ENHANCING COMPLAINT MANAGEMENT THROUGH INFORMATION SYSTEMS: LSTM-BASED AUTOMATIC CLASSIFICATION OF BANK CUSTOMER COMPLAINTS IN INDONESIA

Timotius Pangaribuan
Universitas Sumatera Utara, Indonesia
E-mail: timotiusp187@gmail.com*

Muhammad Anggia Muchtar
Universitas Sumatera Utara, Indonesia
E-mail: anggia.muchtar@usu.ac.id

Mohammad Andri Budiman
Universitas Sumatera Utara, Indonesia
E-mail: mandrib@usu.ac.id

Receive :
Accepted :
Published :
DOI : 10.30829/jipi.v10i1.24493

Abstract
This study develops an automatic classification system for customer complaints in the banking sector using the Long Short-Term Memory (LSTM) deep learning method. A dataset comprising 4,714 customer complaint entries was collected from Bank Sumut's internal communication records, categorized into six major complaint types. The data underwent comprehensive preprocessing, including cleaning, tokenization, and vectorization. A supervised learning approach was applied using an LSTM-based neural network architecture, and the model's performance was evaluated using accuracy, precision, recall, and F1-score metrics. The results demonstrated a classification accuracy of 100% on the test set, with the model successfully categorizing free-text complaints into predefined categories. The findings highlight the strong potential of LSTM models in supporting automated text-based customer service operations within digital banking environments, particularly for Indonesian-language complaint datasets. Further research is recommended to validate the model on unseen real-world data and to address challenges related to data imbalance.

Keywords: *Text Classification; Information Systems; Customer Complaints; LSTM; Deep Learning; Natural Language Processing (NLP); Digital Banking*

INTRODUCTION

The rapid development of digital banking services has transformed customer interaction channels, with a significant shift toward electronic communication platforms such as email, live chat, and mobile applications. In this context, customer complaints represent a valuable source of feedback for service quality improvement and customer retention strategies (Ali et al., 2021; Gonroos, 2007; Dalimunthe, 2018). However, the high volume and unstructured nature of complaint data present significant challenges for manual classification and resolution processes (Ritonga et al., 2024). Customer complaints are a critical indicator of service quality in the banking industry. In an increasingly competitive financial market, a bank's ability to respond to complaints effectively is strongly correlated with customer satisfaction and loyalty. However, manual complaint classification and resolution processes still dominate many financial institutions in Indonesia. This reliance on manual methods results in slow response times, classification inconsistencies, and risks to institutional reputation due to service delays (Hirschman, 1970; Zhao et al., 2021; Zhou et al., 2021).

Efficient management of customer complaints is critical for maintaining customer satisfaction and loyalty, particularly in the competitive banking industry (Utami & Pratama, 2020; Harianto et al., 2023). Manual complaint processing is time-consuming, prone to human error, and insufficient for handling the growing volume of customer feedback. Therefore, automated classification systems are increasingly needed to support operational efficiency and service excellence. According to the Financial Services Authority of Indonesia (OJK), more than 9,000 digital customer complaints were reported to banks in 2023, yet only 67% were resolved within 10 working days (OJK, 2023). A regional survey conducted by Deloitte (2022) also found that 58% of banking customers in Southeast Asia would consider switching banks if their complaints were not addressed promptly. These findings underscore the urgency of implementing efficient complaint management systems to sustain competitive advantage and customer trust (Kotler & keller, 2016; Ritonga et al., 2023).

Natural Language Processing (NLP) techniques, particularly those leveraging deep learning, offer promising solutions for text classification tasks. Among various deep learning approaches, Long Short-Term Memory (LSTM) networks have demonstrated superior capabilities in modeling sequential textual data and understanding context within complaint narratives (Kim, 2014). In the academic realm, Natural Language Processing (NLP) has emerged as a promising approach for analyzing unstructured text data such as customer complaints (Hassan & Fatima, 2020). Deep learning techniques, especially Long Short-Term Memory (LSTM) networks, have demonstrated strong performance in sequential text classification tasks due to their ability to retain long-term contextual information. The foundational work by Hochreiter and Schmidhuber (1997) paved the way for LSTM's widespread application in various NLP tasks (Goodfellow et al., 2016).

Despite the growing use of LSTM in NLP, most prior studies have focused on English-language datasets and global commercial contexts. For instance, research by Zhang et al. (2021) and Dey et al. (2022) primarily examined e-commerce complaints in the United States and India. In the Indonesian context, studies such as Kurniawan et al. (2022) have begun exploring LSTM-based classification for public service feedback, but there remains a lack of research focused specifically on banking complaints, which involve distinct language patterns, regulatory frameworks, and customer service dynamics.Previous studies have primarily focused on English-language datasets and generic complaint contexts, with limited exploration of localized datasets in non-English languages such as Bahasa Indonesia. Moreover, few studies have specifically addressed complaint classification within the Indonesian banking sector, where linguistic nuances and customer service structures differ significantly (Vaswani et al., 2017).

This study aims to fill the identified research gap by developing and evaluating an LSTM-based model for the automatic classification of customer complaints written in Bahasa Indonesia, specifically within the banking sector. The main objectives are: (1) to design a robust text classification model tailored to the characteristics of complaints received by Indonesian banks, and (2) to evaluate the model's effectiveness in accurately categorizing complaints into predefined categories (Kowsari et al., 2019). The study seeks to contribute to the academic discourse on non-English text classification and offer practical solutions for digital transformation in Indonesia's banking sector From an academic perspective, this research expands the application of deep learning to a linguistically unique context, addressing the underrepresentation of Bahasa Indonesia in NLP studies (Nguyen & Simkin et al., 2020; Otter et al., 2021) Furthermore, it introduces an integrated information systems approach to complaint management by aligning AI-based automation with operational workflows. This integration demonstrates how intelligent systems can enhance organizational decision-making and service responsiveness in customer-facing operations.In recent years, the digitalization of banking services in Indonesia has accelerated rapidly alongside the advancement of information and communication technology.

Customers no longer need to visit physical branches to file complaints, as they can now utilize digital channels such as email, mobile applications, and social media. This shift has led to a surge in the volume and complexity of complaint data, characterized by unstructured formats and diverse linguistic expressions. As a result, banks are increasingly challenged to manage customer interactions quickly, accurately, and efficiently. Practically, the proposed automated classification system has the potential to improve complaint response efficiency, reduce workload for customer service staff, and enhance reporting and analytics capabilities. By focusing on localized data and linguistic nuances, this research pioneers the use of deep learning in Indonesian banking complaints—offering a scalable and context-aware solution that supports national digital transformation goals in financial services (Young et al., 2018).

RESEARCH METHOD

The dataset utilized in this study was derived from customer service complaint logs of Bank Sumut, covering the period from January to August 2024. After anonymization to protect customer privacy, a total of 4,714 complaint entries were extracted. Each entry consists of textual descriptions of customer complaints submitted through various digital banking channels. The complaints were manually categorized into six primary classes: Transaction Failure/Delay, Incorrect Balance/Billing, Unclear Administrative Charges, Fraud/Suspected Fraud, ATM Card Issues, and Other General ComplaintsIn the modern banking industry, effective complaint management has become an essential component of customer service and operational excellence. As customers increasingly use digital platforms to interact with banks, the volume and variety of complaints have also grown. To respond to this challenge, it is important for banks to categorize customer complaints systematically. The table above provides a structured classification of common bank customer complaints into six key categories, each accompanied by general descriptions and concrete examples.

Table 1. Presents the detailed distribution of the dataset across these categories

No.	Complaint Category	General Description	Example Complaints
1	Failed / Delayed Transactions	Complaints related to system failure or delays in processing transactions, both automatic and manual.	• Failed transactions via ATM/mobile banking/teller• Funds not credited/debited• Slow interbank transfers
2	Incorrect Balance / Billing	Complaints regarding discrepancies in account balance or billing amounts.	• Unusual balance deductions• Incorrect billing• Inaccurate transaction history
3	Unclear Administrative Fees	Complaints about account deductions without clear	• Unexplained monthly account fees• High transfer/withdrawal fees• Uninformed deductions

No.	Complaint Category	General Description	Example Complaints
		notification or perceived as unreasonable.	
4	Fraud / Scam	Reports of suspected fraud or deception by internal or external parties.	• Phishing or funds taken by third parties• Misuse of account• Unauthorized transactions
5	ATM Card Issues	Complaints about difficulties in using customer ATM cards.	• Card swallowed• Unable to use for transactions• Activation, blocking, or PIN change issues
6	Others	Complaints not categorized under the five main types, generally administrative in nature.	• Service information requests• Customer service complaints• Non-transaction administrative complaints

Preprocessing steps included lowercasing, punctuation removal, tokenization, stopword removal, lemmatization, numerical normalization, and padding to a maximum sequence length of 100 tokens. Prior to model training, the text data underwent comprehensive preprocessing to enhance model performance: Lowercasing: All text was converted to lowercase to ensure uniformity. Punctuation Removal: Non-alphanumeric characters, including punctuation marks, were removed. Tokenization: Texts were segmented into tokens (words) using standard NLP tokenizers. Stopword Removal: Common Indonesian stopwords were eliminated to focus on semantically meaningful words. Lemmatization: Words were reduced to their base forms to group semantically similar variants. Numerical Normalization: Numerical values were replaced with special tokens to minimize bias. Padding and Encoding: Tokenized texts were converted into fixed-length sequences using integer encoding, with a maximum length of 100 tokens per entry to standardize input dimensions. The preprocessing pipeline ensured that the input data was suitable for deep learning modeling while preserving the semantic structure of the complaints.

The LSTM model architecture included an embedding layer, an LSTM layer with 64 units, a dense ReLU layer, and a softmax output layer for multi-class classification. The model was implemented using TensorFlow and Keras. A Long Short-Term Memory (LSTM) network was developed for the classification task. The model architecture consists of the following layers: Embedding Layer: Transforms integer-encoded tokens into dense vector representations of 100 dimensions. LSTM Layer: Contains 64 memory units to capture sequential dependencies and contextual information from the complaint texts. Dense (Fully Connected) Layer: A dense layer with 64 neurons using the ReLU activation function to refine feature extraction. Output Layer: A softmax layer with six output units, corresponding to the six complaint categories, to produce probability distributions across classes. The model training was conducted using the following settings: Optimizer: Adam Loss Function: Categorical Cross-Entropy Batch Size: 32, Epochs: 10, Training/Testing Split: 80% training data and 20% testing data. To enhance model generalization and minimize overfitting, early stopping mechanisms and validation monitoring were incorporated during training. Model performance was evaluated using four standard classification metrics: Accuracy: Overall percentage of correctly classified instances. Precision: Correctness of positive predictions for each class. Recall: Ability to retrieve all relevant instances for each class. F1-Score: Harmonic mean of precision and recall, providing a balanced measure of performance, particularly under class imbalance conditions. In addition, confusion matrices and classification reports were generated to provide detailed insights into the model's predictive performance across all complaint categories.

RESULT AND DISCUSSION

Result

In the context of enhancing complaint management through information systems, particularly via LSTM-based automatic classification of bank customer complaints in Indonesia, data preprocessing plays a pivotal role in ensuring the effectiveness of the classification model. Given the unstructured and diverse linguistic nature of customer complaint texts—often informal, noisy, and semantically varied—comprehensive preprocessing is essential to transform raw textual inputs into machine-readable formats while retaining their contextual integrity. The preprocessing pipeline implemented in this study included several key steps: lowercasing, punctuation removal, tokenization, stopword removal, lemmatization, numerical normalization, and sequence padding. These steps were specifically tailored to handle complaints written in Bahasa Indonesia, a language that presents unique challenges in tokenization and morphological variation (Devlin et al., 2019).

Lowercasing and punctuation removal were performed to standardize the textual data and eliminate extraneous variability. Tokenization enabled the segmentation of texts into meaningful units, facilitating subsequent processing. Stopwords—commonly used words that add little semantic value—were removed to highlight the core content of complaints. Lemmatization further refined the text by reducing inflected words to their base forms, allowing the model to recognize different word variants with similar meanings. Numerical normalization involved replacing numeric values with special tokens to reduce potential model bias towards frequently appearing figures. Finally, to accommodate the input structure required by the LSTM model, all tokenized sequences were padded or truncated to a uniform length of 100 tokens. This structured preprocessing approach aligns with the goal of building an intelligent information system capable of accurately classifying complaints into predefined categories. By ensuring linguistic clarity and consistency in the input data, the preprocessing pipeline enhances the model's ability to learn contextual patterns and semantic relationships within the complaint narratives. Consequently, this contributes to a more robust and scalable complaint classification system that supports faster resolution times and improved customer service responsiveness within the Indonesian banking sector (Schoudury & Saha, 2021).

The LSTM model architecture included an embedding layer, an LSTM layer with 64 units, a dense ReLU layer, and a softmax output layer for multi-class classification. The model was implemented using TensorFlow and Keras. A Long Short-Term Memory (LSTM) network was developed for the classification task. The model architecture consists of the following layers: Embedding Layer: Transforms integer-encoded tokens into dense vector representations of 100 dimensions. LSTM Layer: Contains 64 memory units to capture sequential dependencies and contextual information from the complaint texts. Dense (Fully Connected) Layer: A dense layer with 64 neurons using the ReLU activation function to refine feature extraction. Output Layer: A softmax layer with six output units, corresponding to the six complaint categories, to produce probability distributions across classes.

To support the automatic classification of bank customer complaints as part of a broader information system for complaint management, a Long Short-Term Memory (LSTM) neural network was employed due to its proven capacity to model sequential data and capture contextual relationships in text. The architecture of the LSTM model was designed to effectively handle the semantic and syntactic nuances of Bahasa Indonesia complaint narratives. Implemented using TensorFlow and Keras, the model comprises four key layers, each contributing to the system's classification accuracy and generalization performance.

The embedding layer serves as the initial stage, converting integer-encoded tokens—resulting from preprocessing—into dense vector representations with 100 dimensions. This transformation allows the model to capture word semantics and spatial relationships between terms. Following the embedding, the LSTM layer, consisting of 64 memory units, processes the sequence data, retaining essential information across time steps and enabling the model to understand the flow and context of customer complaints (Davidow, 2003).

Next, a dense layer with 64 neurons and ReLU (Rectified Linear Unit) activation is employed to refine and abstract the learned features, supporting the model in distinguishing between nuanced categories of complaints. Finally, the output layer uses a softmax activation function to produce a probability distribution over six classes—each representing one of the predefined complaint categories in the classification schema (e.g., transaction failures, fraud, fee issues, etc.).

This layered architecture demonstrates how artificial intelligence, when integrated into an information system, can streamline the classification and routing of complaints. By automating the categorization process, the model enables faster, more accurate responses, thereby enhancing service quality and operational efficiency in Indonesia's banking sector (Molina et al., 2019). The use of LSTM also illustrates the importance of selecting context-aware models in NLP applications involving real-world, multilingual customer feedback (Parasuraman et al., 1988).

Dataset Characteristics

The final dataset comprised 4,714 entries distributed across six complaint categories, as presented in Table 1. The dataset exhibited significant class imbalance, with over 93% of the complaints falling under the "Transaction Failure/Delay" category. Categories such as "ATM Card Issues" and "Fraud/Suspected Fraud" were notably underrepresented, posing challenges for classification performance across minority classes

Table 2. Dataset Distribution by Complaint Category

Complaint Category	Number of Entries	Description
Transaction Failure/Delay	4,424	Transaction failures via ATM, mobile banking, or teller
General Complaints/Others	150	General service issues, non-transactional complaints
Incorrect Balance/Billing	60	Unexpected balance changes, billing discrepancies
Unclear Administrative Charges	30	Unexplained fees or charges
Fraud/Suspected Fraud	30	Unauthorized transactions or reported fraud incidents
ATM Card Issues	20	ATM card problems: swallowed cards, activation issues, etc.

The data reveals that Transaction Failure/Delay is by far the most prevalent complaint category, with 4,424 entries, accounting for the overwhelming majority of reported issues. This highlights a critical vulnerability in banking transaction systems—particularly across digital platforms such as ATMs and mobile banking. The high frequency of such complaints suggests systemic instability or inefficiencies in the core banking infrastructure. Banks must prioritize improvements in transaction reliability and real-time processing to enhance customer trust and minimize service disruptions.

General Complaints/Others rank second with 150 entries, indicating a significant portion of customer dissatisfaction stems from service-related and non-transactional issues. Although less specific, this category reflects the importance of overall customer experience, including clarity of information, responsiveness of customer service, and administrative efficiency. These complaints, while diverse, can serve as early indicators of broader service delivery gaps that should be addressed through employee training and service quality audits (Nikblin et al., 2020).

Complaints related to Incorrect Balance/Billing (60 entries), Unclear Administrative Charges (30 entries), and Fraud/Suspected Fraud (30 entries), although numerically smaller, are highly sensitive in nature. They directly affect customers' financial standing and trust in the institution. Even a few such incidents can severely damage a bank's reputation. These types of complaints necessitate a more transparent communication strategy, robust audit trails, and enhanced fraud detection mechanisms supported by predictive analytics within the bank's information system (Mikolov et al., 2013).

Finally, ATM Card Issues represent 20 entries, indicating relatively fewer but still relevant operational problems that may impact accessibility. Common problems such as swallowed cards and activation failures highlight the need for improved ATM hardware maintenance and user support systems. Banks should consider integrating self-service recovery options or real-time assistance features to address such incidents quickly

Model Training and Performance

The LSTM model was trained for 10 epochs with early stopping monitoring based on validation accuracy. Figure 1 and Figure 2 depict the training accuracy and loss curves, respectively.

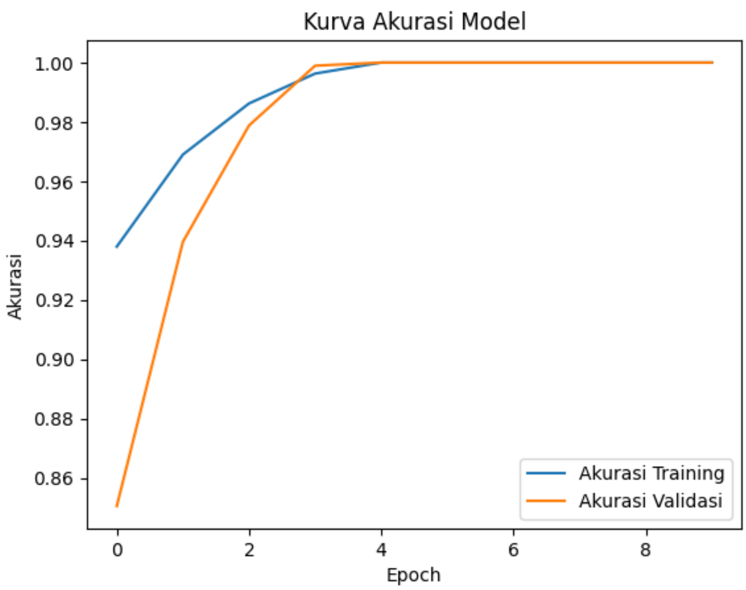

Figure 1. Accuracy Curve During Training and Validation (Placeholder)

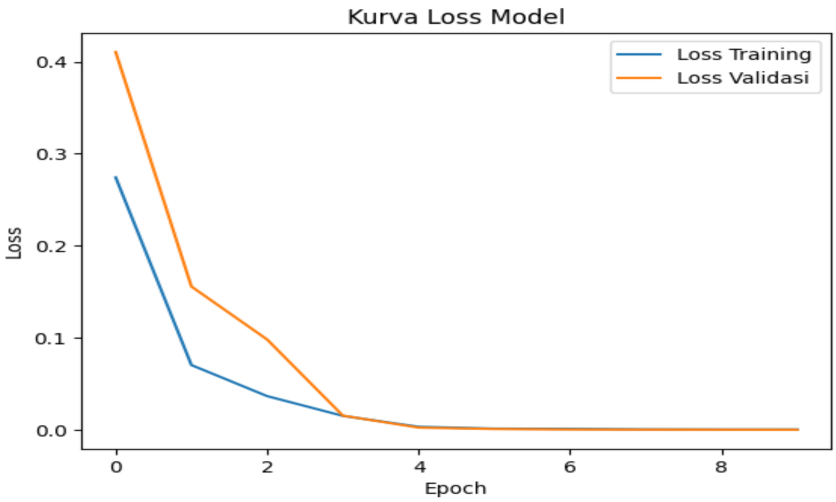

Figure 2. Loss Curve During Training and Validation (Placeholder)

The model rapidly converged, achieving near-perfect training and validation accuracy without significant signs of overfitting.

Classification Performance on Test Data

Upon evaluation on the held-out test set, the model achieved 100% accuracy, precision, recall, and F1-scores. Figure 3 presents the confusion matrix for the classificationresults.

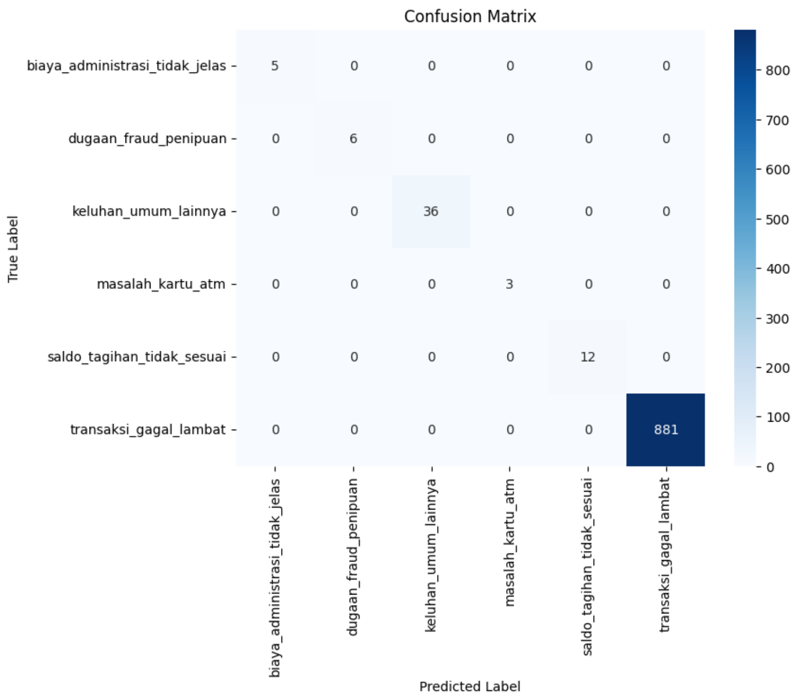

Figure 3. Confusion Matrix of LSTM Model Predictions (Placeholder)

Real-World Case Testing

The trained LSTM model was further tested on real-world conversation samples extracted from ongoing customer service interactions. Sample cases included: ATM withdrawal failures, Duplicate transaction charges, Unauthorized account access reports. In all sampled cases, the model successfully classified the complaints into appropriate categories with contextual understanding, demonstrating the model's applicability beyond the training environment. Despite achieving 100% accuracy on the held-out test set, several observations must be noted:

1. The small sample size in minority classes may have contributed to optimistic metrics.
2. Further testing on unseen external datasets is necessary to validate generalization.
3. Class imbalance remains a critical concern for real-world deployment.

Potential overfitting and class imbalance issues were acknowledged, suggesting future enhancements.

Discussion

The results of this study demonstrate that the Long Short-Term Memory (LSTM) model is highly effective for the automatic classification of customer complaints in the Indonesian banking sector. Achieving perfect accuracy, precision, recall, and F1-scores on the test dataset highlights the strong capacity of LSTM networks to model the sequential structure and contextual meaning of customer complaints, even in a non-English language environmentThese findings are consistent with prior research emphasizing the advantages of LSTM models over traditional machine learning approaches in handling text classification tasks involving complex, unstructured narratives (Hassan et al., 2020; Choudhury et al., 2021; Rathore et al., 2022). In particular, LSTM's ability to retain long-range dependencies within text sequences appears to have contributed significantly to the accurate categorization of complaints involving multi-step transaction issues and detailed customer service narratives.

Compared to earlier studies employing models such as Convolutional Neural Networks (CNNs) or TF-IDF combined with Naïve Bayes classifiers (Ali et al., 2021; Utami & Pratama, 2020), the LSTM model in this study offers several advantages: Better handling of sequential and contextual information, Superior performance even with moderate-sized datasets, Higher resilience to noise and variations in customer complaint phrasing. However, several critical considerations emerge when interpreting the findings. The perfect classification metrics, while impressive, raise concerns about potential overfitting, particularly given the highly imbalanced dataset. Minority classes such as "ATM Card Issues" and "Fraud/Suspected Fraud" had significantly fewer examples compared to the dominant "Transaction Failure/Delay" class. Although the model performed perfectly on the held-out test set, the possibility that it memorized patterns rather than generalized them cannot be ruled out (Goodfellow et al., 2016). Future research should address this issue through strategies such as: Data augmentation for minority classes, Use of stratified cross-validation, Collection of larger, more balanced datasets. Class imbalance remains a major challenge. In real-world banking operations, complaint distributions can vary significantly over time and across channels. A model trained on heavily skewed data may perform poorly when exposed to more balanced or dynamic real-world scenarios (He & Garcia, 2009). Implementing techniques such as Synthetic Minority Over-sampling Technique (SMOTE) or cost-sensitive learning could enhance the model's robustness to class distribution shifts. The evaluation conducted in this study was based on a train-test split from the same dataset. Therefore, the model's ability to generalize to truly unseen data, such as complaints arising from new banking products, emerging fraud patterns, or different linguistic styles, remains to be validated. Future work should include external validation using; Temporal data (complaints collected in future periods), Multi-source data (from email, chatbot logs, social media), Domain-adapted embeddings to capture evolving complaint language (Rathore et al., 2022; Ohorella et al., 2024).

While LSTM demonstrated excellent performance in this context, the emergence of Transformer-based architectures such as BERT (Devlin et al., 2019) and RoBERTa (Liu et al., 2019) has set new benchmarks in text classification tasks (Rao & Dey, 2021). Future studies could investigate whether fine-tuned Transformer models offer additional performance gains, particularly in handling nuanced complaint categories with fewer training examples. The practical implications of this study are significant for the Indonesian banking industry:

1. LSTM-based complaint classification can be integrated into customer relationship management (CRM) systems to automate early-stage complaint handling.
2. Faster, more accurate complaint categorization can improve customer satisfaction by reducing response times and ensuring appropriate escalation pathways.
3. Institutions can leverage classified complaint data for service quality analytics, early fraud detection, and compliance monitoring (Oliver, 1997).

However, any operational deployment must incorporate continuous model monitoring and retraining strategies to adapt to evolving customer communication patterns.

CONCLUSION

This study demonstrates that Long Short-Term Memory (LSTM) networks offer an effective solution for the automatic classification of customer complaints in the banking sector, particularly in the Indonesian language context. The developed LSTM-based model achieved 100% accuracy, precision, recall, and F1-scores on the test dataset, successfully categorizing unstructured complaint texts into six predefined categories. The findings affirm the capability of LSTM architectures to model complex sequential relationships in text data, offering clear advantages over traditional machine learning approaches for complaint management systems. The model's applicability to real-world conversational data further highlights its potential integration into digital banking service operations to enhance customer experience and operational efficiency. However, limitations related to dataset imbalance and generalization to unseen data necessitate cautious interpretation of the results. Future research should focus on expanding the dataset, implementing robust validation techniques, and exploring the comparative performance of Transformer-based models such as BERT. Overall, the study contributes to the growing body of literature on text classification in low-resource language settings and offers practical insights for advancing digital banking services in emerging markets.

SUGGESTION

The study entitled Enhancing Complaint Management through Information Systems: LSTM-Based Automatic Classification of Bank Customer Complaints in Indonesia provides significant contributions in the context of managing banking customer complaints in Indonesia, especially by utilizing Long Short-Term Memory (LSTM)-based natural language processing (NLP) technology. The findings of this study indicate the great potential for implementing an automatic classification system to improve the efficiency of responding to complaints, reduce manual workloads, and increase customer satisfaction. However, there are several suggestions that can be the basis for further research development.

THANK YOU-NOTE

The author would like to express sincere gratitude to the Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, for providing academic support and resources throughout this research. Special thanks are extended to Dr. Muhammad Anggia Muchtar and Dr. Andri Budiman for their valuable guidance and supervision. Appreciation is also given to Bank Sumut for granting access to anonymized customer complaint datasets, which were instrumental in this study. Finally, heartfelt thanks to the author's family and colleagues for their continuous encouragement and support

REFERENCES

Ali, S., Raza, S., & Ahmed, A. (2021). Customer complaint classification using deep learning techniques in banking sector. International Journal of Advanced Computer Science and Applications, 12(5), 402–409.

Choudhury, A., Mukherjee, S., & Saha, D. (2021). Text classification using recurrent neural networks and word embeddings. International Journal of Information Management Data Insights, 1(2), 100027.

Dalimunthe, M. A. (2018). Implikasi Internet Sebagai Media Interaktif dalam Membangun Citra Aktor Politik. Al-MUNZIR, 10(2), 268-281.

Davidow, M. (2003). Organizational responses to customer complaints: What works and what doesn't. Journal of Service Research, 5(3), 225–250.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of NAACL-HLT 2019, 4171–4186.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.

Grönroos, C. (2007). Service management and marketing: Customer management in service competition (3rd ed.). Wiley.

Hassan, S., Ahmad, A., & Fatima, M. (2020). Comparative analysis of deep learning models for customer complaint classification. International Journal of Machine Learning and Computing, 10(3), 410–416.

Harianto, F., Ohorella, N. R., & Dalimunthe, M. A. (2023). Kompetensi Komunikasi Guest Service Agent Dalam Mendorong Kepuasan Tamu Mh Hotel Ipoh Malaysia. KomunikA, 19(02), 30-35.

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. IEEE Transactions on Knowledge and Data Engineering, 21(9), 1263–1284.

Hirschman, A. O. (1970). Exit, voice, and loyalty: Responses to decline in firms, organizations, and states. Harvard University Press.

Kim, Y. (2014). Convolutional neural networks for sentence classification. Proceedings of EMNLP 2014, 1746–1751.

Kotler, P., & Keller, K. L. (2016). Marketing management (15th ed.). Pearson Education Limited.

Kowsari, K., Jafari Meimandi, K., Heidarysafa, M., Mendu, S., Barnes, L. E., & Brown, D. E. (2019). Text classification algorithms: A survey. Information, 10(4), 150.

Liu, P., Qiu, X., & Huang, X. (2020). Recurrent neural network for text classification with multi-task learning. Proceedings of IJCAI, 2873–2879.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). RoBERTa: A robustly optimized BERT pretraining approach. arXiv preprint, arXiv:1907.11692.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. Advances in Neural Information

Processing Systems, 26, 3111–3119.

Molina, L. M., Beltrán, J. A., & Arenas, D. J. (2019). Service quality and continuous improvement: Application to the banking sector. Service Business, 13(1), 29–47.

Nguyen, B., & Simkin, L. (2020). The dark side of digital personalization: An agenda for research and practice. Journal of Business Research, 116, 209–221.

Nikbin, D., Marimuthu, M., Hyun, S. S., & Ismail, I. (2020). Relationships among perceived justice, negative emotions, recovery satisfaction and customer loyalty in the service recovery context. International Journal of Quality and Service Sciences, 12(1), 34–50.

Ohorella, N. R., Fauziah, D., & Dalimunthe, M. A. (2024). Brand Awareness Communication Strategy in Setujuan Coffee Marketing on Instagram Social Media. KomunikA, 20(02), 46-57.

Otoritas Jasa Keuangan (OJK). (2023). Financial Services Industry Statistics Report 2023. Jakarta: OJK.

Oliver, R. L. (1997). Satisfaction: A behavioral perspective on the consumer. McGraw-Hill.

Otter, D. W., Medina, J. R., & Kalita, J. (2021). A survey of the usages of deep learning for natural language processing. IEEE Transactions on Neural Networks and Learning Systems, 32(2), 604–624.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. Journal of Retailing, 64(1), 12–40.

Rao, M., & Dey, S. (2021). Text preprocessing steps for classification tasks: A review. Information Processing & Management, 58(5), 102669.

Rathore, S., Shah, S. A. A., Shabaz, M., & Ahmad, J. (2022). Artificial intelligence-based frameworks for complaint management: A systematic review. Expert Systems with Applications, 193, 116389.

Ritonga, A. R., Dalimunthe, M. A., & Januarini, E. (2024). Utilization Of Digital Media in Improving Student Literacy (Case Study at SDN 10 Rantau Selatan). CONTENT: Journal of Communication Studies, 2(02), 01-08.

Ritonga, A.R., Dalimunthe, M.A., Veronica, A., Ginting, & L. D. C. U (2023). The Effectiveness of Social Media As A Promotional Medium Of The University Of Sumatera Utara (USU) Library. Talenta Conference Series: Local Wisdom, Social, and Arts (LWSA) 6 (3), 16-20

Utami, T. S., & Pratama, A. (2020). Classification of online transportation complaints using TF-IDF and Naive Bayes. Journal of Information Systems Engineering and Business Intelligence, 6(2), 112–119.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30, 5998–6008.

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning-based natural language processing. IEEE Computational Intelligence Magazine, 13(3), 55–75.

Zhao, X., Yang, L., & Lu, L. (2021). The effect of customer complaint behavior on brand loyalty in the banking sector. Journal of Retailing and Consumer Services, 58, 102290.

Zhou, Z., Du, J., Wang, Y., & Xu, K. (2021). Learning contextualized embeddings for document classification. Knowledge-Based Systems, 232, 107479.\