Spektrum Industri
Vol. 23, No. 2, 2025, pp. 293-305
ISSN 1693-6590
https://journal3.uad.ac.id/index.php/spektrum/index

Performance Analysis of the IndoBERT–Prophet Hybrid Model for Logistics Applications

Erna Mulyati *, Maniah, Noviana, Nia Pardede

Universitas Logistik dan Bisnis Internasional, Bandung, 40151, Indonesia
* Corresponding Author: ernamulyati@ulbi.ac.id

ARTICLE INFO

Article history
Received September 12, 2025
Revised October 12, 2025
Accepted October 27, 2025

Keywords
Courier services;
Digital transformation;
E-commerce;
IndoBERT;
Logistics performance;

ABSTRACT

The increasing competition in Indonesia's logistics sector, particularly in digital courier applications, highlights the need for advanced analytical tools capable of understanding and predicting customer sentiment in real time. However, current sentiment analysis methods often lack contextual depth and predictive capability, limiting their practical value for decision-making. This study aims to develop and validate an integrated analytical framework that combines diagnostic and predictive analytics for logistics performance evaluation. The framework integrates a fine-tuned IndoBERT model for sentiment classification and a Prophet model for time-series forecasting, allowing the analysis of user reviews while accounting for external service disruptions. Empirical validation was conducted using 924 user reviews from the PosAja! application by PT Pos Indonesia. The IndoBERT model achieved an impressive 99.5% accuracy, effectively identifying two main complaint categories: application functionality issues and delivery delays. The Prophet forecasting component successfully modeled sentiment trends, revealing spikes in negative sentiment that strongly correlated with technical service disruptions, such as COD feature failures and server maintenance. The results confirm the framework's robustness in both diagnosing and forecasting sentiment dynamics. User sentiment proved to be a sensitive real-time indicator of service stability and operational performance. The validated IndoBERT–Prophet hybrid framework provides a novel, data-driven approach for proactive decision-making and continuous service improvement in the logistics industry.

1. Introduction

The logistics service sector plays a vital role in supporting global economic activities, especially amidst rapid technological advancement and the ongoing Industrial Revolution 4.0. Intense competition compels companies to innovate and continuously improve service quality to meet increasingly complex customer needs (Winkelhaus & Grosse, 2020). In Indonesia, empirical studies have confirmed that logistics service quality strongly influences customer satisfaction and loyalty (Khairi & Cahyadi, 2023). Thus, evaluating digital service quality to maintain customer satisfaction has become an essential topic for companies like PT Pos Indonesia (Damanik et al., 2025; Maulana et al., 2025).

This competitive landscape is heavily influenced by the rapid growth of e-commerce, which has significantly transformed consumer behavior in Indonesia (E Mulyati et al., 2025). Reviews generated from these transactions are crucial because online word of mouth has been proven to directly influence consumer perceptions and purchasing decisions (Alnoor et al., 2024; Handoyo, 2024; Wahyuningjati & Purwanto, 2024). Furthermore, feedback from digital platforms is a key driver for evaluating and improving supply chain and logistics service performance (Suali et al., 2024; Tan et al., 2024). This phenomenon is supported by technological advancement and increasing internet penetration (Apriani et al., 2021). The resulting surge in online transactions has led to the flourishing of businesses in the courier service industry, which has begun to experience increasingly tight competition (Mulyati & Hamidin, 2022).

The method of goods delivery, as a result of increasing e-commerce transactions, has led to the flourishing of businesses in the courier service industry. The courier service industry has begun to experience increasingly tight competition since the emergence of various courier service providers (Mulyati & Hamidin, 2022). According to the 2025 Top Brand Index for Courier Services in Indonesia, the courier service market is experiencing significant growth, driven by the corporate sector, which includes central and regional governments, companies, and various industries.


COURIER SERVICES

Brand	TBI	
J&T Express	33.30%	TOP
JNE Express	29.10%	TOP
TIKI	10.60%	TOP
Pos Indonesia	7.30%	
DHL	7.20%	

Fig. 1. Top Brand Index of Courier Services in Indonesia in 2025
Source: https://www.topbrand-award.com

Based on Fig. 1, This intense competition presents challenges for established players like PT Pos Indonesia. Market data from the 2025 Top Brand Index shows that while the company maintains a presence (7.3%), it ranks fourth behind major competitors J&T Express (33.3%), JNE Express (29.1%), and TIKI (10.6%), highlighting the competitive pressure in the traditional courier market.

In the digital domain, the challenge is even greater, as shown in Table 1. PT Pos Indonesia's official application, PosAja!, demonstrates significantly lower user adoption (500 thousand+ downloads) compared to competitors like J&T Express (50 million+) and My JNE (5 million+). This gap in digital market penetration signals potential issues in meeting user expectations and serves as a relevant case study to validate the need for more advanced analytical frameworks capable of diagnosing and addressing user sentiment effectively within the competitive Indonesian logistics landscape.

Despite periodic updates to the PosAja! application, negative user sentiments and low-star ratings remain consistently high. This recurring pattern suggests a fundamental misalignment between developer improvements and user expectations, indirectly affecting the company's operational and logistics performance (Septyarani & Nurhadi, 2023). While previous studies have attempted to analyze user feedback using classical machine learning techniques like Support Vector Machine (SVM) (Alemerien et al., 2024; Guido et al., 2024; Obiedat et al., 2022; Sukma Rukmana & Wiwik Handayani, 2023; Talaat, 2023), these methods face significant limitations. Firstly, models like SVM struggle to capture complex contextual nuances in user reviews due to their reliance on "bag-of-words" representations, making them less effective in interpreting negation, sarcasm, or intricate linguistic structures. More critically, these traditional approaches are primarily diagnostic and reactive, focusing on analyzing past sentiments rather than proactively forecasting future trends. This limits their strategic value in a dynamic market. The absence of an integrated analytical framework capable of combining deep contextual diagnosis with proactive predictive forecasting represents a significant methodological gap in analyzing digital logistics service performance in Indonesia.

Table 1. Data on the Number of Digital Logistics Application Downloads

Ranking	Applications	Number of Downloads
1	Gojek (GoSend)	100 million +
2	J&T Express Indonesia	50 million +
3	Grab (GrabExpress)	15 million +
4	Lalamove	10 million +
5	MyJNE	5 million +
6	Paxel	1 million +
7	Deliveree	1 million +
8	PosAja!	500 thousand +
9	SiCepatExpress	100 thousand +

To overcome these limitations, this research proposes an innovative integrated analytical framework. This framework uniquely combines two state-of-the-art components: (1) the IndoBERT model, leveraging its deep bidirectional contextual understanding for accurate diagnostic sentiment analysis specifically tailored for the Indonesian language, and (2) the Prophet model for proactive time-series forecasting, capable of incorporating external event impacts (Yaqoob et al., 2023). The novelty lies in the synergistic integration of these advanced diagnostic and predictive capabilities within a single, unified framework, designed to provide a holistic and actionable understanding of user sentiment dynamics for the PosAja! application.

Therefore, the primary objective of this research is to design and empirically validate the effectiveness of this novel integrated analytical framework, utilizing user reviews from the PosAja! application as a case study for validation. The main contribution lies in providing a validated and measurable methodology for logistics companies to leverage predictive sentiment analysis as a real-time performance indicator, thereby enabling better data-driven decision-making for continuous service improvement in the digital era.

To process this data, this research is based on the evolution of technology in the field of Natural Language Processing (NLP). Sentiment analysis is a method used to automatically extract, understand, and process text data to determine the underlying feelings or attitudes within an opinion (Liu, 2020). In the context of digital services, this technique is highly useful for evaluating public response to service quality based on reviews from social media and other digital platforms. Sentiment analysis, also commonly referred to as opinion mining, is a field of study that analyzes people's opinions, sentiments, evaluations, attitudes, and emotions toward entities such as products, services, and topics (Medhat et al., 2014).

A more granular approach within this field is Aspect-Based Sentiment Analysis (ABSA) (Draskovic & Milanovic, 2025; Yang et al., 2025), which allows businesses to extract and interpret this information to gain detailed insights into specific product and service attributes. The crucial first step in ABSA is aspect extraction, which aims to pinpoint and extract the critical aspects of a product or service referenced in the text. This process is conceptually identical to the topic identification performed in this study. The current state-of-the-art has shifted from classic machine learning methods like Support Vector Machine (SVM), which, like other methods such as Naïve Bayes and Decision Trees, has limitations in understanding context, towards deep learning models based on the Transformer architecture (Lin, 2023), such as BERT (Bidirectional Encoder Representations from Transformers) (Kim, 2022).

In general, deep learning approaches have revolutionized many fields due to their ability to learn data representations at multiple levels of abstraction (LeCun et al., 2015). This model offers superior contextual understanding capabilities. This process is conceptually identical to the topic identification conducted in this study. The use of Transformer-based models such as BERT has been proven highly effective for Aspect-Based Sentiment Analysis (ABSA), even by constructing auxiliary sentences to enhance the model's focus on specific aspects (Akhtar et al., 2019; Imron et al., 2023).

Furthermore, this research also integrates the pillar of Predictive Analytics as a step to move from diagnostic analysis (what happened) to proactive analysis (what will happen). Sukma Rukmana & Wiwik Handayani (2023) specifically describe a Service Science Design framework for analyzing how Customer Experience on the PosAja! application can be engineered to increase user satisfaction. This aligns with foundational marketing theory, which posits that customer satisfaction is a key measure of how well a service meets or exceeds user expectations (Jarumaneeroj et al., 2023).

That study underscores that improving application performance is not just a technical matter of bug fixes, but a strategic effort that must be rooted in a holistic understanding of the user's journey and perceptions. Although previous research has applied SVM to PosAja! reviews, a research gap exists in building an integrated, end-to-end analytical framework. To enhance proactive customer satisfaction prediction, recent research has proposed multimodal learning frameworks that leverage big data, integrating user behavior, geographic, and product attribute data, to significantly improve accuracy in consumer satisfaction forecasting across cross-border e-commerce contexts (Zhang & Guo, 2024). This study fills that gap by combining three technical pillars, (1) high-accuracy sentiment diagnosis with IndoBERT, (2) automated root cause identification through Topic Modeling (a form of aspect extraction), and (3) proactive trend forecasting via predictive analysis, which are holistically aimed at accelerating logistics performance improvement for a strategic state-owned enterprise in Indonesia.

2. Method

This study employs the Knowledge Discovery in Databases (KDD) method, a systematic process for extracting valuable insights from large datasets. The KDD process in this study is adapted for the purpose of sentiment and predictive analysis, which includes five main stages: (1) Data Selection, (2) Data Pre-processing, (3) Data Transformation, (4) Data Mining, and (5) Evaluation & Interpretation as shown in Fig. 2.

All of these stages will be implemented using the Python programming language via the Google Colaboratory platform, with the core IndoBERT model used for text analysis.

2.1. Data Selection

The initial dataset consists of 1200 reviews, which, after being labeled based on the score, are distributed as follows:

1. 460 Positive reviews (38.3%)
2. 413 Negative reviews (34.4%)
3. 327 Neutral reviews (27.3%)

This distribution still shows that the number of negative sentiments is significant and nearly matches the number of positive sentiments, which strengthens the justification for further analysis. After this cleaning process, a final valid sample of 924 unique user reviews was obtained for analysis.

These 924 reviews are distributed into 354 positive reviews, 318 negative reviews, and 252 neutral reviews. This relatively balanced distribution between the positive and negative sentiment classes is an ideal condition for training a classification model to reduce potential bias. Nevertheless, the high number of negative sentiments serves as a strong foundation for conducting an in-depth investigation into the primary sources of user dissatisfaction. To conduct this investigation accurately, a reliable classification model is required. Therefore, the next stage is to evaluate the performance of the pre-trained BERT model to ensure its validity before it is used for further analysis.

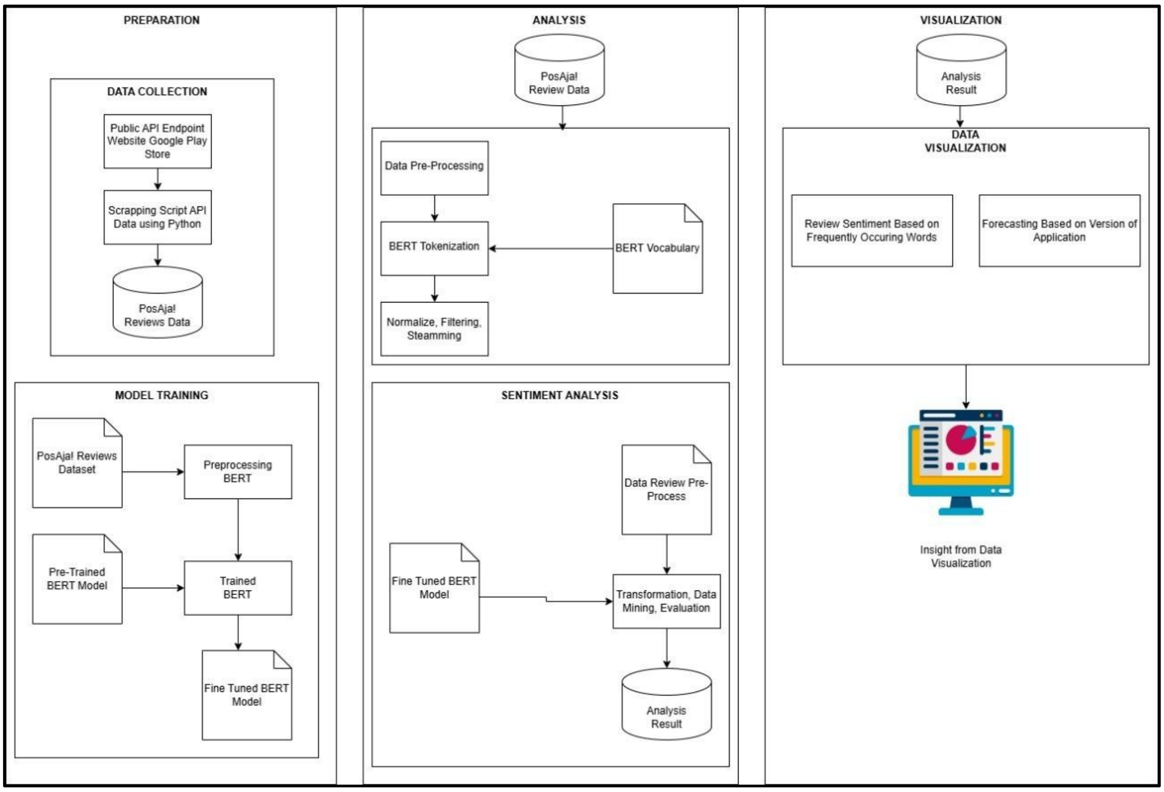

Fig. 2. The proposed method

2.2. Data Processing and Mining Framework

The methodological framework began with a Pre-processing stage to ensure data quality. This involved cleaning the dataset by handling missing values, such as imputing the app version based on chronological data, and removing duplicate reviews to obtain a final, valid dataset. The review text itself was cleaned through a series of steps including case folding and normalization. For the forecasting analysis, the review data was then aggregated to create a monthly time-series dataset of negative sentiments.

Following this, the Data Transformation stage restructured the clean data for modeling. The primary transformation was sentiment labeling, where the 1 – 5star score was converted into numerical categories (Negative, Neutral, Positive). Additionally, the app version text was split into numerical features to enrich the dataset for the classification model.

Finally, the Data Mining stage applied three analytical techniques to the prepared data. First, a state-of-the-art IndoBERT model was trained for 10 epochs to perform sentiment classification on the user reviews. Second, an N-gram frequency analysis was conducted on the negatively classified reviews to perform topic analysis and identify the root causes of user complaints. Third, a **Prophet** time-series model was trained on the monthly negative sentiment data to forecast trends, a model widely recognized for its effectiveness and applicability in business forecasting applications (Abdullah et al., 2021), incorporating extra regressors to account for the impact of specific service disruptions.

3. Results and Discussion

3.1. Model Analysis

The analysis stage in this research begins with a descriptive analysis to understand the composition and characteristics of the user review dataset. The distribution of the number of reviews based on manually assigned sentiment labels is presented in Fig. 3.

Fig. 3. Distribution of User Review Sentiment Labels

From a small total sample of 924 reviews, 354 positive, 318 negative, and 252 neutral reviews were identified. This relatively balanced distribution between the positive and negative sentiment classes is an ideal condition for training a classification model to reduce potential bias. Nevertheless, the high number of negative sentiments serves as a strong foundation for conducting an in-depth investigation into the primary sources of user dissatisfaction. To conduct this investigation accurately, a reliable classification model is required. Therefore, the next stage is to evaluate the performance of the pre-trained BERT model to ensure its validity before it is used for further analysis.

In the pre-processing stage, the sentiment classification model was trained for 10 epochs with the objective of optimizing its ability to recognize sentiment patterns. The effectiveness of this learning process was carefully monitored by observing the decrease in the loss value at each epoch. The consistent decrease in the loss value, as visualized in Fig. 3, indicates that the model successfully converged and effectively learned the data representation.

After the training process was completed, the model's performance was quantitatively tested on a test set to assess its ability to classify sentiment on new data. The comprehensive evaluation results, which include precision, recall, and f1-score metrics for each class, are presented in the form of a Classification Report in Fig. 4 and Fig. 5.

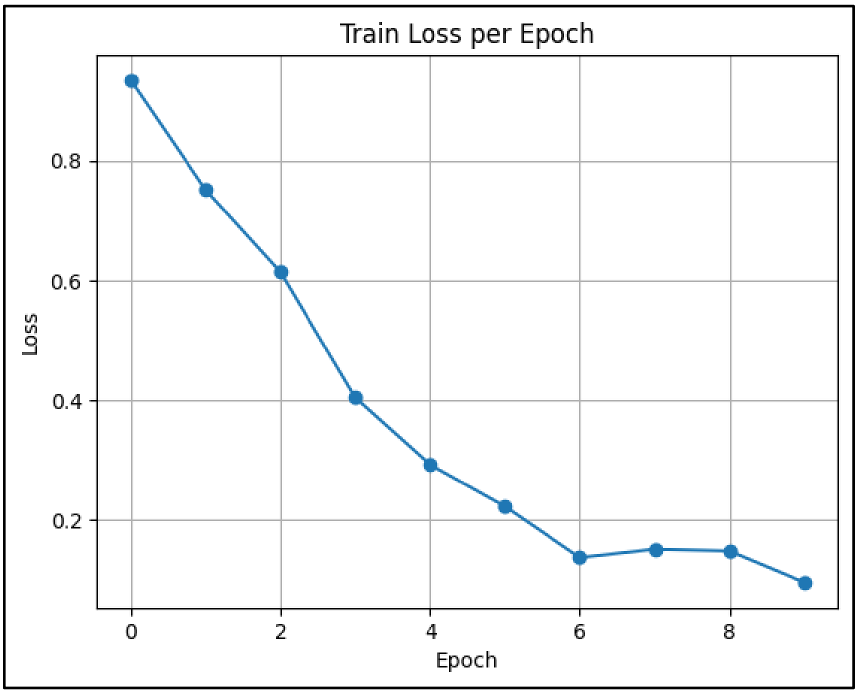

Fig. 4. Loss Value Graph During Model Training

Classification Report (Tabular):

	precision	recall	f1-score	support
Negatif	1.0000	1.0000	1.0000	74.0000
Netral	0.8889	1.0000	0.9412	8.0000
Positif	1.0000	0.9902	0.9951	102.0000
accuracy	0.9946	0.9946	0.9946	0.9946
macro avg	0.9630	0.9967	0.9788	184.0000
weighted avg	0.9952	0.9946	0.9947	184.0000

Fig. 5. Classification Report

Overall, the model demonstrated highly satisfactory performance with an accuracy rate reaching 99.5%. To validate the superiority of this result, the model's performance was compared against a classic machine learning method, Support Vector Machine (SVM), which has been applied in previous relevant research. The IndoBERT model's accuracy significantly outperforms the SVM model, which achieved an accuracy of 88.1% in a study on PosAja! application reviews (Ghatora et al., 2024). This performance gap empirically proves that the Transformer architecture in IndoBERT, with its ability to understand bidirectional context, is superior in capturing the nuances of user review sentiment compared to the bag-of-words approach commonly used by SVM.

This indicates that the model is almost always correct in predicting the sentiment of a review. When analyzed further per class, the model is capable of classifying Negative and Positive sentiments with near-perfect F1-scores of 1.000 and 0.995, respectively. These values signify that the model is highly reliable in detecting these two primary sentiments. For the Neutral class, although the metrics are high, these results must be interpreted with caution. The very small amount of supporting data (only 8 samples) makes it less representative, and the high performance could be due to the small amount of data during testing. Based on the combination of a convergent training process and very high evaluation metrics, especially for the Negative and Positive classes, the model is deemed highly feasible and valid. Therefore, this model will be used in the next analysis stage to identify the main topics from reviews with negative sentiment.

3.2. Negative Sentiment Analysis

After the classification model was validated and proven to have reliable performance in the previous sub-chapter, the next stage of analysis focuses on answering the main research question: what are the most frequently expressed complaint topics by application users? For this purpose, the pre-trained BERT model is applied to filter and analyze reviews that are specifically identified as having negative sentiment.

To extract complaint topics from these reviews, a word frequency analysis was conducted to identify the key terms most often used in the context of complaints. The results of this word frequency analysis on negative reviews are presented visually in Fig. 6.

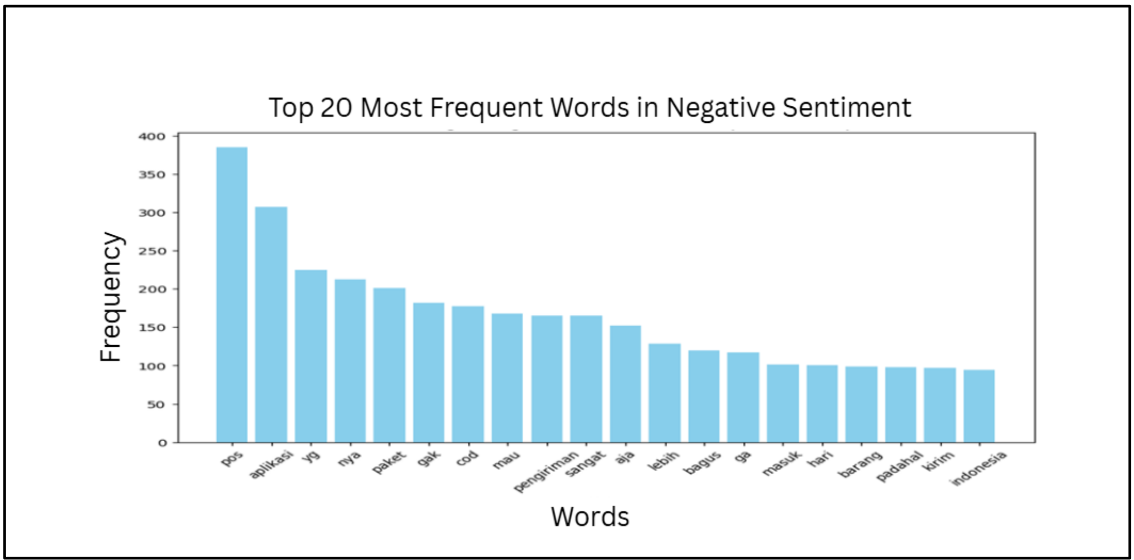

Fig. 6. Top 20 Word Frequency Distribution in Negative Reviews, Highlighting Application and Delivery Issues

Based on Figure 6, these findings can be interpreted through the lens of Service Quality theory (Shaiful, 2020; Zeithaml, 1988). Complaints related to "application functionality" (e.g., 'gak masuk,' 'error') directly reflect quality dimensions of the digital platform, such as reliability and ease of use. Meanwhile, complaints surrounding the "delivery process" (e.g., 'paket gak sampai,' 'pengiriman lama') highlight deficiencies in the core logistics service quality dimension, namely timeliness. This finding confirms that in the context of e-logistics, customer perception of quality is multidimensional, encompassing both digital and physical experiences.

3.3. Forecasting Sentiment Analysis

As a further analysis, time-series modeling was conducted to understand the patterns and trends in the number of negative sentiments over time. The results from the modeling using Prophet are presented in Fig. 7.

The graph displays the historical data of monthly complaint numbers (black dots) along with the model's prediction line (blue line) and its uncertainty interval (light blue area). Visually, it can be observed that the volume of negative sentiment moved in a relatively stable fluctuation pattern from 2020 until the end of 2024. However, a significant anomaly was detected in the period of December 2024 to January 2025, where a very drastic spike in complaints occurred, far exceeding the previous historical pattern.

Vol. 23, No. 2, 2025, pp. 293-305

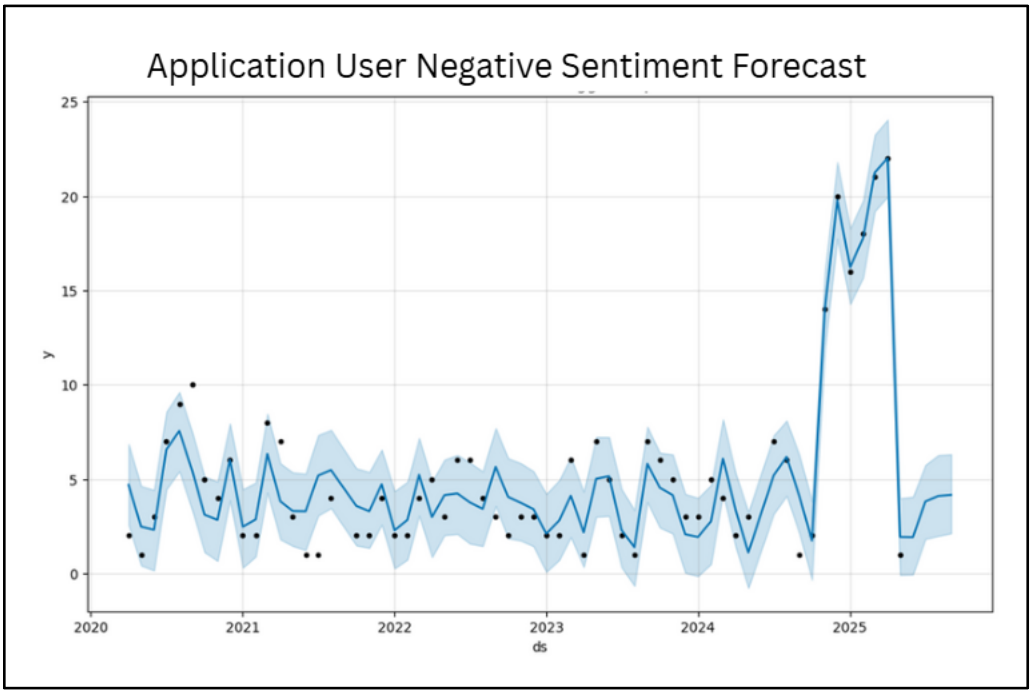

Fig. 7. Negative Sentiment Trend Forecast, showing a Significant Anomaly Correlated with Service Disruptions

This sharp spike was successfully validated and found to be strongly correlated with real-world events confirmed by internal parties at Pos Indonesia. Based on the information received, during the anomalous period, there was a systemic disruption to the Cash on Delivery (COD) feature, which prevented users from placing orders. In addition, the application also experienced service instability due to a server maintenance process that caused an 'on/off' condition. The combination of these two crucial technical problems directly triggered a wave of complaints from users, which is clearly recorded in the data.

This trend analysis effectively proves that the volume of negative sentiment is a very sensitive and fast-reacting indicator of technical service stability. This finding underscores how disruptions to core features or server problems can instantly damage the user experience and be recorded in the review data. Furthermore, the model projects that if no similar disruptions occur in the future, the number of complaints is expected to return to a lower, normal level, as shown by the post-spike prediction line.

A deeper analysis of the forecast model's components, as presented in Fig. 8, can provide further quantitative evidence. The trend graph (top panel) shows that the long-term trend of negative sentiment was actually tending to decrease from 2020 to 2025. Meanwhile, the yearly component (middle panel) identifies a recurring seasonal pattern each year. The most significant is the bottom panel (extra_regressors_additive), which isolates the impact of the service disruption. This panel shows that the drastic spike at the end of 2024 was almost entirely caused by external factors (COD and server disruptions), not by a change in the natural trend or seasonal patterns. Time-series modeling quantitatively proves that the volume of complaints is highly responsive to service stability, with a strong correlation between spikes in negative sentiment and periods of technical service disruptions, such as problems with the COD feature and servers.

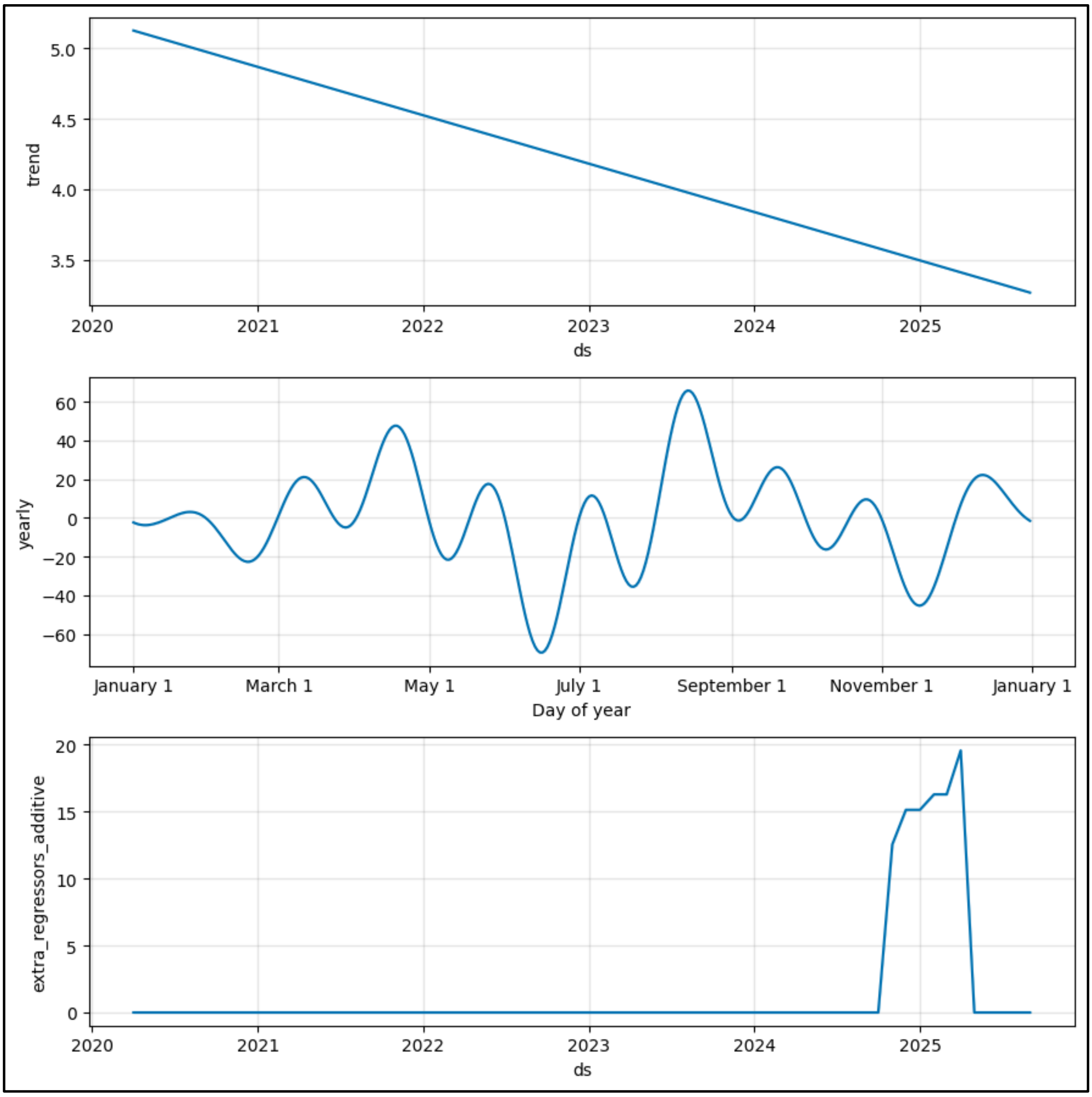

Fig. 8. Details of the Forecasting of Negative Sentiment Trends for 2020 – 2025

4. Conclusion

This research has successfully designed and validated an integrated analytical framework that is proven effective for diagnosing problems and predicting customer sentiment trends at PT Pos Indonesia. The findings show that the fine-tuned IndoBERT model is capable of performing with 99.5% accuracy, validating its capability as a reliable diagnostic tool. The diagnosis identified that the main user complaints are focused on two crucial areas: application functionality issues and obstacles in the delivery process. Furthermore, the predictive analysis using Prophet quantitatively proves a strong correlation between spikes in negative sentiment and periods of technical service disruptions. The primary contribution of this work is the demonstration that this integrated framework, which combines high-accuracy diagnostics with predictive forecasting, can serve as a measurable, data-driven accelerator for improving logistics performance. Based on the research findings, it is recommended that PT Pos Indonesia prioritize evidence-based improvements in the two areas proven to be the main sources of complaints: the stability and functionality of the application, as well as the reliability of core services like COD and delivery timeliness. The analytical framework developed in this study can be implemented as an early warning system to proactively detect operational issues by monitoring shifts in customer sentiment (He et al., 2013), allowing for faster response and mitigation. This effort should be supported by developing internal human resource capacity in data analytics and establishing a data-driven feedback loop for continuous improvement.

This study has several limitations. The analysis was based on data from a single source (the Google Play Store) and the topic identification relied on N-gram frequency rather than more advanced semantic clustering techniques. These limitations open several avenues for future research. Subsequent studies could deepen the predictive model by comparing various techniques like machine learning or digital twins and integrating it with real-time operational data from IoT technologies. Additionally, it is recommended to conduct a similar analysis on positive sentiment reviews to identify service strengths and to validate this framework through real-world field experiments to measure its long-term impact on cost efficiency and customer satisfaction.

Author Contribution: Erna Mulyati was responsible for the conceptualization and supervision. Maniah and Noviana conducted the data analysis and methodology. Nia Pardede was responsible for the literature review and writing the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors would like to thank the Institute for Research and Community Service (LPPM) of the International Logistics and Business University for the support provided for this research.

Conflict of interest: The authors declare that there is no conflict of interest.

References

Abdullah, A. S., Ruchjana, B. N., Jaya, I. G. N. M., & Soemartini. (2021). Comparison of SARIMA and SVM model for rainfall forecasting in Bogor city, Indonesia. *Journal of Physics: Conference Series*, *1722*(1), 012061. https://doi.org/10.1088/1742-6596/1722/1/012061

Akhtar, M. S., Chauhan, D., Ghosal, D., Poria, S., Ekbal, A., & Bhattacharyya, P. (2019). Multi-task Learning for Multi-modal Emotion Recognition and Sentiment Analysis. *Proceedings of the 2019 Conference of the North*, 370–379. https://doi.org/10.18653/v1/N19-1034

Alemerien, K., Al-Ghareeb, A., & Alksasbeh, M. Z. (2024). Sentiment analysis of online reviews: a machine learning based approach with TF-IDF vectorization. *Journal of Mobile Multimedia*. https://ieeexplore.ieee.org/abstract/document/10976588/

Alnoor, A., Tiberius, V., Atiyah, A. G., Khaw, K. W., & Yin, T. S. (2024). How positive and negative electronic word of mouth (eWOM) affects customers' intention to use social commerce? A dual-stage multi group-SEM and ANN analysis. *International Journal of Human-Computer Interaction*. https://doi.org/10.1080/10447318.2022.2125610

Apriani, E., Indriani, Y., & Adawiyah, R. (2021). Pengambilan Keputusan, Sikap Dan Kepuasan Konsumen Terhadap Paket Nasi Liwet Di Rmsasa Bandar Lampung. *Jurnal Ilmu-Ilmu Agribisnis*, *8*(2), 325. https://doi.org/10.23960/jiia.v9i2.5106

Damanik, E. O. P., Tarigan, W. J., Tampubolon, J., & Simanjuntak, D. C. Y. (2025). Digital Transformation of PT Pos: Customer Loyalty In The Era of Security and Data Technology. *Jurnal Ekuilnomi*. http://jurnal.usi.ac.id/index.php/ekuilnomi/article/view/1299

Draskovic, D., & Milanovic, S. (2025). Aspect-based sentiment analysis of user-generated content from a microblogging platform. In *Journal of Big Data*. Springer. https://doi.org/10.1186/s40537-025-01244-0

Ghatora, P. S., Hosseini, S. E., Pervez, S., Iqbal, M. J., & Shaukat, N. (2024). Sentiment Analysis of Product Reviews Using Machine Learning and Pre-Trained LLM. *Big Data and Cognitive Computing*, *8*(12), 199. https://doi.org/10.3390/bdcc8120199

Guido, R., Ferrisi, S., Lofaro, D., & Conforti, D. (2024). An overview on the advancements of support vector machine models in healthcare applications: a review. In *Information*. mdpi.com. https://www.mdpi.com/2078-2489/15/4/235

Handoyo, S. (2024). Purchasing in the digital age: A meta-analytical perspective on trust, risk, security, and e-WOM in e-commerce. In *Heliyon*. cell.com. https://www.cell.com/heliyon/fulltext/S2405-8440(24)05745-1

He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management*, *33*(3), 464–472. https://doi.org/10.1016/j.ijinfomgt.2013.01.001

Imron, S., Setiawan, E. I., & Santoso, J. (2023). Aspect based sentiment analysis marketplace product reviews using BERT, LSTM, and CNN. In *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*. academia.edu. https://www.academia.edu/download/115178772/767.pdf

Jarumaneeroj, P., Ramudhin, A., & Barnett Lawton, J. (2023). A connectivity-based approach to evaluating port importance in the global container shipping network. *Maritime Economics & Logistics*, *25*(3), 602–622. https://doi.org/10.1057/s41278-022-00243-9

Khairi, L. I., & Cahyadi, E. R. (2023). Pengaruh Logistics Service Quality Terhadap Customer Satisfaction dan Customer Loyalty Pada Pengguna JNE dan J&T Express di Jabodetabek. *Jurnal Aplikasi Bisnis Dan Manajemen*, *9*(2), 671. https://doi.org/10.17358/jabm.9.2.671

Kim, J. (2022). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. mlg.postech.ac.kr. http://mlg.postech.ac.kr/~jtkim/courses/2022-spring-trends-in-ml/materials/05_bert.pdf

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lin, W. (2023). *BERT pre-training of deep bidirectional transformers for language understanding GPT generative pre-trained transformer LaMDA language model for*.

Liu, B. (2020). *Sentiment Analysis*. Cambridge University Press. https://doi.org/10.1017/9781108639286

Maulana, M. N., Maniah, M., & Lestiani, M. E. (2025). The Influence of Digital Competence and ITC to Digital Transformation and the Implication on Company Performance in Tenggarong Branch Post Office. *Journal of Advanced*. http://karyailham.com.my/index.php/arca/article/view/250

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

Mulyati, E, Rachmatullah, M. I. C., & Firmansyah, A. S. (2025). Sentiment Analysis of Pospay Application Reviews Using the Bert Deep Learning Method. In *Jurnal Teknik Informatika*. https://journal.uinjkt.ac.id/ti/article/view/41116

Mulyati, E., & Hamidin, D. (2022). Pemetaan Layanan Jasa E-Commerce Di Kota Bandung Menggunakan Metode Multidimensional Scaling. *Matrik: Jurnal Manajemen Dan Teknik Industri Produksi*, *23*(1), 39. https://doi.org/10.30587/matrik.v23i1.3596

Obiedat, R., Qaddoura, R., Ala'M, A. Z., & Al-Qaisi, L. (2022). Sentiment analysis of customers' reviews using a hybrid evolutionary SVM-based approach in an imbalanced data distribution. *Ieee Access*. https://ieeexplore.ieee.org/abstract/document/9706209/

Septyarani, T. A., & Nurhadi, N. (2023). Pengaruh Kualitas Pelayanan dan Kepuasan Pelanggan terhadap Loyalitas Pelanggan. *Widya Cipta: Jurnal Sekretari Dan Manajemen*, *7*(2), 218–227. https://doi.org/10.31294/widyacipta.v7i2.15877

Shaiful, G. N. (2020). Servqual: A multiple-item scale for measuring consumer perc. In *J. Retail*.

Suali, A. S., Srai, J. S., & Tsolakis, N. (2024). The role of digital platforms in e-commerce food supply chain resilience under exogenous disruptions. In *Supply Chain Management: An International Journal*. emerald.com. https://doi.org/10.1108/SCM-02-2023-0064

Sukma Rukmana, & Wiwik Handayani. (2023). Analysis of Service Science Design through Customer Experience to Increase PosAja Application User Satisfaction. *Kontigensi: Jurnal Ilmiah Manajemen*, *11*(2). https://doi.org/10.56457/jimk.v11i2.402

Talaat, A. S. (2023). Sentiment analysis classification system using hybrid BERT models. In *Journal of Big Data*. Springer. https://doi.org/10.1186/s40537-023-00781-w

Tan, J., Wong, W. P., Tan, C. K., Jomthanachai, S., & Lim, C. P. (2024). Blockchain-based Logistics 4.0: enhancing performance of logistics service providers. *Asia Pasific Journal of Marketing and Logistics*. https://doi.org/10.1108/APJML-07-2023-0650

Wahyuningjati, T., & Purwanto, E. (2024). Exploring the influence of electronic word of mouth and customer reviews on purchase decisions: A study of trust as a mediating factor in the Shopee Marketplace. In *MindVanguard. Behav*. library.acadlore.com. https://library.acadlore.com/MVBB/2024/2/2/MVBB_02.02_01.pdf

Winkelhaus, S., & Grosse, E. H. (2020). Work Characteristics in Logistics 4.0: Conceptualization of a qualitative assessment in order picking. *IFAC-PapersOnLine*, *53*(2), 10609–10614. https://doi.org/10.1016/j.ifacol.2020.12.2816

Yang, R., Ju, X., Guo, C., Ding, L., Li, M., & Zhang, B. (2025). A unified review of aspect sentiment triplet extraction methods in aspect-based sentiment analysis. *Knowledge and Information Systems*. https://doi.org/10.1007/s10115-025-02519-x

Yaqoob, A., Aziz, R. M., & Verma, N. K. (2023). Applications and techniques of machine learning in cancer classification: a systematic review. In *Human-Centric Intelligent Systems*. Springer. https://doi.org/10.1007/s44230-023-00041-3

Zeithaml, V. A. (1988). *Servqual: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality dalam Journal of Retailing: Spring*.

Zhang, X., & Guo, C. (2024). Research on Multimodal Prediction of E-Commerce Customer Satisfaction Driven by Big Data. *Applied Sciences*, *14*(18), 8181. https://doi.org/10.3390/app14188181