BIM Proceedings

Bali International Meeting on Business, Advanced engineering, and Leadership sustainability Impact
BIM University
DOI: 10.xxxxxxxxxxxxxxxxxxxx/

Sentiment and Topic Modeling of Livin Reviews for Consumer Trust Insights

Andika Haidar Ihsan[1], Diniffa Azzahra[2], and Muhammad Imam Baihaqi[3]

[1]Islamic University of Indonesia, St. Kaliurang, Umbulmartani, Ngemplak, Special Region of Yogyakarta, Indonesia, 55584
[2]Islamic University of Indonesia, St. Kaliurang, Umbulmartani, Ngemplak, Special Region of Yogyakarta, Indonesia, 55584
[3]Islamic University of Indonesia, St. Kaliurang, Umbulmartani, Ngemplak, Special Region of Yogyakarta, Indonesia, 55584

e-mail: 24522053@students.uii.ac.id

Abstract

The rapid growth of mobile banking in Indonesia has driven the adoption of financial super-apps like Livin' by Mandiri, where consumer trust serves as a critical prerequisite for engagement. While traditional survey-based methods dominate existing literature, they often suffer from limited sample sizes and inherent biases. To address this gap, this study integrates Natural Language Processing (NLP) techniques, specifically Support Vector Machine (SVM) and Latent Dirichlet Allocation (LDA), to analyze user reviews and characterize consumer trust perceptions. A dataset of 3,000 user reviews was scraped from the Google Play Store and underwent a comprehensive text pre-processing pipeline. Sentiment classification using an SVM model with a TF-IDF approach revealed a highly polarized distribution, consisting of 59.3% negative, 40.1% positive, and 0.6% neutral reviews. While the SVM model achieved robust classification for the positive and negative classes, it struggled with the severe class imbalance of the neutral group. Subsequent LDA topic modeling isolated each sentiment category to uncover the underlying thematic structures. The results indicate that ease of use (*mudah*) and functional helpfulness (*membantu*) act as primary drivers for positive trust formation. Conversely, negative sentiment is heavily clustered around system reliability and login failures (*sulit*, *login*). These findings imply that resolving authentication stability should be prioritized to mitigate trust erosion, providing actionable intelligence for digital financial service developers

***Keywords**: Sentimen Analysis, SVM, LDA, Bank*

INTRODUCTION

The fast development of digital technology has changed the nature of financial services, especially with the adoption of mobile banking, which has enabled real-time transactions and financial management via smartphones [1]. In Indonesia, this growth has been fueled by the COVID-19 pandemic that massively drove smartphone and internet penetration [2]. To counter this trend, PT Bank Mandiri (Persero) Tbk introduced Livin' by Mandiri, a super-app that combines banking, investment, insurance, and lifestyle services into one platform, with tens of millions of registered users in 2023 [3].

User reviews on the Google Play Store and Apple App Store are a true reflection of consumer

DOI: 10.xxxxxxxxxxxxxxxxxxxx/

feelings. They come from real opinions and aren't tampered with, making them incredibly valuable. These online reviews play a crucial role in influencing new users' decisions to adopt apps and in keeping existing users loyal [4]. In the context of mobile banking, consumer trust is the key prerequisite for engaging in financial transactions. Trust is a multidimensional construct that includes perceptions of the security, reliability, competence, and benevolence of the service provider [5]. Empirical work based on the Technology Acceptance Model (TAM) shows that trust is one of the strongest predictors for mobile banking adoption and retention [6].

Natural Language Processing (NLP) techniques, especially sentiment analysis, have been widely used to analyze large scale unstructured text data. Among various algorithms, Support Vector Machine (SVM) has been proven to be superior for sentiment classification in high dimensional feature space with limited labeled data and can achieve accuracy more than 85% on mobile application reviews [7][8]. But sentiment analysis only shows the polarity of opinions, not the thematic dimensions underlying it. Latent Dirichlet Allocation (LDA) solves this problem by automatically uncovering latent topic structure in document corpora, i.e., the recurring dominant themes without manual labeling [9][10]. Combining both approaches provides more comprehensive insights, not only identifying what users feel, but also identifying which specific service aspects drive those feelings [11].

There are still many research gaps, even though the literature on mobile banking is growing. Studies in Indonesia are still overwhelmingly dominated by survey-based studies, which are subject to bias and limited sample size [12]. Furthermore, the scientific literature rarely discusses NLP-based analysis of Indonesian super-apps such as Livin' by Mandiri that combine banking and non-banking services. The present study aims to fill these gaps by integrating SVM and LDA on a corpus of reviews on Livin' by Mandiri for a comprehensive characterization of consumer trust perceptions. In particular, this study aims to: (1) develop an SVM classifier for classifying positive, negative and neutral sentiment; (2) identify latent thematic structures through LDA that correspond to known antecedents of consumer trust in digital banking; and (3) synthesize the outputs of both analytical frameworks to yield practical implications for platform developers and researchers in digital financial services.

METHOD

Data Collection

Web scraping is an automated method for extracting large quantities of data from online platforms in a structured and efficient manner [13]. The data was collected by scraping on the Google Play Store with the target of Livin' by Mandiri application in this study. The reviews were sorted by Newest to display the most recent user experiences, with a total sample of 3,000 reviews collected.

Data pre-processing

In natural language processing, text pre-processing is an important step that transforms the raw, unstructured text into a clean and consistent format that can be processed computationally [14]. Before analysis, a pre-processing pipeline was performed to improve the quality of the data in the following stages:

1. Sentiment Labeling
2. Case Folding
3. Cleansing of URLs, Mentions, Hashtags, and Numbers
4. Removal of Non-ASCII Characters (Emojis) & Punctuation
5. Whitespace Normalization
6. Slang Normalization (Non-standard Words)

DOI: 10.xxxxxxxxxxxxxxxxxxxx/

7. Stopword Removal

The first step was the labeling of reviews based on star ratings [15]: 1–2 were negative, 3 neutral and 4–5 positive. The second stage, case folding, made all uppercase letters lowercase, for example, "Beberapa" became "beberapa". The third stage removed elements without sentiment value. SVM algorithm depends only on emotionally nuanced words . Links to websites ("http…"), user tags ("@"), hashtags ("#") and numbers ("123…") were removed. In stage 4, all emoji characters (like 👍 and 😂) and other unicode symbols were removed, then all punctuation marks (periods, commas, exclamation marks) were removed because the TF-IDF algorithm can't process emoji characters. The fifth stage normalized whitespace, removing any double spaces left over from the deletions in the previous stages. The sixth stage was to normalize slang and informal abbreviations like "bgt," "ga," and "apk" to their standard forms, "banget," "tidak," and "aplikasi," so that the weights of words were not fragmented [15]. The last step was stopword removal to remove common words that have no sentiment value (e.g. 'yang', 'di', 'ke') and application specific terms (e.g. 'livin', 'mandiri').

SVM Sentiment Classification

The Support Vector Machine (SVM) is a supervised machine learning algorithm that finds an optimal hyperplane to separate classes of data in a high-dimensional feature space, and has been shown to perform strongly and consistently in text classification tasks [16]. In the present study, we pulled out features using a TF-IDF (Term Frequency-Inverse Document Frequency) approach first, then fed them into SVM , which basically weights every term by its occurrence inside a single document relative to how often that same term shows up across all documents . This sequence is meant to capture the discriminative importance of the words, kind of like separating the signal from the background, not just counting stuff blindly. To evaluate the generalization ability of the model, the dataset was split by simple random sampling, where 80% (2,382 rows) was used as training data to build the classification model and the rest 20% (596 rows) was used as testing data to validate the accuracy and prediction performance of the model on unseen data.

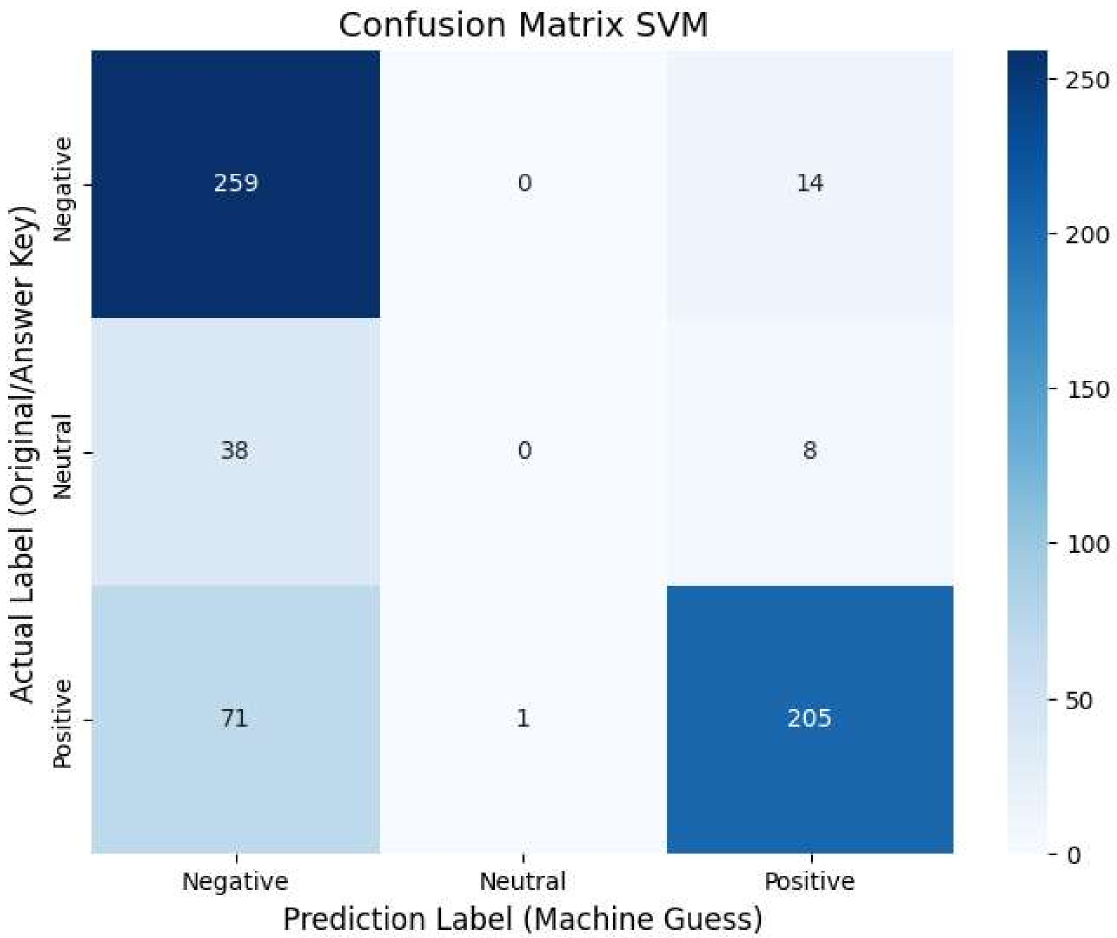
LDA Topic Modeling

Latent Dirichlet Allocation (LDA) is a generative probabilistic model that assumes each document is a mixture of latent topics and each topic is characterized by a probability distribution over words, thus allowing the unsupervised detection of hidden thematic structures in large text corpora [9]. The cleaned review texts obtained through pre-processing in this study were first transformed into a Bag-of-Words matrix, which represents the frequency of each term in documents [16]. In order to obtain more focused insights, the LDA implementation was separated by the three sentiment classification categories reviews tagged 'Positive,' 'Neutral' and 'Negative.' For each category, the LDA model was run with the number of topics set to $k = 1$. This sentiment isolation strategy was used to derive the most probable keyword distributions that represent the single most dominant topic within each sentiment group as a basis for interpreting the specific context behind user opinions.

RESULTS AND DISCUSSION

This section presents the findings we got from using SVM based sentiment analysis and LDA topic modeling on the Livin by Mandiri review corpus. The results are set up in order like, first we check the SVM classification model performance, then we look at sentiment distribution, and in the end we show the LDA topic modeling output across three sentiment categories. Each subsection kind of talks through what we found in relation to the research objectives, plus what it means for consumer trust in digital banking services, whether it shows more confidence or hesitation over time:

DOI: 10.xxxxxxxxxxxxxxxxxxxx/

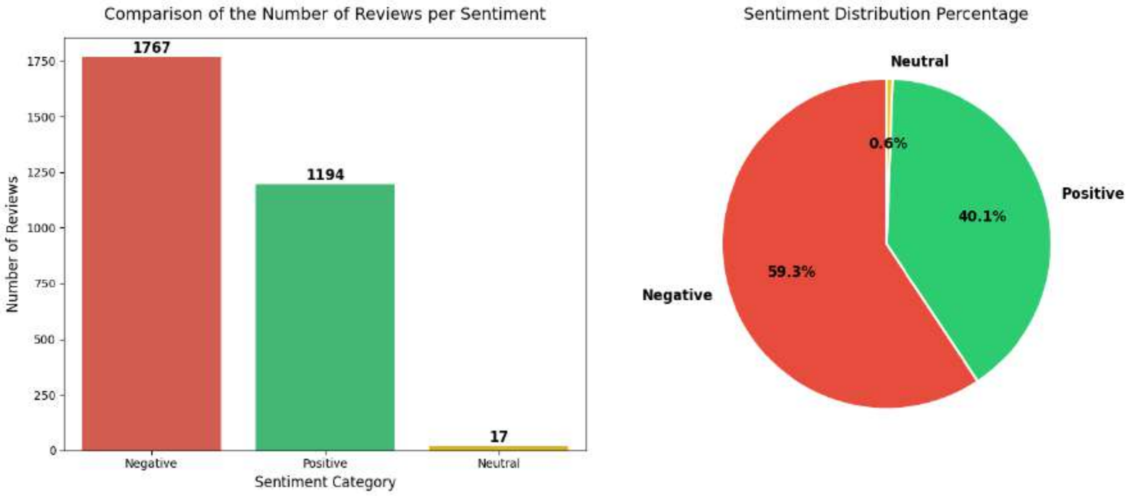


The performance of the SVM classification model was kind of evaluated using a confusion matrix on the testing dataset, which had 596 reviews total. As it's shown in Figure 1, the model managed to classify correctly 259 of 273 negative reviews, and 205 of 277 positive ones, however then it really struggled with the neutral class, where all 46 of the neutral test samples ended up misclassified. More precisely, 38 got predicted as negative , and 8 got predicted as positive , and in the end none of those 46 were correctly spotted. This kind of result, is most likely due to the strong class imbalance during training, since there were only 17 neutral reviews across the whole dataset, so the model didn't get enough examples to actually learn the subtle distinguishing cues of neutral sentiment, or at least not well enough.

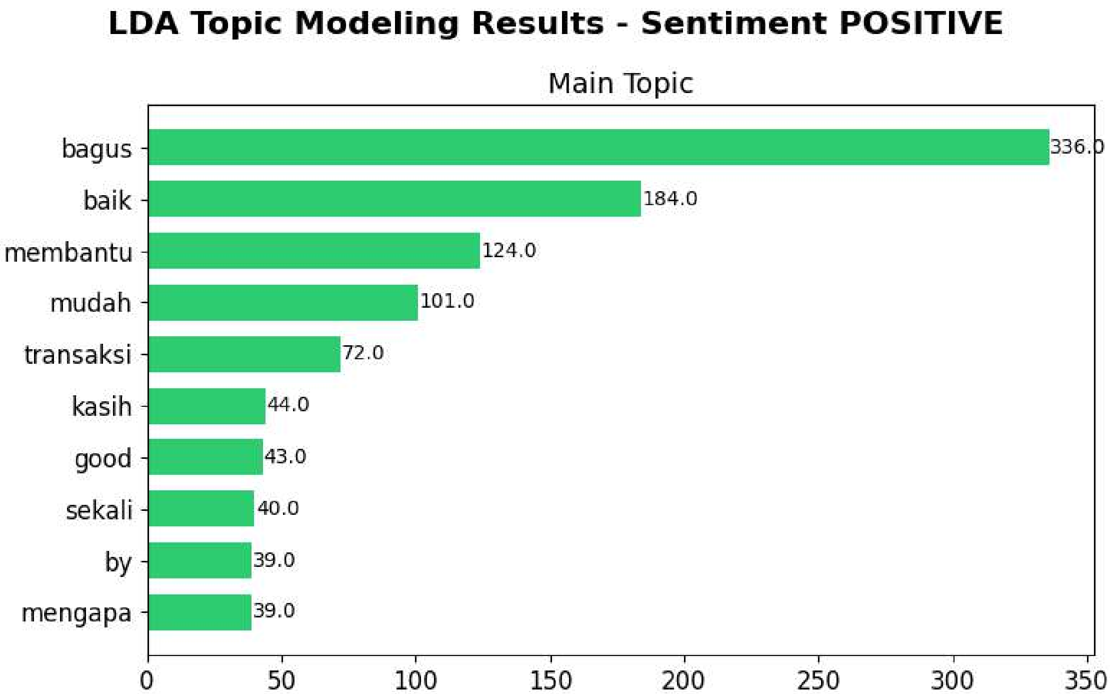
Overall, the SVM model showed a fairly ok classification behavior for the negative and positive groups, which kinda lines up with earlier reports on how SVM is sturdy in high dimensional text classification chores. That said the total inability to sort the neutral mood out, kind of flags a well known restriction of SVM in unbalanced multi class contexts, where the less represented kind of sentiment gets swallowed by the bigger adjacent classes when the hyperplane is being tuned. Looking ahead, future efforts could use an over sampling route like SMOTE, or also tweak the class weights so the minority class recall becomes more reliable.

Sentimen Distribution

DOI: 10.xxxxxxxxxxxxxxxxxxxx/



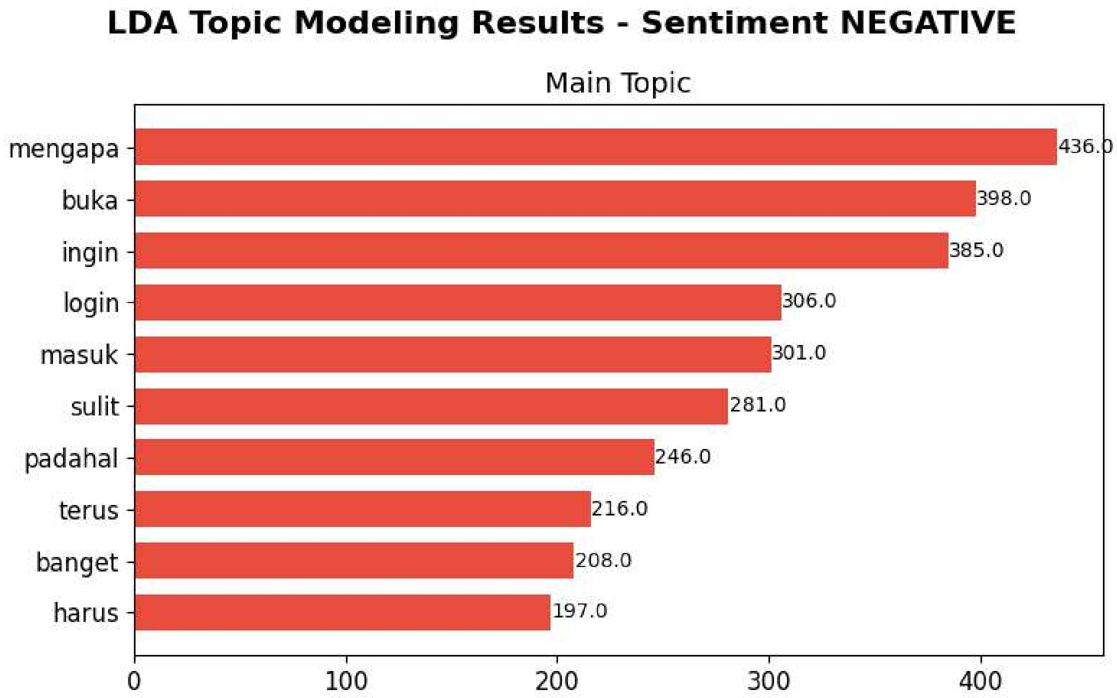
The sentiment labeling outcomes, derived from user star ratings across the 3,000 gathered reviews, showed this pretty uneven spread across the three sentiment categories. Like in Figure 2, the negative reviews were basically the lead group, with 1,767 reviews, 59.3%, then positive reviews came next with 1,194 entries, 40.1%. Meanwhile the neutral reviews were just 17 reviews, only 0.6% . So, this lopsided pattern makes it look like Livin' by Mandiri reviewers on the Google Play Store are mostly stepping in after a negative experience, which matches earlier research on mobile app reviews too where annoyed users tend to show up more often on review platforms than happy ones. Also, the very small share of neutral sentiment hints that people's views on the app are sort of polarized, like opinions swing strongly one way or the other, instead of staying balanced.



The LDA topic modeling applied to the 1,194 positive reviews identified a dominant topic, and it looks like it's mostly about keywords such as bagus (good/great, 336) and baik (good/well, 184) then membantu (helpful, 124) , mudah (easy, 101) and transaksi (transaction, 72) as shown in Figure 3. This grouping of words kind of shows user satisfaction that sticks around three main aspects: the overall application quality, how simple they feel it is to use, and the

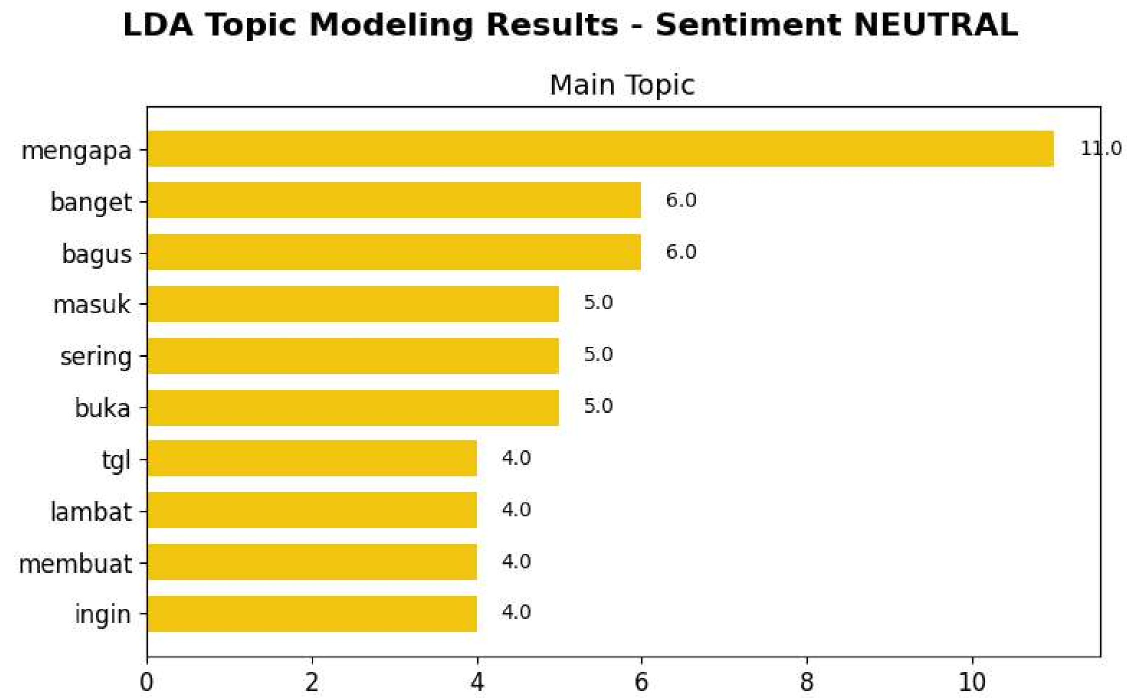
DOI: 10.xxxxxxxxxxxxxxxxxxxx/

convenience of doing transactions. The fact that membantu is so frequent together with mudah, implies users really like the application's practical usefulness and the straightforward navigation too, these are like key triggers of consumer trust according to the Technology Acceptance Model. Also, transaksi being present a lot hints that when the transactions run smoothly and reliably, it becomes a real driver of a better positive experience, so the platform's core banking function feels more trustworthy.



The LDA analysis of the 1,767 negative reviews produced a dominant topic dominated by terms such as *mengapa* (why, 436), *buka* (open, 398), *ingin* (want, 385), *login* (306), *masuk* (enter/log in, 301), *sulit* (difficult, 281), and *padahal* (even though, 246), as displayed in Figure 4. This topic structure kinda reveals that the main source of user dissatisfaction is clustered around login and access failures, users repeatedly say they feel frustrated because they can't open the app or they can't successfully authenticate. The co-occurrence of mengapa sulit and padahal suggests an emotionally charged complaint pattern, like users show bewilderment plus frustration, even though they still wanna use the application (ingin) over and over. These findings directly point to system reliability and security accessibility as the most critical dimensions eroding consumer trust in Livin' by Mandiri, which matches theoretical frameworks that basically say perceived security and system dependability are core foundations for digital banking trust.

DOI: 10.xxxxxxxxxxxxxxxxxxxx/



The LDA modeling of the 17 neutral reviews, as shown in Figure 5, identified keywords including *mengapa* (11), *banget* (6), *bagus* (6), *masuk* (5), *sering* (often, 5), *buka* (5), *lambat* (slow, 4), and *tgl* (date, 4). With the really tiny size of this segment, the topic structure really should be read with care or sort of cautious. Still, the together appearance of those good words (bagus) and the negative signals (sering, lambat, masuk) in the same topic seems aligned with how ambivalent neutral sentiment users are. They usually recognize the aplikasi's overall quality in a calm way , but at the same time they mention repeat issues like loading that is slow, or access that comes and goes, sometimes even intermittent. The word *tgl* (date) may suggest that some users associate their complaints with specific recurring dates, possibly related to periodic system maintenance or high-traffic transaction periods.

Synthesizing the outputs of both the SVM classification and LDA topic modeling, three key dimensions of consumer trust emerge as central to user experience with Livin' by Mandiri. First, *reliability and system accessibility* constitutes the most critical trust determinant, as evidenced by the dominance of login and access-related complaints in negative sentiment. Second, *ease of use and functional utility* emerges as the primary driver of positive trust formation, reflected in terms such as *mudah* and *membantu* among satisfied users. Third, *transactional confidence* users' belief that financial transactions will be executed correctly and securely appears as a supporting pillar of positive sentiment. These findings suggest that PT Bank Mandiri should prioritize resolving authentication and login stability issues as the most immediate lever for improving consumer trust, while continuing to invest in interface simplicity and transaction reliability to sustain positive user engagement

CONCLUSION

DOI: 10.xxxxxxxxxxxxxxxxxxxx/

(The conclusion summarizes the main findings of the study and emphasizes the key contributions of the research. This section should demonstrate how the research objectives have been achieved and should provide a concise synthesis of the most significant insights derived from the study.

Authors may also present recommendations for future research, policy implications, or practical applications, particularly in areas where further investigation is needed. The conclusion should not merely repeat previous sections but should provide a reflective synthesis of the research outcomes and their broader implications.)

DECLARATION OF GENERATIVE AI

During the preparation of this manuscript, the author used generative AI for language refinement, grammar correction, and structural improvement of the content. After using these tools, the author reviewed, edited, and carefully validated the content to ensure its accuracy and academic integrity. The author is fully responsible for the final content of this publication.

ACKNOWLEDGEMENTS (if any)

The authors would like to express their sincere appreciation to the Department of Industrial Engineering, Faculty of Industrial Technology, Universitas Islam Indonesia (UII), for providing the academic facilities, supportive learning environment, and resources that greatly contributed to the completion of this study.

REFERENCES

[1] A. Yasin, R. Fatima, A. N. Ghazi, and Z. Wei, “Python data odyssey: Mining user feedback from google play store,” *Data Brief*, vol. 54, p. 110499, Jun. 2024, doi: 10.1016/j.dib.2024.110499.

[2] OJK, “OJK Annual Report 2022,” 2022.

[3] Bank Mandiri, “ANNUAL REPORT BMRI 2023 (FINAL) - 22 Mei 2024,” 2024.

[4] P. H. C. Samanmali and R. A. H. M. Rupasingha, “Sentiment analysis on google play store app users’ reviews based on deep learning approach,” *Multimed. Tools Appl.*, vol. 83, no. 36, pp. 84425–84453, Apr. 2024, doi: 10.1007/s11042-024-19185-w.

[5] R. C. Mayer, J. H. Davis, and F. D. Schoorman, “An Integrative Model of Organizational Trust,” *The Academy of Management Review*, vol. 20, no. 3, p. 709, Jul. 1995, doi: 10.2307/258792.

[6] X. Luo, H. Li, J. Zhang, and J. P. Shim, “Examining multi-dimensional trust and multi-faceted risk in initial acceptance of emerging technologies: An empirical study of mobile banking services,” *Decis. Support Syst.*, vol. 49, no. 2, pp. 222–234, May 2010, doi: 10.1016/j.dss.2010.02.008.

[7] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.

[8] M. D. Bimantara and I. Zufria, “Text Mining Sentiment Analysis on Mobile Banking Application Reviews using TF-IDF Method with Natural Language Processing Approach,” *JINAV: Journal of Information and Visualization*, vol. 5, no. 1, pp. 115–123, Jul. 2024, doi: 10.35877/454RI.jinav2772.

[9] D. M. Blei, A. Y. Ng, and J. B. Edu, “Latent Dirichlet Allocation Michael I. Jordan,” 2003.

[10] T. L. Griffiths and M. Steyvers, “Finding scientific topics,” *Proceedings of the National Academy of Sciences*, vol. 101, no. suppl_1, pp. 5228–5235, Apr. 2004, doi: 10.1073/pnas.0307752101.

[11] L. Çallı, “Exploring mobile banking adoption and service quality features through user-

DOI: 10.xxxxxxxxxxxxxxxxxxxx/

generated content: the application of a topic modeling approach to Google Play Store reviews," *International Journal of Bank Marketing*, vol. 41, no. 2, pp. 428–454, Mar. 2023, doi: 10.1108/IJBM-08-2022-0351.

[12] H. Adiningtyas and A. S. Auliani, "Sentiment analysis for mobile banking service quality measurement," *Procedia Comput. Sci.*, vol. 234, pp. 40–50, 2024, doi: 10.1016/j.procs.2024.02.150.

[13] E. wina and T. Mauritsius, "Data-Driven Insights for Mobile Banking App Improvement: A Sentiment Analysis and Topic Modelling Approach for SimobiPlus User Reviews," *International Journal of Engineering Trends and Technology*, vol. 72, no. 6, pp. 347–360, Jun. 2024, doi: 10.14445/22315381/IJETT-V72I6P132.

[14] M. A. Palomino and F. Aider, "Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis," *Applied Sciences*, vol. 12, no. 17, p. 8765, Aug. 2022, doi: 10.3390/app12178765.

[15] J. R. Jim, M. A. R. Talukder, P. Malakar, M. M. Kabir, K. Nur, and M. F. Mridha, "Recent advancements and challenges of NLP-based sentiment analysis: A state-of-the-art review," *Natural Language Processing Journal*, vol. 6, p. 100059, Mar. 2024, doi: 10.1016/j.nlp.2024.100059.

[16] Y. Mao, Q. Liu, and Y. Zhang, "Sentiment analysis methods, applications, and challenges: A systematic literature review," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 4, p. 102048, Apr. 2024, doi: 10.1016/j.jksuci.2024.102048.