processing cleans the data from useless information, which does not help in the model training processing. For the Marketplace Customer Reviews dataset, four data pre-processing steps were used.

Cleaning: To remove noise in the data, elimination of words that are proven not to provide information in a sentence such as Hyperlinks, Hashtags, Symbols, and punctuation, to ensure that words with banca marks or words without punctuation marks, are considered as the same word. Changing capital letters to lowercase letters is also done in this step (Case Folding).

Tokenization: As illustrated in figure 4, Tokenization is the process of converting a Text Segment into smaller units, referred to as Tokens. These tokens can consist of words, word fragments, or punctuation marks, thus enabling more effective NLP analysis and processing.

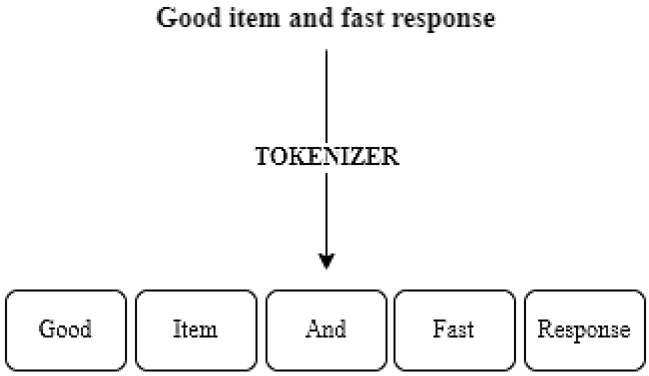

Figure 4. Tokenization Process, showing the transformation of raw text into tokens.

Filtering: Using NLTK a NLP system that removes or eliminates stop-words. Some of the many stop-words in Indonesian are "dari", "dan", "yang", "dengan", "untuk", "adalah", and some other words that have no meaning. These are words that often appear in a sentence, but do not provide information, and will be removed.

Stemming: Stemming means the removal of prefixes, suffixes, insertions, and appendages from a word to get its basic or root form. Stemmer Factory algorithm, a rule-based method of Library Sastrawi version 1.0.1 developed by Hanif Amal Robbani on January 18, 2016, is used to perform Stemming on Indonesian words or sentences.

2.3. Sentiment Analysis

Sentiment analysis is used to analyze people's opinions, sentiments, evaluations, attitudes, and emotions written by people about face-to-face activities. The focus of sentiment analysis is opinions that express negative or positive emotions or sentiments. Neutral expressions must also be considered [21]. Lexicon-based methods are used to determine sentiment by summing up the scores for each word that represents sentiment in a sentence. A lexicon is a collection of information. The first step is to determine the sentiment word scores, both positive and negative. Each word in the sentence has a certain score: words with positive sentiment are given a score of 1, words with negative sentiment are given a score of -1, and words with neutral sentiment are given a score of 0. The score is calculated by summing them all. A sentence is considered a positive sentiment if the total score is greater than 0, but if the sum is less than 0, then the sentence is a negative sentiment. Conversely, if the sum is greater than 0, then the sentence is a positive sentiment [22][23].

2.4. Long Short Term Memory (LSTM)

LSTM is a universal RNN, when the time span is much longer than the size known with important events, this algorithm is suitable for learning from experience to classify, process, and predict time series data [24]. LSTM consists of three gates: the input gate, the forget gate, and the output gate. Each gate performs a specific function in controlling the flow of information. The input gate decides how to update the internal state based on the current input and the previous internal state. The forget gate determines how much of the previous internal state should be forgotten. Lastly, the output gate regulates the influence of the internal state on the system. Here are the mathematical formulas used in LSTM:

1. Forgate Gate (Controlling how much the old value in the memory cell is remembered.)

$$f_t = \sigma(W_{fx} X_t + W_{ih} h_{t-1} + W_{fc} c_{t-1} + b_f) \quad (1)$$

This is done by multiplying the previous input $ht-1$ and the current input Xt with weights Wf and Uf, and then passing through a sigmoid activation function to generate a vector containing values between 0 and 1. This determines how important each element of the previous memory $ct-1$ is to be forgotten.

2. Input Gate (Controlling the extent to which the new input value affects the value in the memory cell)

$$i_t = \sigma(W_{ix} X_t + W_{ih} h_{t-1} + W_{ic} c_{t-1} + b_i) \quad (2)$$

Where It is the value of the input gate, Xt is the current input, $ht-1$ is the previous hidden state, Wi and Ui are the weights, and bi is the bias.

3. Memory Update (The main component of LSTM that allows the network to retain information over long periods of time)

$$c_t = f_t \circ c_{t-1} + i_t \circ \emptyset\,(W_{cx} x_{t-1} + W_{ch} h_{t-1} + b_c) \tag{3}$$

This is done by multiplying the value of the forget gate with the previous memory cell $ct-1$, and multiplying the value of the input gate with the new value Ct. The new memory cell is then obtained by adding the results of these two operations.

4. Output Gate (Controlling how much the values in the memory cell are used to compute the output of the LSTM)

$$o_t = \sigma(W_{ox}\, X_t + W_{oh}\, h_{t-1} + W_{oc}\, c_t + b_o) \tag{4}$$

$$h_t = o_t \circ \emptyset\,(c_t) \tag{5}$$

This involves two steps: first, deciding how the memory cell will be affected by its current value, and second, generating the output ht. by using the tanh activation function on the updated memory cell, which is then multiplied by the output gate [25].

2.5. Gated Recurrent Unit (GRU)

The Gated Recurrent Unit (GRU) is another variant of the RNN architecture that addresses short-term memory problems and offers a simpler structure compared to LSTM. GRU can only control information within the unit as it does not have additional memory cells to store information. The GRU unit consists of three main components: the update gate, the reset gate, and the current memory content. These gates allow GRU to selectively update and utilize information from previous time steps, enabling it to capture long-term sequential dependencies. The formulation of GRU is as follows:

$$Z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \tag{6}$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \tag{7}$$

$$\tilde{h}_t = \tanh(W_h \cdot [r_{t-1}, x_t] + b_c) \tag{8}$$

$$h_t = (1 - Z_t) \times h_{t-1} + Z_t \times \tilde{h}_t \tag{9}$$

Description,

Z_t	: gate update
r_t	: reset gate
$\tilde{h}_t$	: candidate hidden state
h_t	: final hidden state
h_{t-1}	: state
σ	: sigmoid gate network layer
tanh	: network layer activation function tanh
W_h, W_z, W_r	: weight matrix

2.6. Bidirectional Gated Recurrent Unit (Bi-GRU)

The Bi-GRU is the most advanced type of Recurrent Neural Network (RNN) with less complexity compared to other types of RNN. Bi-GRU is a modification of GRU that allows information to flow forward and backward in a sequence of data. This is achieved by using two sets of GRU units working together: one set flows forward through the sequence of data, while the other flows backward. In this way, Bi-GRU can capture information from both directions in the data sequence, which is often beneficial in tasks such as sentiment analysis where the context of words before and after can influence overall understanding. Bi-GRU functions as a better and more efficient Window-Based Feature Extractor. Bi-GRU Architecture for Sentiment Classification view as figure 5.

The Input Layer consists of 5000 words from all reviews, which are fed into the Embedding Layer to create a 16-dimensional vector for each word. Two Bi-GRU Layers are used to classify the text shown in Figure 3. The Output Layer consists of nodes with Sigmoid Activation function as it is a Binary Classification.

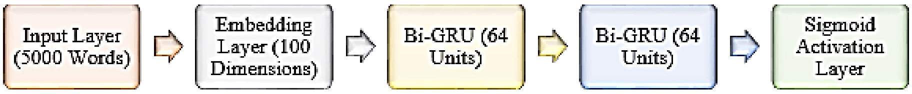

Figure 3. Bi-GRU Architecture for Sentiment Classification.

2.7. Performance Evaluation

Performance evaluation on sentiment classification is important to identify positive or negative reviews from customers. Evaluation criteria for measuring classification performance that are often used are F1-Score, Accuracy, Recall, Precision, and Computational Time. The Accuracy, Precision, Recall, and F1-Score equations are shown in equations (10), (11), (12), and (13).

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{10}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{11}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{12}$$

$$\text{F1} - \text{Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} \times \text{Recall}} \tag{13}$$

Description,

Accuracy : The closeness of the measured value to the predicted result
Precision : Positive predictive value
Recall : Positive sensitivity value
F1 − Score : A measure of the balanced combination of Precision and Recall
TP (True Positive) : The number of positive cases that were correctly identified
TN (True Negative) : The number of negative cases that were correctly identified
FP (False Positive) : The proportion of negative data identified as positive
FN (False Negative): The proportion of positive data that is identified as negative

3. RESULTS AND DISCUSSION

In this research, the data used is a collection of Indonesian product review data labeled with emotions and sentiments. The data is collected from one of the giant e-commerce in Indonesia called Tokopedia. The dataset contains product reviews of one product on Tokopedia that uses Indonesian language. Each product review is annotated with Customer Rating, to support further research. The methods that will be used in this research are LSTM, Bidirectional-GRU, GRU.

Training Sample and Test Sample were divided into 80%: 20% by using the Holdout Technique. For all three Deep Learning Architectures, model training is performed for 10 Epochs with a Batch Size of 64, and the RMSProp Optimizer so as not to increase the occurrence of computational errors (Loss) using Binary Cross-Entropy during the training stage in all three Deep Learning Architectures.

Evaluation and Comparison of the three Deep Learning Architectures is done with Classification Report which provides a better interpretation of the research results. Accuracy Metrics of the three Deep Learning Architectures, using Precision, Recall, Accuracy, F1-Score, and Computational Time can be seen in the table 1 below.

Table 1. Performance Evaluation of Bi-GRU, GRU, and LSTM Architecture Models.

Evaluation Measure	Bi-GRU	GRU	LSTM
Accuracy	0,93	0,92	0,92
Precision	0,96	0,91	0,91
Recall	0,89	0,93	0,93
F1-Score	0,92	0,92	0,92
Computational Time (Second)	0,035	0,019	0,032

From the table above, it can be observed that Bi-GRU has Accuracy of 0.93, Precision of 0.96 R, better than the other two methods which both have Accuracy of 0.92 and Recall of 0.91. GRU and LSTM both have a Recall of 0.93, 4% better than Bi-GRU. In terms of F1-Score, all three Deep Learning Architectures have the same value, which is 0.92. GRU has a Computational Time performance of 0.019 seconds, better than LSTM which have Computational Time performance of 0.0032 seconds, and Bi-GRU which have Computational

Time performance of 0.035 seconds, which makes GRU better than the other two methods in terms of Computational Time.

The ROC curves of the three Deep Learning Architectures show that Bi-GRU performs 1% better than GRU and LSTM to classify positive and negative sentiments from Tokopedia Marketplace customer reviews, because Bi-GRU has unique characteristics in finding patterns or slogans in a sentence. It can be observed that, Bi-GRU has outperformed all other approaches for sentiment classification for its unique characteristics to find pattern or catchphrase in sentences. Bi-GRU have also performed better than the traditional method [1] [4] used for sentiment classification on Marketplace Customer Review dataset. From the above discussions, we can conclude that Bi-GRU is the best-suited architecture for performing sentiment classification on Tokopedia Customer Reviews.

4. CONCLUSION

Based on the results of sentiment analysis research using deep learning methods BI-GRU, GRU, and LSTM has been successfully carried out. The data processing technique is separated into Train Sample and Testi Sample divided into 80%: 20% by using the Holdout Technique. The accuracy results on the BI-GRU algorithm with 0.93% and Precision of 0.96, better than the other two methods which both have Accuracy of 0.92 and Recall of 0.91. GRU and LSTM both have a Recall of 0.93, 4% better than Bi-GRU. In terms of F1-Score, the three Deep Learning Architectures have the same value, which is 0.92. From the results of the discussion above, there are still some tuning parameters that can be tried to improve the accuracy of the model built. Some of these tuning parameters are trying to vary the number of units used, then the size of the feature data formed, the number of epochs, the dropout value used, and the length of the processed data sequence.

REFERENCES

[1] Luh Kadek Budi Martini, Ejasa Sembiring, and Frandy Paulus, “Customer Online Customer Rating Dan Online Customer Review Terhadap Keputusan Pembelian Di Toko Modeliafashion Pada Marketplace Tokopedia Jakarta,” J. Appl. Manag. Account. Sci., vol. 4, no. 1, pp. 15–24, 2022, doi: 10.51713/jamas.v4i1.67.

[2] R. Sutoyo, S. Achmad, A. Chowanda, E. W. Andangsari, and S. M. Isa, “PRDECT-ID: Indonesian product reviews dataset for emotions classification tasks,” Data Br., vol. 44, p. 108554, 2022, doi: 10.1016/j.dib.2022.108554.

[3] H. Wan, M. Q. Mei, J. Yan, J. Xiong, and L. Wang, “How does apology matter? Responding to negative customer reviews on online-to-offline platforms,” Electron. Commer. Res. Appl., vol. 61, no. July, p. 101291, 2023, doi: 10.1016/j.elerap.2023.101291.

[4] R. D. Irawan and R. . Nurlinda, “Pengaruh E-Service Quality, Online Customer Review dan Sales Promotion Terhadap Minat Beli Ulang,” J. Manaj. dan Bisnis Ekon., vol. 1, no. 4, pp. 1–25, 2023, doi: 10.54066/jmbe-itb.v1i4.660.

[5] O. Manullang, C. Prianto, and N. H. Harani, “Analisis Sentimen Untuk Memprediksi Hasil Calon Pemilu Presiden Menggunakan Lexicon Based Dan Random Forest,” J. Ilm. Inform., vol. 11, no. 02, pp. 159–169, 2023, doi: 10.33884/jif.v11i02.7987.

[6] 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, 2022, doi: 10.7717/PEERJ-CS.1044.

[7] A. Adak, B. Pradhan, and N. Shukla, “Sentiment Analysis of Customer Reviews of Food Delivery Services Using Deep Learning and Explainable Artificial Intelligence: Systematic Review,” Foods, vol. 11, no. 10, 2022, doi: 10.3390/foods11101500.

[8] C. Intelligence and Neuroscience, “Retracted: Real-Time Twitter Spam Detection and Sentiment Analysis using Machine Learning and Deep Learning Techniques,” Comput. Intell. Neurosci., vol. 2023, pp. 1–1, 2023, doi: 10.1155/2023/9810910.

[9] F. Fitroh and F. Hudaya, “Systematic Literature Review: Analisis Sentimen Berbasis Deep Learning,” J. Nas. Teknol. dan Sist. Inf., vol. 9, no. 2, pp. 132–140, 2023, doi: 10.25077/teknosi.v9i2.2023.132-140.

[10] R. Naquitasia, D. H. Fudholi, and L. Iswari, “Analisis Sentimen Berbasis Aspek pada Wisata Halal dengan Metode Deep Learning,” J. Teknoinfo, vol. 16, no. 2, p. 156, 2022, doi: 10.33365/jti.v16i2.1516.

[11] W. Widayat, “Analisis Sentimen Movie Review menggunakan Word2Vec dan metode LSTM Deep Learning,” J. Media Inform. Budidarma, vol. 5, no. 3, p. 1018, 2021, doi: 10.30865/mib.v5i3.3111.

[12] M. A. H. Wadud, M. M. Kabir, M. F. Mridha, M. A. Ali, M. A. Hamid, and M. M. Monowar, “How can we manage Offensive Text in Social Media - A Text Classification Approach using LSTM-BOOST,” Int. J. Inf. Manag. Data Insights, vol. 2, no. 2, p. 100095, 2022, doi: 10.1016/j.jjimei.2022.100095.

[13] B. Subramanian, B. Olimov, S. M. Naik, S. Kim, K. H. Park, and J. Kim, “An integrated mediapipe-optimized GRU model for Indian sign language recognition,” Sci. Rep., vol. 12, no. 1, pp. 1–16, 2022, doi: 10.1038/s41598-022-15998-7.

[14] R. Sowmiya, G. Sivakamasundari, and V. Archana, "Facial Emotion Recognition using Deep Learning Approach," Int. Conf. Autom. Comput. Renew. Syst. ICACRS 2022 - Proc., vol. 2022, pp. 1064–1069, 2022, doi: 10.1109/ICACRS55517.2022.10029092.

[15] R. K. Hario Laskito Ardi,Eko Sediyono, "Support Vector Machine Classifier for Sentiment Analysis of Feedback Marketplace with a Comparison Features at Aspect Level," Int. J. Innov. Res. Adv. Eng., vol. 4, no. 11, 2017, doi: 10.26562/IJIRAE.2017.NVAE10080.

[16] A. N. Rohman, R. Luviana Musyarofah, E. Utami, and S. Raharjo, "Natural Language Processing on Marketplace Product Review Sentiment Analysis," 2020 2nd Int. Conf. Cybern. Intell. Syst. ICORIS 2020, 2020, doi: 10.1109/ICORIS50180.2020.9320827.

[17] N. R. Bhowmik, M. Arifuzzaman, and M. R. H. Mondal, "Sentiment analysis on Bangla text using extended lexicon dictionary and deep learning algorithms," Array, vol. 13, p. 100123, 2022, doi: 10.1016/j.array.2021.100123.

[18] M. M.Abdelgwad, T. H. A Soliman, A. I.Taloba, and M. F. Farghaly, "Arabic aspect based sentiment analysis using bidirectional GRU based models," J. King Saud Univ. - Comput. Inf. Sci., vol. 34, no. 9, pp. 6652–6662, 2022, doi: 10.1016/j.jksuci.2021.08.030.

[19] B. Liu, "Sentiment Analysis Essentials," Cambridge Univ. Press, 2015.

[20] B. Liu, M. Hu, and J. Cheng, "Opinion observer," p. 342, 2005, doi: 10.1145/1060745.1060797.

[21] D. H. Wahid and A. SN, "Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity," IJCCS (Indonesian J. Comput. Cybern. Syst., vol. 10, no. 2, p. 207, 2016, doi: 10.22146/ijccs.16625.

[22] G. Chellasamy, S. K. Arumugasamy, and S. Govindaraju, "Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID- 19 . The COVID-19 resource centre is hosted on Elsevier Connect , the company ' s public news and information ," Diabetes Metab Syndr., vol. 14(4), no. January, pp. 337–339, 2020.

[23] P. Sugiartawan, A. A. Jiwa Permana, and P. I. Prakoso, "Forecasting Kunjungan Wisatawan Dengan Long Short Term Memory (LSTM)," J. Sist. Inf. dan Komput. Terap. Indones., vol. 1, no. 1, pp. 43–52, 2018, doi: 10.33173/jsikti.5.

[24] A. K. Gangwar and V. Ravi, "A Novel BGCapsule Network for Text Classification," SN Comput. Sci., vol. 3, no. 1, pp. 1–14, 2022, doi: 10.1007/s42979-021-00963-4.

[25] F. M. Shiri, T. Perumal, N. Mustapha, and R. Mohamed, "A Comprehensive Overview and Comparative Analysis on Deep Learning Models: CNN, RNN, LSTM, GRU," 2023, doi: 10.48550/arXiv.2305.17473.