Basic and Applied Computing for Digital Innovation (BACODING)
Volume 01, No. 02, August, 2026, pp. 49-57
Received: 25 July 2026; Revised: 13 August 2026; Accepted: 16 August 2026
https://bacoding.id/index.php/journal/
E-ISSN: 3163-5881
DOI:

Research Article

User Sentiment Analysis Towards Online Loan Services Based on Social Media Data X Using the Bert Method

Dwy Nafila Radjaloa[1], Saiful do Abdullah[2], Muhammad Fhadli[3]
[1,2,3]Department of Informatics, Faculty of Engineering, Universitas Khairun, Ternate 97728, Indonesia

* **Correspondence:** Email: saiful.abdullah@unkhair.ac.id

Abstract: Shopee PayLater (SPayLater) is an online lending service widely used by Indonesians. The high usage of this service has led to diverse user opinions regarding the experience of using SPayLater on social media platform X. The large number of comments generated makes the manual analysis process inefficient, so an automated method capable of quickly and accurately recognizing user sentiment tendencies is needed. This study aims to analyze user sentiment towards the SPayLater service and evaluate the performance of the IndoBERT model in classifying comments based on social media platform X data. The research data consists of 800 comments collected manually using keywords related to SPayLater. Next, the data went through preprocessing stages including data cleaning, case folding, tokenizing, and normalization before being divided into training data and test data for the IndoBERT model fine-tuning process. The fine-tuned IndoBERT model achieved an accuracy of 97.50%, a precision of 97.56%, a recall of 97.50%, and an F1-score of 97.46%. In the classification results, negative sentiment dominated with 392 comments, followed by neutral sentiment with 293 comments and positive sentiment with 115 comments. This achievement proves that IndoBERT is capable of classifying user sentiment with a very high level of accuracy, making this method suitable for sentiment analysis of online lending services based on social media data.

Keywords: Sentiment Analysis, Online Loans, SPaylater, Social Media X, IndoBERT

1. Introduction

The rapid development of the digital era has made technology a vital part of everyday life, transforming the way people interact, work, and communicate. The advent of the Industrial Revolution 4.0 has spurred innovations across many sectors, including the financial sector, which has given rise to new business models and disrupted the conventional banking system. One important innovation emerging from this development is Financial Technology (Fintech), which offers more efficient and secure financial services through modern transaction processes. Among the various Fintech services available, online loans (pinjol) have become a popular solution for people seeking fast and convenient credit access [1].

Online loans can be defined as financial services that utilize applications or information technology

and operate entirely online. These services fall under the category of financial technology (Fintech), which is a financial innovation developed by non-bank institutions to provide faster and more accessible financial access. The rapid development of Fintech has also driven the emergence of online loans, making it easier for people to access financial services [2].

The growth of online lending services is marked by the use of the internet as a medium for transactions, including in various banking activities. One social media platform frequently used by the public to express opinions and experiences regarding online lending services is the app X, formerly known as Twitter. The number of X users continues to increase annually, generating enormous volumes of data. This phenomenon is known as Big Data, which refers to large-scale, rapidly changing, and complex data sets that require specialized processing techniques for effective and meaningful use, one of which is sentiment analysis [2]-[3].

Sentiment analysis is a process used to determine public opinion regarding a particular object. In the context of online lending, sentiment analysis is crucial for understanding public perception of these services [4] – [6], particularly regarding the various risks that frequently arise, such as high interest rates, personal data leaks, and fraudulent activities that harm users. Therefore, this study was conducted to determine public opinion regarding online lending applications through the social media platform X using sentiment analysis. The results can be used to understand the risks and potential for fraud in online lending services [7].

Sentiment analysis can be performed using either deep learning or machine learning [8]. Studies [9–10] have applied deep learning models such as LSTM, whereas [11–16] have employed machine learning models. The BERT (Bidirectional Encoder Representations From Transformers) method is a transformer-based architecture model designed to deeply understand word context using a bidirectional approach, reading sentences from left to right and right to left simultaneously. This bidirectional capability allows BERT to capture the meaning and nuances of words in a sentence more accurately, thus proving effective for various tasks such as text classification and sentiment analysis [9]. In this study, IndoBERT, a BERT variant specifically retrained using an Indonesian corpus, is used to classify comments from social media users regarding online lending services into positive, negative, and neutral sentiment categories.

Based on the problem description, this study focused on the Shopee PayLater (SPayLater) service. This service is one of the most widely used online lending products by Indonesians and has a high comment volume on social media platform X, making it a representative target for sentiment analysis. IndoBERT was chosen as the text classification approach due to its ability to understand Indonesian context bidirectionally and its ability to produce good performance in sentiment analysis tasks. With this approach, user sentiment in this study is grouped into three categories, namely positive, negative, and neutral.

2. Method and materials

2.1 Dataset

The dataset is the initial and fundamental step in this research. The news dataset was collected from various sources to ensure sufficient content variety for the clustering and classification process. A total of 80.121 data sets were collected.

2.2 Model Usulan

The research stages were carried out systematically to ensure the process ran smoothly. The research flow can be seen in Figure 1.

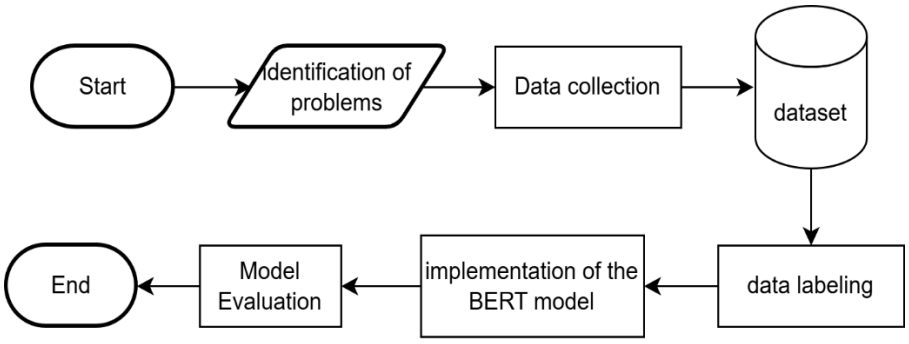

Figure 1. Research Stage Flowchart

2. 1 Problem Identification

The problem addressed in this study is how to identify and analyze user sentiment towards the Shopee PayLater (SPayLater) service based on opinions posted on social media platform X. This analysis aims to determine public perceptions of the quality of SPayLater's service.

2. 2 Data Collection

Data collection in this study was conducted manually through direct observation on social media platform X. The data collected consisted of user comments or posts discussing the Shopee PayLater (SPayLayer) service using search keywords such as SPayLater, Shopee PayLater, PayLater bills, and PayLater interest. The data collection process involved reading and transcribing each relevant comment into a structured Excel file to facilitate further data processing. This resulted in 800 comments, which were then divided into 80% training data and 20% test data before sentiment labeling and modeling using IndoBERT.

2. 3 Data Preprocessing

This stage is crucial in sentiment analysis, as the comment data obtained from social media platform X is still in raw text form, containing various irrelevant elements, such as punctuation, emojis, abbreviations, and non-standard words. The preprocessed data can be seen in Table 1.

Table 1. Comments to be processed

Original Comment	Sentiment label
"*SPayLater* prosesnya cepat banget, limit langsung aktif 👍 "	Positive
"Bunga *SPayLater* lumayan tinggi bikin berat"	Negative
"*SPayLater* sudah terdaftar OJK jadi lebih aman"	Neutral

2. 4 Data Labelling

The data labeling stage is carried out to assign a sentiment category to each comment that has undergone preprocessing. This labeling is important because it will serve as a reference in training the IndoBERT model to distinguish comments based on their polarity. Labeling is done manually (manual

annotation) by reading each comment and determining its sentiment polarity based on the meaning of the opinion contained in the text. In this study, three categories were used: positive, negative, and neutral. The results of the data labeling for the three sample categories can be seen in Table 2.

Table 2. Data Labeling Results

Original Comment	*Result of Preprocessing*	Sentiment Label
"*SPayLater* prosesnya cepat banget, limit langsung aktif 👍 "	*SPayLater* proses cepat limit aktif	Positive
"Bunga *SPayLater* lumayan tinggi bikin berat"	bunga *SPayLater* tinggi berat	Negative
"*SPayLater* sudah terdaftar OJK jadi lebih aman"	*SPayLater* daftar ojk aman	Neutral

2. 5 Implemented indobert

The BERT model implementation used in this study was IndoBERT with the indobenchmark/indobert-base-p1 pre-trained model, a BERT variant developed specifically for Indonesian language processing. IndoBERT was chosen based on its ability to understand Indonesian context more accurately than general-language (multilingual) BERT models, making it more suited to the characteristics of the Indonesian comment data used in this study. The implementation process was carried out through fine-tuning, adjusting the IndoBERT pre-trained model to suit the characteristics of the research data. The stages of BERT model implementation in this study are as follows:

1. Text Tokenization: The processed comment data is converted into tokens using the BERT tokenizer. A special token [CLS] is added at the beginning of the text and [SEP] at the end as a marker.
2. Conversion to Numeric Form: The resulting tokens are converted into numeric IDs according to the IndoBERT dictionary, then adjusted for length using padding or truncation, with a maximum token length of 128.
3. Embedding Formation: Each token is represented as a vector consisting of a token embedding, a position embedding, and a segment embedding.
4. Bidirectional Encoding: The IndoBERT encoder processes the embedding vector bidirectionally to understand the context of the word from both the left and right sides.
5. Classification Layer: The final token representation [CLS] is fed into the classification layer, using the softmax activation function to generate sentiment class probabilities.
6. Model Fine-Tuning: All IndoBERT parameters are retrained using this research dataset to match the characteristics of comment data related to the Shopee PayLater (SPayLater) service.
7. Sentiment Prediction: The model predicts sentiment as positive, negative, or neutral based on the highest probability.
8. Mathematically, BERT utilizes a self-attention mechanism calculated using the equation shown in equation 1.

$$Attention\ (Q,K,V) = softmax\left(\frac{KQ^T}{\sqrt{d_k}}\right) V \quad (1)$$

Where Q (Query), K (Key), and V (Value) are the input vector representations, while d_k is the vector dimension. This mechanism allows the model to understand the relationship between words contextually.

The final token representation [CLS] is then fed into the classification function, the formula for which can be seen in equation 2.

$$y = softmax\ (wh + b) \quad (2)$$

The highest probability value determines the sentiment class, the formula for which can be seen in equation 3

$$y = argmax(y) \quad (3)$$

2. 6 Model Evaluation

Model evaluation was conducted to assess the extent to which the fine-tuned IndoBERT model accurately classified sentiment in the test data. The evaluation process was conducted by comparing the model's predictions against the actual labels in the test data to determine the degree of agreement between the model's predictions and the actual data. Based on this comparison, four evaluation metrics were calculated: accuracy, precision, recall, and F1-score.

Given that this study involved three sentiment classes (positive, negative, and neutral), precision, recall, and F1-score values were first calculated for each class and then combined into a single final value using a weighted average method, which is an average weighted by the amount of data in each class. This allows the evaluation results to better reflect the model's overall performance, considering the proportion of data in each sentiment class.

3. Result And Discussion

3.1. Dataset

The data used in this study were obtained manually through direct observation on social media platform X using search keywords such as SPayLater, Shopee PayLater, PayLater bills and PayLater interest, then copied into a structured format in the form of an Excel file to obtain 800 comment data used as the main dataset in this study. The dataset was then divided into 640 training data (80%) and 160 test data (20%) using the train_test_split function from the Scikit-Learn library. The results of the identification of irrelevant elements such as emojis, punctuation, symbols, non-standard abbreviations that require a preprocessing stage before they can be processed by the model. Sample comment data along with sentiment labels used in this study can be seen in Table 3.

Table 3. Dataset crawling results

COMMENT DATA	SENTIMENT LABEL
ucap seseorang yg udh lunas tagihannya tp masih tergoda bebelian	Negatif
hari ni doang kan lu, besok jg make lagi	Negatif
bener ka, sekarang bawaannya ga tenang banget	Negatif
sama banget kaaak huhu rindu banget hidup tenang	Negatif
jujur iya banget, tapi InsyaAllah aku akhir tahun debut free	Netral
aku juga kangen, walaupun kere kaya gembel, nyesel banget dari kecil sampe numpuk	Negatif
Spaylatter gue stress banget sih ga lagi2 minjemin akun buat sodara kredit hp mana bayar suka telat	Negatif
sama banget ka.. kangen bgt masa-masa hidup tenang dan ga was-was sama teror	Negatif
Bener bangettt, masa-masa yg aku kangenin sebelum kena panjul. Hidup lebih tenang	Negatif

3.2. Data distribution

The sentiment distribution bar chart is visualized using the matplotlib library to illustrate the distribution of data across each sentiment category in the dataset. Each sentiment category is displayed in a different color: green for positive sentiment, red for negative sentiment, and orange for neutral sentiment, for ease of visualization. This bar chart shows the proportional comparison of data between the three sentiment classes. For more clarity, see Figure 2.

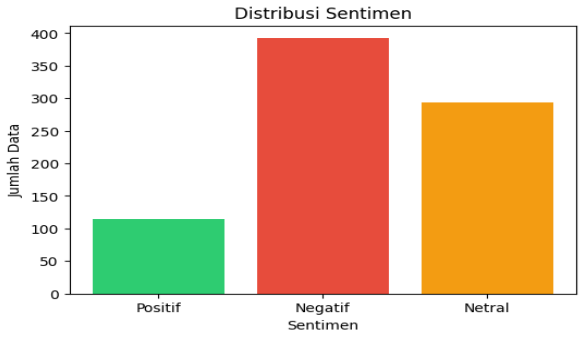

Figure 2. data distribution each class

3.3. Indobert implementation results

The BERT model was implemented using the IndoBERT variant with the indobenchmark/indobert-base-p1 pre-trained model. The implementation process included text tokenization, data conversion to the PyTorch dataset format, and fine-tuning using Trainer from the Transformers library. Briefly, the testing parameters used in the IndoBERT model training process in this study are presented in Table 4 below.

Table 4. BERT Model Testing Parameters

Parameter	Value
Model pra-training	'indobenchmark/indobert-base-p1
Number of classes	3
Training and test data ratio	80:20
Maximum token length	128
Total *epoch*	3
Training batch size	8
Evaluation batch size	8
Training equipment	GPU T4

After the parameter settings were applied, testing was carried out with five variations of epoch values, namely 3, 5, 10, 30, and 100, using other parameters that still refer to Table 4. The test results can be seen in Table 5 below.

Table 5. Test Results with Epoch Value Variations

Epoch	*Accuracy*	*Precision*	*Recall*	F1-*Score*
3	97,50%	97,56%	97,50%	97,46%
5	85,62%	87,17%	85,62%	86,05%
10	80,00%	81,26%	80,00%	79,60%
30	80,62%	80,44%	80,62%	80,45%
100	77,50%	77,99%	77,50%	77,57%

Based on Table 5, the best performance was achieved at epoch = 3 with an accuracy of 97.50%,

precision of 97.56%, recall of 97.50%, and an F1-score of 97.46%. Model performance then tended to decline as the number of epochs increased, reaching its lowest point at epoch = 100 with an accuracy of 77.50%. This decline indicates overfitting, where the model increasingly adapts to the training data, resulting in a decline in its generalization ability to the test data. This is a normal occurrence considering the relatively small sample size of 800. Based on these results, epoch = 3 was chosen as the final configuration because it produced the best performance across all evaluation metrics.

In addition to testing the number of epochs, additional testing was conducted on the batch size to ensure optimal training parameters. Testing was conducted on a fixed number of epochs, namely epoch = 3, by comparing three batch sizes, namely 5, 8, and 10. The batch size also affects the number of training steps, namely the number of model weight updates performed during the training process, calculated by the number of training data divided by the batch size, then multiplied by the number of epochs. Because the number of training data used is fixed, namely 640 data, the smaller the batch size used, the more steps are produced. The test results are presented in Table 6 below.

Table 6. Batch Size Test Results on Epoch 3

Batch size	**Total Step**	***Accuracy***	***Precision***	***Recall***	**F1-*Score***
5	384	88,12%	88,09%	88,12%	88,10%
8	240	97,50%	97,56%	97,50%	97,46%
10	192	88,12%	88,72%	88,12%	88,34%

Based on Table 6, batch size affects the number of model weight updates during training. Smaller batch sizes result in more frequent but less stable weight updates, while larger batch sizes result in more stable but less frequent updates. A batch size of 8 provides the optimal balance between the two, resulting in the highest classification performance compared to batch sizes of 10 and 5. Based on these results, a batch size of 8 was chosen as the final parameter.

3.4. Confussion matrix results

The confusion matrix was visualized using the sns.heatmap function from the Seaborn library to show a comparison between the model's predictions and the actual labels for each sentiment class. The confusion matrix results in Figure 4 represent testing with an initial epoch value of 3.

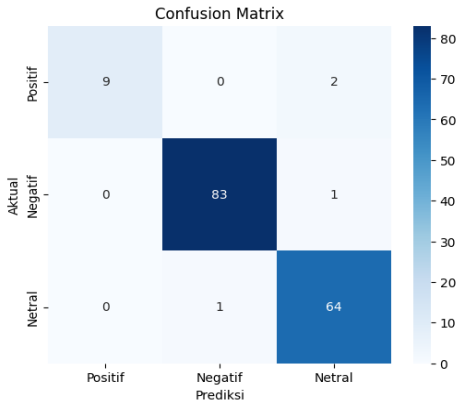

Figure 4. *Confusion Matrix result*

Based on Figure 4, the model demonstrates its best performance in classifying the negative class, with 83 out of 84 data points correctly predicted. The neutral class follows, with 64 out of 65 data points correctly predicted, while the positive class achieves 9 correct predictions out of 11, with 2 positive instances misclassified as neutral. This misclassification stems from the fact that positive comments in

the dataset are far less numerous than those in other classes, and their linguistic patterns sometimes resemble those of neutral comments.

4. Conclusion

Based on the research results, it can be concluded that the sentiment analysis of social media users X towards the Shopee PayLater (SPayLater) service produced three sentiment categories from 800 comment data, namely positive with 115 comments, negative with 392 comments, and neutral with 293 comments. These results indicate that most users gave a negative response to the SPayLater service, especially regarding the interest rate that is considered high and billing problems. The IndoBERT model used in this study was able to classify sentiment well, as evidenced by the accuracy value of 97.50%, precision 97.56%, recall 97.50%, and F1-score 97.46%. This performance is superior to the Support Vector Machine (SVM), Naïve Bayes and Random Forest methods in previous studies. In addition, future research needs to handle unbalanced data to see comparisons with this study.

5. Reference

[1]. Ardiansyah, M. A., Alamsyah, M., & Arif, M. F. (2024). Analisis Sentimen Twitter Tentang Pinjaman Online di Indonesia Menggunakan Metode Random Forest. *Jutisi : Jurnal Ilmiah Teknik Informatika Dan Sistem Informasi*, 13(2), 1410. https://doi.org/10.35889/jutisi.v13i2.2215

[2]. Utami, D. S., & Erfina, A. (2021). Analisis Sentimen Pinjaman Online di Twitter Menggunakan Algoritma Support Vector Machine (SVM). *SISMATIK (Seminar Nasional Sistem Informasi Dan Manajemen Informatika)*, 1(1), 299–305.

[3]. Hosnia, A., (2026). Automated News Classification System Using K-Means Clustering-Based Labeling and LSTM Deep Learning. *Basic and Applied Computing for Digital Innovation*. 1(2). pp, 41-48.

[4]. Nurdin, T. A., Alexandri, M. B., Sumadinata, W., & Arifianti, R. (2023). Sentiment analysis of user preference for old vs new fintech technology using SVM and NB algorithms. *Management Systems in Production Engineering*, 31(4), 373–380. https://doi.org/10.2478/mspe-2023-0041

[5]. Cui, Y., & Liu, L. (2022). Investor sentiment-aware prediction model for P2P lending indicators based on LSTM. *PLoS ONE*, 17(1), e0262539. https://doi.org/10.1371/journal.pone.0262539

[6]. Siering, M. (2023). Peer-to-peer (P2P) lending risk management: Assessing credit risk on social lending platforms using textual factors. *ACM Transactions on Management Information Systems*, 14(3). https://doi.org/10.1145/3589003

[7]. Victory, T., & Wazaumi, D. D. (2026). Analisis Sentimen Pengguna Aplikasi Pinjaman Online Melalui Media Sosial X (Twitter) Menggunakan Metode Support Vector Machine. *Jurnal Minfo Polgan*, 14(2), 2116–2130. https://doi.org/10.33395/jmp.v14i2.15300

[8]. Karanikola, A., Davrazos, G., Liapis, C. M., & Kotsiantis, S. (2023). Financial sentiment analysis: Classic methods vs. deep learning models. *Intelligent Decision Technologies*, 17, 893–915. https://doi.org/10.3233/IDT-230478

[9]. Imaduddin, H., A'la, F. Y., & Nugroho, Y. S. (2023). Sentiment analysis in Indonesian healthcare applications using IndoBERT approach. *International Journal of Advanced Computer Science and Applications*, 14(8). https://doi.org/10.14569/IJACSA.2023.0140813

[10]. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - *Proceedings of the Conference*, 1(Mlm), 4171–4186

[11]. Ahmad, K. (2023). Analisis Sentimen Pinjaman Online Akulaku dan Kredivo dengan Metode Support Vector Machine (SVM). *Jurnal Mandalika Literature*, 4(4), 323–332. https://doi.org/10.36312/jml.v4i4.2045

[12]. Alta Zahir, M. R., Ramdani, B., & Dwi Saputra, A. (2024). Analisis Sentimen Terhadap Ulasan Aplikasi Pinjaman Online (PINJOL) di Google Play Store Menggunakan Naive Bayes Classifier. *Jurnal Riset Informatika dan Teknologi Informasi*, 1(2), 61–64. https://doi.org/10.58776/jriti.v1i2.39

[13]. Ghozali, M. I., Sugiharto, W. H., & Iskandar, A. F. (2023). Analisis Sentimen Pinjaman Online Di Media Sosial Twitter Menggunakan Metode Naive Bayes. *KLIK: Kajian Ilmiah Informatika dan Komputer*, 3(6), 1340–1348. https://doi.org/10.30865/klik.v3i6.936

[14]. Afandi, R., Afdal, M., Novita, R., & Mustakim, M. (2024). Analisis Sentimen Masyarakat Terhadap Pinjaman Online di Twitter Menggunakan Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor. *Building of Informatics, Technology and Science (BITS)*, 6(2), 596–605. https://doi.org/10.47065/bits.v6i2.5300

[15]. Arya Wibisono, Y., Afdal, M., & Novita, R. (2024). Implementasi Algoritma Random Forest Untuk Analisa Sentimen Data Ulasan Aplikasi Pinjaman Online di Google Play Store. *Building of Informatics, Technology and Science (BITS)*, 6(2). https://doi.org/10.47065/bits.v6i2.5350

[16]. Gunadarma, I. K. A. L., Sasmita, G. M. A., & Trisna, I. N. P. (2025). Analisis Kinerja K-Nearest Neighbors (K-NN) untuk Analisis Sentimen Aplikasi Pinjaman Online X. *Jurnal Ilmiah Merpati (Menara Penelitian Akademika Teknologi Informasi)*, 12(3), 204. https://doi.org/10.24843/jim.2024.v12.i03.p07