Vol.4, No.10, Page 2619- 2634, 2026 https://doi.org/10.59628/jast.v4i10.3463 ISSN 2958-9568 JAST

SUN PUBLISHER

مجلة جامعة صنعاء للعلوم التطبيقية والتكنولوجيا

Sana'a University Journal of Applied Sciences and Technology

https://journals.su.edu.ye/index.php/jast/

A Scoping Review of Sentiment Analysis for Customer Satisfaction in Digital Banking: Mapping the Research Gap Toward Multimodal Integration

Maimona Al-Maamari * and Ghaleb Al-Gaphari

Department of Computer Science, Faculty of Computer and Information Technology, Sana'a University, Sana'a, Yemen.

*Corresponding author: *mimwna.almamri@su.edu.ye*

Abstract

Sentiment analysis has become increasingly important for interpreting customer-generated content across social media and digital technologies. However, research on digital banking continues to rely predominantly on text-only data, with limited attention to multimodal integration. This scoping review mapped the literature on sentiment analysis for customer satisfaction in digital banking, identified research gaps, and outlined future directions. The review followed the PRISMA-ScR guidelines and searched six databases—ScienceDirect, IEEE Xplore, Wiley Online Library, ACM Digital Library, SpringerLink, and Google Scholar—for English-language studies published from 2015 to 2026; the final search was conducted on May 10, 2026. Of 1,631 records identified, 35 studies met the inclusion criteria. Text-only data were used in 31 studies (88.6%), multi-source text was used in one study (2.9%), text combined with survey data in two studies (5.7%), and numerical time-series data in one study (2.9%); no included study integrated text, audio, and video within a single sentiment analysis framework. The literature was geographically concentrated and relied mainly on traditional machine-learning and hybrid approaches, whereas transformer/LLM-based methods, explainable AI, and real-time approaches remained limited. The absence of banking-specific multimodal datasets is a key barrier to progress. Future research should prioritize privacy-preserving multimodal datasets, multilingual and culturally diverse models, explainable AI, and real-time deployment to support more reliable assessment of customer experience in digital banking.

Article Info

Keywords:
Sentiment Analysis, Digital Banking, Scoping Review, Multimodal, Customer Satisfaction, Text-Based Data, Traditional Machine Learning, Hybrid Approach.

Article History:
Received: 03-July-2026,
Revised: 07-August-2026,
Accepted: 12-August-2026,
Published: 28 October 2026.

1. INTRODUCTION

Sentiment analysis, also known as opinion mining, focuses on evaluating the feelings and opinions produced by individuals through various modalities and converting these opinions into measurable information. This field has demonstrated its ability to absorb and accurately understand the vast amounts of data generated by customers. Sentiment analysis has become increasingly important in analyzing the content of social media and digital technologies, as it has been exploited in various fields, including education and finance. In the field of education [1], sentiment analysis techniques based on transformers, such as BERT, have been used to monitor and evaluate students' views on the effectiveness of teaching. In financial services, sentiment analysis techniques have been used to predict cryptocurrency trends, achieving high prediction accuracy using advanced models such as the HSIF model [2]. Similarly, mobile banking applications such as Seabank [3] have revolutionized customer service by enhancing ease of use and accessibility. Focusing on improving customer satisfaction, system quality, information, and awareness of perceived risks increases satisfaction and the intention to continue using these applications while reducing abandonment [4]. Satisfied customers are more likely to recommend the use of banking applications than dissatisfied customers [5]. A study by Liem et al. [6] found that a transfer learning model achieved outstanding performance in sentiment analysis tasks within digital banking. A study [7] showed that analyzing customer reviews can help banks identify

©2026 The Author(s). This article is licensed under CC By 4.0

weaknesses and strengths in their services and better understand customer concerns. Current research on digital banking still relies primarily on text-based data [8], neglecting the integration of multimodal approaches. This limitation hinders the understanding of customer sentiment in digital banking environments. Recent studies have shown that relying on integrated modalities provides a better understanding of implicit emotions [9, 10]. Previous reviews of multimodal sentiment analysis have prioritized fields such as marketing [11], public services [12], and health [13], with little attention paid to digital banking contexts. This reveals a research gap centered on the absence of an integrated framework that combines multimodal sentiment analysis with customer satisfaction assessment for digital banking. This scoping review followed the PRISMA-ScR guidelines and is, to the best of our knowledge, the first scoping review in digital banking. It highlights the need to integrate multiple data modalities into customer satisfaction research in digital banking. This scoping review contributes to the development of a map of current research, including datasets, methodologies, and application areas employed in previous studies, while also identifying future research directions toward multimodal integration and explainable AI in digital banking. More specifically, we are interested in answering the following main research question: What is the current state of sentiment analysis research in digital banking? Accordingly, the following research questions are addressed:

RQ1: What are the most prominent application domains of sentiment analysis in digital banking?

RQ2: What methodologies are currently used for sentiment analysis in digital banking?

RQ3: What types of data are used in sentiment analysis for digital banking?

RQ4: What are the research gaps in sentiment analysis for digital banking?

This scoping review aims to map the current research on sentiment analysis in digital banking and identify research gaps by analyzing the included studies.

2. REVIEW METHODOLOGY

2.1. Study Design

To answer the research questions, a scoping review of the published literature on sentiment analysis in digital banking was conducted. Given the diversity of study designs, a scoping review was considered more appropriate than a systematic review, which primarily focuses on evaluating the effectiveness of specific models or approaches. This review was conducted in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [14] guidelines to ensure reproducibility and transparency. In this review, no pre-established protocol was used; however, the review followed specific eligibility criteria, research procedures, and study selection processes to ensure transparency and consistency.

2.2. Defining Research Questions

The research questions were identified and formulated using the population–concept–context (PCC) framework recommended for scoping reviews. The population comprised customers and users of digital banking services. The concept focused on sentiment analysis approaches and the models used to analyze sentiments and opinions in digital banking. The context included digital banking environments, such as mobile banking applications, online banking, digital wallets, fintech platforms, and chat banking systems. The development of the review questions, eligibility criteria, and search strategy was guided by the PCC framework.

2.3. Search Strategy

The search strategy was developed based on the main objectives and research questions of the review, as well as the fundamental concepts of sentiment analysis, customer satisfaction, and digital banking. The searches were conducted in the following databases: ScienceDirect, IEEE Xplore, Wiley Online Library, ACM Digital Library, SpringerLink, and Google Scholar. The search was limited to studies published in English from 2015 to 2026. The final search was conducted on May 10, 2026. Initially, a search was conducted using the term "sentiment analysis" to check for any ongoing or planned comprehensive or systematic reviews on this topic on the PROSPERO, OSF, and Figshare platforms. The review protocol was not registered. Preliminary searches retrieved many records. Therefore, the strategy was improved by reducing the number of Boolean operators, adding another Boolean operator, and including an additional exclusion criterion. The final search query was as follows: ("sentiment analysis" OR "opinion mining" OR "emotion recognition") AND ("digital banking" OR "online banking" OR "mobile banking" OR "financial services" OR "fintech") AND ("customer review" OR "user feedback" OR "customer experience" OR "social media" OR "chatbot"). The query was executed with slight variations in the search terms because each database has a different indexing system. However, the scope and core concept of the review remained unchanged. The search strategies used for each database are shown in Table (1).

2.4. Eligibility Criteria

After conducting the database search, a set of eligibility criteria was established to determine the studies relevant to the objectives of this scoping review. The inclusion

©2026 The Author(s). This article is licensed under CC By 4.0

Table[1]: Search Queries and Retrieved Records by Database.

Database	Query string	records
Google Scholar	("sentiment analysis") AND ("digital banking") AND ("customer review" OR "user feedback")	286
IEEE Xplore	("sentiment analysis" OR "opinion mining") AND ("banking" OR "fintech") AND ("customer review" OR "chatbot" OR "user feedback")	43
ScienceDirect	("sentiment analysis" OR "opinion mining") AND ("banking" OR "fintech" OR "financial services") AND ("customer review" OR "user feedback" OR "chatbot")	584
Wiley Online Library	(" sentiment analysis" OR "opinion mining") AND ("digital banking" OR "fintech") AND ("customer review" OR "user experience")	24
SpringerLink	("sentiment analysis" OR "opinion mining" OR "emotion recognition") AND ("digital banking" OR "online banking" OR "mobile banking" OR "fintech") AND ("customer review" OR "user feedback" OR "customer experience")	234
ACM Digital Library	("sentiment analysis" OR "opinion mining") AND ("banking" OR "fintech") AND ("customer review" OR "chatbot" OR "user feedback")	460
Total	1631	

Table[2]: Inclusion and Exclusion Criteria.

Inclusion Criteria	**Exclusion Criteria**
The study was published in English between 2015 and 2026.	The study was written in a language other than English.
The study was published in peer-reviewed scientific journals or conference proceedings.	The study was a report, editorial, thesis, book, or patent that had not been peer-reviewed
The study addressed sentiment analysis or opinions in the context of banking services and customer satisfaction.	The study did not address sentiment analysis, banking services, or customer satisfaction.

and exclusion criteria are shown in Table (2).

2.5. Study Selection Process

An optimized search string was used. The search yielded 1,631 records. After applying filters to the selected databases for journal articles, conference papers, relevance to sentiment analysis and digital banking, English-language studies, and the publication period from 2015 to 2026, the search yielded 1,097 results, which were imported into Zotero reference management software. After removing web pages, books, and dissertations, 632 records were obtained. After removing duplicates, 612 records were obtained. The results were exported into CSV format and manually screened using Microsoft Excel. After reading the study titles and then the abstracts by the reviewers, 131 studies were obtained, and the remaining studies were excluded because they did not address sentiment analysis for customer satisfaction in digital banks. After applying filtering using full-text review, 46 studies met all inclusion criteria. Of these, 35 studies were included. Eight were excluded by the authors after they agreed to exclude them because they did not address customer satisfaction. Three studies were excluded because the full texts were not available in English, which was an a priori inclusion criterion [15–17]. No statistical calculations, such as Cohen's kappa coefficient, were used to assess inter-reviewer reliability. However, the reviewers independently examined titles and abstracts and evaluated the full texts. Any disagreements regarding study inclusion were resolved through discussion until a consensus was reached. The PRISMA diagram in Figure (1) illustrates the selection process.

2.6. Quality Assessment

This scoping review followed a scoping review methodology. Therefore, no formal methodological quality or risk-of-bias assessment was conducted for the included studies. This decision was consistent with the objectives of this scoping review, which were to map the breadth and characteristics of research on sentiment analysis for customer satisfaction in digital banking, including application domains, data types, analytical methodologies, sentiment-classification approaches, and research gaps, rather than to evaluate intervention effectiveness or determine the comparative certainty of evidence. In scoping reviews, critical appraisal of included studies is optional and may be undertaken depending on the objectives and purpose of the review [14]. Accordingly, studies were included based on the predefined eligibility criteria rather than their methodological quality. Nevertheless, the absence of a formal quality assessment means that the methodological robustness and risk of bias of individual studies could not be systematically evaluated or compared. Therefore, the findings presented in this review should be interpreted primarily as a mapping of the

©2026 The Author(s). This article is licensed under CC By 4.0

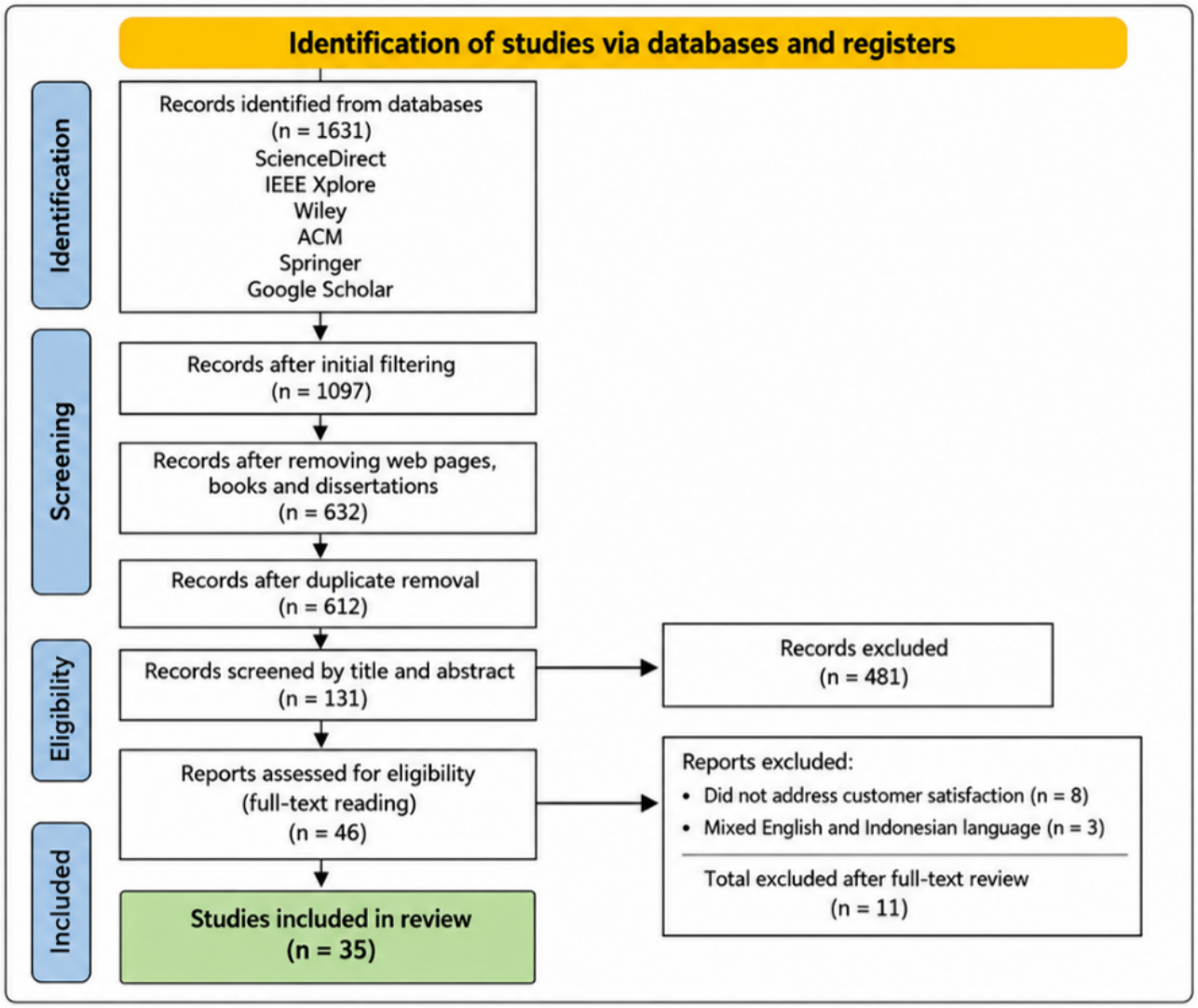

Figure 1. PRISMA flow diagram for study selection.

existing research landscape rather than as an assessment of the quality, reliability, or certainty of the available evidence.

2.7. Data Extraction and Compilation

The following data were extracted from the included studies: standard bibliographic characteristics, such as the year of publication and the authors' names, the country of publication and the application area (e.g., Digital Banking Services; Digital Banking and Mobile Banking); data type (text, time series, audio, video, or multimodal) and sentiment classes; data collection sources (e.g., customer reviews, social media); analytical methodologies (Machine Learning, Deep Learning, Hybrid Approaches, Transformer/LLM-based Approaches, and Lexicon/Statistical/Topic Modeling Approaches); model names; and research gaps. The reviewers extracted the data using a standardized data extraction form developed in Microsoft Excel. Before full data extraction, the standardized data extraction form was piloted on five selected studies, representing 14.3% of the 35 studies included in the review. The pilot subset was selected to reflect variation in application domains, data types, and methodological approaches. Based on the pilot test, the form was refined to improve the consistency of data extraction by clarifying the classification of data types, standardizing the five methodological categories, distinguishing sentiment classes from rating-based outcomes, and ensuring consistent recording of dataset characteristics, model names, and reported research gaps. The revised form was then applied to all 35 included studies. The extracted data were summarized quantitatively where appropriate and qualitatively where necessary. The included studies were categorized into five methodological groups to facilitate comparison among the included studies. The reviewers classified the studies after reading them, examining the methodologies used in each study, and reviewing the models and feature extraction techniques. The reviewers agreed on the following methodological categories: Traditional Machine Learning (ML), Deep Learning (DL), Hybrid Approaches, Transformer/LLM-based Approaches,

©2026 The Author(s). This article is licensed under CC By 4.0

and Lexicon/Statistical/Topic Modeling Approaches. In addition, the reviewers independently identified the application areas after reviewing the study title, objective, and dataset used in each study. The application areas were then categorized into five groups after the researchers reached a consensus: Digital Banking Services, Digital Banking and Mobile Banking, Digital Wallets and Payment Systems, Credit Card and Payment Services, and Conversational Banking Systems. Finally, these classifications enabled us to identify areas and methodologies that have not yet been explored in the field of digital banking. The extracted data are summarized in Table (3).

Table[3]: Characteristics of the Included Studies (continued in Appendix A)

Ref	Year	Country	Domains	Data type	Sentiments	Approaches
[18]	2025	Indonesia	Digital Wallets and Payment Systems	Text	Positive / Negative	Deep Learning (DL)
[19]	2024	United Kingdom	Digital Banking Services	Text	Positive / Neutral / Negative	Hybrid Approach
[20]	2026	India	Digital Wallets and Payment Systems	Text + Survey	Positive / Neutral / Negative	Hybrid Approach
[21]	2024	Indonesia	Digital Banking Services	Text	Positive / Neutral / Negative	Traditional Machine Learning (ML)
[22]	2026	Indonesia	Digital Wallets and Payment Systems	Text	Positive / Neutral / Negative	Hybrid Approach
[23]	2025	Indonesia	Digital Banking Services	Text	Positive / Negative	Traditional Machine Learning (ML)
[24]	2024	Indonesia	Digital Banking and Mobile Banking	Text	Positive / Neutral / Negative	Lexicon / Statistical / Topic Modeling
[25]	2024	Indonesia	Digital Banking and Mobile Banking	Text	Positive / Negative	Traditional Machine Learning (ML)
[26]	2024	Indonesia	Digital Banking and Mobile Banking	Text	Positive / Negative	Traditional Machine Learning (ML)
[27]	2025	Indonesia	Digital Banking and Mobile Banking	Text	Positive / Neutral / Negative	Hybrid Approach
[28]	2023	Bangladesh	Digital Banking and Mobile Banking	Text	Positive / Negative	Traditional Machine Learning (ML)
[29]	2025	Indonesia	Digital Banking and Mobile Banking	Text	Positive / Neutral / Negative	Lexicon / Statistical / Topic Modeling
[30]	2024	Indonesia	Digital Banking and Mobile Banking	Text	Positive / Negative	Traditional Machine Learning (ML)
[31]	2024	USA	Digital Banking Services	Text	Positive / Neutral / Negative	Hybrid Approach
[32]	2025	Indonesia	Digital Banking and Mobile Banking	Text	Positive / Negative	Transformer / LLM-based Approaches
[33]	2025	Saudi Arabia	Digital Banking and Mobile Banking	Text	Positive / Negative / Conflict	Hybrid Approach
[34]	2024	India	Digital Wallets and Payment Systems	Text	Positive / Negative	Traditional Machine Learning (ML)
[35]	2025	Yemen	Digital Banking and Mobile Banking	Text	Positive / Neutral / Negative	Hybrid Approach

©2026 The Author(s). This article is licensed under CC By 4.0

[36]	2026	Indonesia	Digital Banking and Mobile Banking	Text	Positive / Neutral / Negative	Hybrid Approach
[37]	2023	Indonesia	Conversational Banking Systems	Text	Positive / Neutral / Negative	Hybrid Approach
[38]	2025	Indonesia	Digital Banking and Mobile Banking	Text	Positive / Neutral / Negative	Deep Learning (DL)
[39]	2025	Türkiye	Digital Banking and Mobile Banking	Numerical (Time Series)	Implicit (Rating-based)	Deep Learning (DL)
[40]	2026	Indonesia	Digital Banking and Mobile Banking	Text	Positive / Negative / Neutral	Lexicon / Statistical / Topic Modeling
[41]	2023	Indonesia	Digital Banking Services	Text	Positive / Negative / Neutral	Traditional Machine Learning (ML)
[42]	2024	Turkey	Digital Banking and Mobile Banking	Text	Positive / Negative / Neutral	Hybrid Approach
[43]	2021	USA	Credit Card and Payment Services	Multi-source Text	Positive / Negative / Neutral	Hybrid Approach
[44]	2024	Indonesia	Digital Banking and Mobile Banking	Text	Positive / Negative / Neutral	Deep Learning (DL)
[45]	2024	Nigeria	Digital Wallets and Payment Systems	Text	Positive / Negative / Neutral	Lexicon / Statistical / Topic Modeling
[46]	2026	Indonesia	Digital Banking Services	Text	Positive / Negative	Hybrid Approach
[47]	2025	Southeast Europe	Digital Banking and Mobile Banking	Text	Positive / Negative / Neutral	Hybrid Approach
[48]	2025	Indonesia	Digital Banking and Mobile Banking	Text	Positive / Negative	Traditional Machine Learning (ML)
[49]	2024	Indonesia	Digital Banking and Mobile Banking	Text	Positive / Negative	Lexicon / Statistical / Topic Modeling
[50]	2025	Tunisia	Digital Banking and Mobile Banking	Text	Positive / Negative / Neutral	Transformer / LLM-based Approaches
[51]	2024	Indonesia	Digital Banking and Mobile Banking	Text + Survey	Positive / Negative / Neutral	Hybrid Approach
[52]	2025	Indonesia	Digital Banking and Mobile Banking	Text	Positive / Negative	Transformer / LLM-based Approach

3. RESULTS

In this scoping review, the results obtained from the selected studies were analyzed using both numerical and thematic analyses.

3.1. Overview of Included Studies (Main Research Question)

A total of 35 studies related to sentiment analysis in digital banking were identified, covering the period from 2015 to 2026. The findings indicated a rapid growth in sentiment analysis research within the field of digital banking, especially in recent years. No eligible studies were identified from 2015 to 2020 or in 2022; the first eligible publication appeared in 2021, followed by a marked increase in 2024 and 2025, as shown in Figure (2). The findings also showed the dominance of unimodal text-based data over multi-source textual data (reviews, chats, surveys) during this period. In contrast, multimodal approaches that integrate text, audio, and

©2026 The Author(s). This article is licensed under CC By 4.0

video data remain underutilized. This finding highlights a significant gap in the current literature.

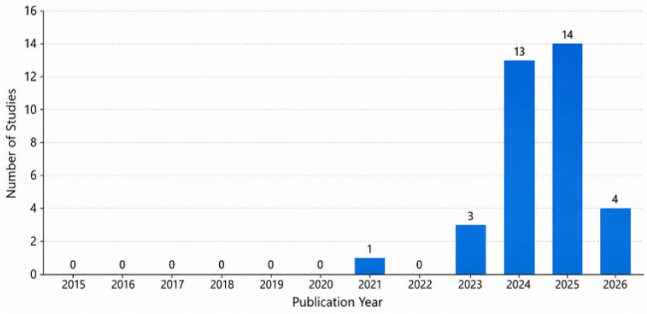

Figure 2. Distribution of included studies by publication year; no studies were identified in 2015–2020 or 2022.

3.2. Domains of Application in Digital Banking (RQ1)

The included studies were classified into five categories based on their characteristics, as shown in Figure (3). Digital banking and mobile banking accounted for 22 studies. Second, digital banking services accounted for six studies, reflecting the studies' focus on customer satisfaction with digital banking services. Third, digital wallets and payment systems accounted for five studies. Finally, emerging areas received limited attention, with conversational banking systems and credit card and payment services each represented by only one study.

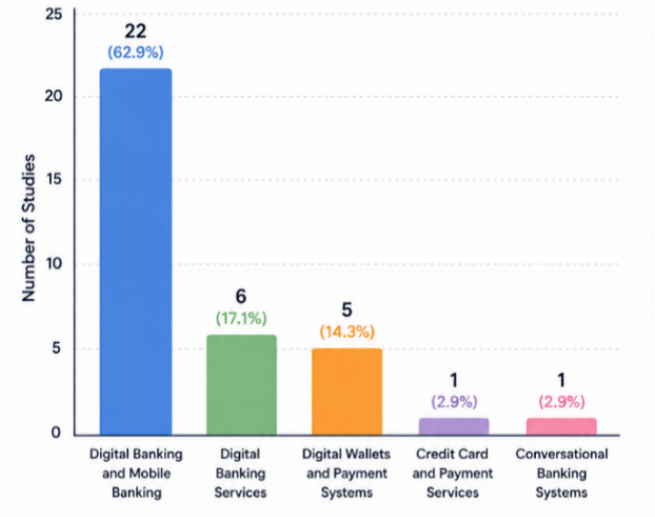

Figure 3. Distribution of included studies by application domain.

3.3. Methodologies Used in Previous Studies (RQ2)

Figure (4) summarizes the methodological approaches used in the included studies. The studies were classified into five categories: traditional machine learning (ML), deep learning (DL), hybrid approaches (ML+DL, ML + lexicon-based, and hybrid AI frameworks), lexicon/statistical/topic modeling, and transformer/LLM-based approaches. Methodologies were classified as hybrid when they combined two or more methodological families, while the remaining methodologies were retained as defined in the original studies. Deep learning was employed in four studies. These studies utilized CNN and LSTM techniques, achieving reported accuracies between 91% and 98%. Lexicon/statistical/topic modeling approaches were employed in five studies, achieving classification performance that reached 84%. Some of these studies did not mention accuracy but addressed sentiment classification, topic modeling, and customer satisfaction indicators. Transformer- or LLM-based approaches were employed in three studies. These approaches achieved accuracies ranging from 43% to 99%, relying on models such as LLMs, BERT, IndoBERT, and ABSA. This indicates that despite advances in these models, their application in digital banking environments remains limited. The studies included varied considerably in terms of dataset sizes, models used, and performance metrics; therefore, direct comparisons of the performance of these models should not be considered reliable standards.

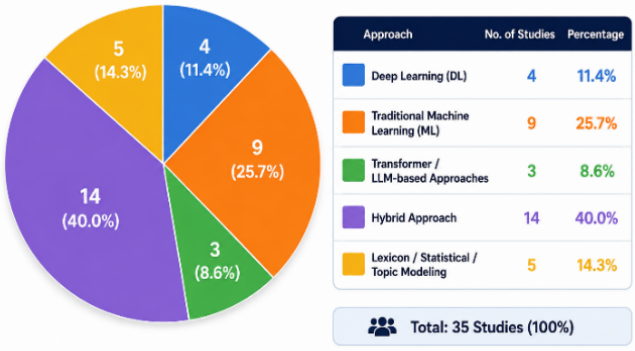

Figure 4. Methodological approaches used in the included studies.

3.4. Types of Data Used in Digital Banking Services (RQ3)

The included studies utilized textual data, multi-source text, numerical data (time series), or text combined with survey data, as shown in Figure (5). Most studies analyzed textual data (31 studies, 88.6%). In contrast, only one study (2.9 %) used multi-source text (e.g., reviews,

©2026 The Author(s). This article is licensed under CC By 4.0

chats, and surveys). Two studies (5.7 %) employed a combination of text and survey data, using customer feedback collected from the Google Play Store and surveys distributed to participants. Only one study (2.9%) analyzed numerical data (time series) [39], relying on a Google Play Store review dataset. Overall, the findings indicate that most studies relied on textual data, neglecting multimodal data such as video and audio, which remain largely untapped.

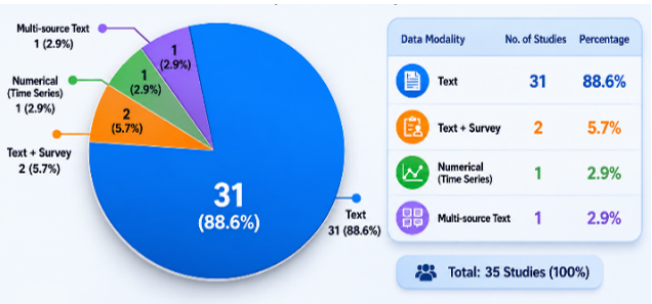

Figure 5. Data types used in the included studies.

3.5. Geographical Distribution of Included Studies in Digital Banking

This section presents the geographical distribution of the included studies on digital banking, as illustrated in Figure (6). Most studies focused on Southeast Asia, particularly Indonesia, which accounted for the largest number of studies related to digital banking, with a total of 22 studies. India, Turkey, and the USA followed, with 2 studies each. In contrast, the Middle East and other regions were represented by only a few studies. These findings revealed an uneven geographical distribution of research related to sentiment analysis in digital banking.

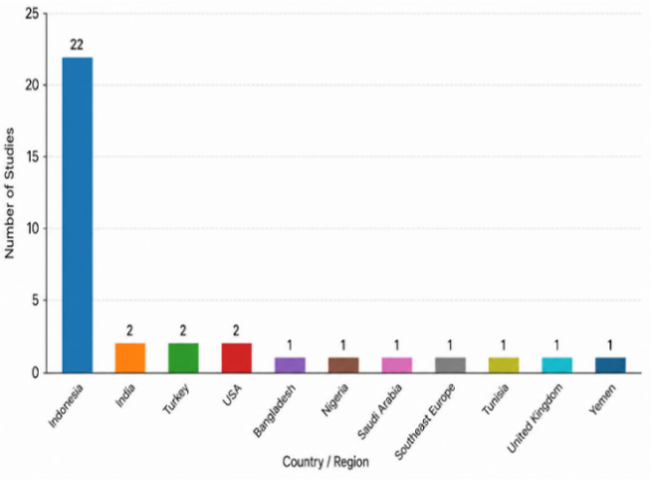

Figure 6. Geographical distribution of the included studies.

3.6. Sentiment Classification of the Included Studies

The findings of this review revealed that the majority of studies relied on a three-class classification (positive, negative, and neutral) to categorize customer emotions in digital banking, representing 22 studies, as illustrated in Figure (7). However, some studies relied on a two-class classification (positive and negative), which was used in 12 studies. Only one study focused on sentiment scores, in which sentiment was classified using numerical values. Overall, the findings indicate that most studies adopted a three-class classification for sentiment analysis in digital banking.

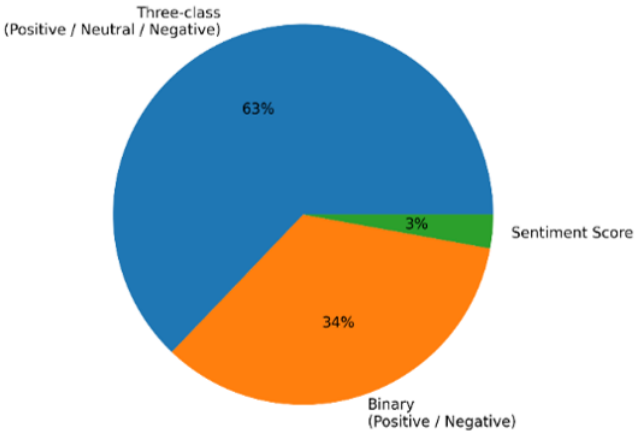

Figure 7. Sentiment-classification schemes used in the included studies.

3.7. Research Gaps in Sentiment Analysis for Digital Banking (RQ4)

The included studies revealed four principal research gaps. First, modality integration is almost absent. Most studies analyzed text only, whereas audio, video, and synchronized multimodal signals received little or no attention. Second, banking-specific multimodal datasets are lacking, and many studies relied on relatively small, single-platform datasets. Third, advanced methods—including deep learning, transformer architectures, and large language models—were applied less frequently than traditional machine-learning and hybrid approaches. Finally, explainable AI was rarely incorporated; most studies emphasized predictive performance without providing sufficient interpretation of model decisions. These gaps constrain generalizability, real-time deployment, and the ability to capture implicit or conflicting customer emotions.

4. DISCUSSION

©2026 The Author(s). This article is licensed under CC By 4.0

4.1. Principal Findings

This scoping review identified and summarized research on sentiment analysis in digital banking. Five research questions addressed the current state of the field, application domains, methodologies, data types, and research gaps. The review included 35 studies published between 2015 and 2026, a period characterized by rapid advances in artificial intelligence, particularly transformer-based models such as BERT and GPT [53]. Text-only data dominated because customer reviews are readily available, inexpensive to collect, and less computationally demanding than synchronized audio-visual data. Digital and mobile banking were the most common application domains because app stores and social-media platforms generate large volumes of customer feedback. Traditional machine learning and hybrid approaches were used more frequently than transformer/LLM-based methods, probably because many datasets were small and because conventional models require fewer computational resources. However, these approaches often depend on manually designed features and may capture contextual and implicit sentiment less effectively. Transformer and hybrid AI approaches offer stronger contextual modeling but introduce higher computational costs, longer training times, privacy concerns, and challenges in explaining model decisions. The evidence was also geographically concentrated, particularly in Indonesia, which limits cross-cultural generalizability. Taken together, these findings indicate that sentiment analysis in digital banking is growing rapidly but remains methodologically fragmented and predominantly text-based.

4.2. Toward Multimodal Sentiment Analysis in Digital Banking

Although multimodal sentiment analysis has been applied in e-commerce [54], healthcare [55], and education [56], none of the included digital-banking studies integrated synchronized text, audio, and video. Multimodal analysis could capture discrepancies between what customers say and how they express it—for example, a neutral written statement accompanied by a frustrated tone or facial expression. However, this potential benefit does not justify indiscriminate collection of audio or video. Adoption should be proportional to a clearly defined service or research purpose and evaluated against privacy, security, fairness, accessibility, regulatory, and operational requirements.

4.2.1. Ethical, Privacy, Security, and Governance Challenges

Banking data are especially sensitive because multimodal records may contain account information, identity documents, faces, voice characteristics, device identifiers, behavioral patterns, and contextual details that enable re-identification. Dataset development should therefore apply data minimization, purpose limitation, explicit and revocable consent where appropriate, retention limits, privacy by design, and security controls throughout the data lifecycle [57, 58]. Identity-verification data should be separated from sentiment-research data, and data-protection and algorithmic-impact assessments should be completed before collection. De-identification must address not only names and account numbers but also quasi-identifiers embedded in speech, images, metadata, and conversational content. Encryption, role-based access, audit logging, secure research environments, and documented incident-response procedures are required. Where centralization is inappropriate, federated learning and differential privacy may reduce raw-data exposure [59, 60], although privacy-utility trade-offs must be reported. Ethical governance must also address demographic, linguistic, and cultural bias. Performance should be audited across relevant groups, including gender, age, accent, dialect, disability, and cultural context, with human oversight and clear mechanisms for contesting errors. Inferred sentiment should not be used as the sole basis for credit eligibility, pricing, fraud labeling, or other high-stakes decisions.

4.2.2. Sources and Development of Banking-Specific Multimodal Datasets

Potential sources for banking-specific multimodal datasets include customer service calls, chatbot and live-chat interactions, video-based customer support sessions, mobile banking application feedback, social media interactions, and behavioral or application-usage data. Where ethically and legally permissible, these sources can be combined to construct aligned datasets containing text, audio, video, and complementary behavioral signals. Dataset development would require preprocessing, anonymization, temporal synchronization across modalities, sentiment annotation, quality control, and appropriate training, validation, and test partitions. Existing benchmark datasets such as CMU-MOSI [61] and CMU-MOSEI [62] may provide useful methodological references; however, they were developed outside the digital banking context and do not capture banking-specific language, customer concerns, privacy requirements, or regulatory constraints. Therefore, developing representative, secure, multilingual, and banking-specific multimodal datasets represents an important direction for future research.

4.3. Limitations of the Review

This review has some limitations. A limitation of this review is that no formal methodological quality or risk-of-bias assessment was conducted for the included studies. Although such an assessment is not mandatory in scoping reviews and was not required to achieve the mapping

©2026 The Author(s). This article is licensed under CC By 4.0

objectives of this study [14], its absence limits the extent to which the methodological robustness and potential biases of individual studies can be evaluated and compared. Consequently, the synthesized findings should not be interpreted as an indication that all identified studies provide equally reliable or methodologically robust evidence. Future reviews may incorporate formal critical appraisal or risk-of-bias assessment to provide a more detailed evaluation of the strength and reliability of the evidence identified in this review. In addition, only studies published in English were included, which may have resulted in the exclusion of relevant studies published in other languages. Most of the included studies were also concentrated in Southeast Asia, particularly Indonesia, which may affect the generalizability of the findings to other geographical and cultural contexts. Although six databases were searched, some relevant studies available in other databases may not have been identified. Publication bias is another possible limitation, as studies with positive or higher-performing results may be more likely to be published. For these reasons, the findings of this review should primarily be considered a mapping of current research trends, methods, data types, and research gaps in sentiment analysis for digital banking.

4.4. Future Research

This scoping review calls for future research to focus on developing multimodal, multilingual, and multicultural models and datasets, as customers may express their satisfaction or dissatisfaction using different expressions, which can affect the performance of sentiment analysis models. Therefore, developing multilingual and multicultural models improves sentiment analysis performance and generalizability. Studies should use customer-disjoint and temporal validation, report macro-F1 and class-wise performance, audit subgroup fairness, and test cross-platform and cross-country generalization. Furthermore, future studies should be directed toward transformer-based architectures and large-scale language models for multilingual digital banking environments. Transformer and multimodal architectures should be compared with simpler baselines under consistent protocols. Research should also examine aspect-based sentiment analysis, temporal sentiment modeling, cross-lingual transfer, sarcasm detection, real-time deployment, and explainable AI such as SHAP and LIME. Governance studies should evaluate informed consent, data minimization, secure access, human oversight, and the limitations of using inferred emotion in high-stakes banking decisions. Federated learning, differential privacy, secure aggregation, and carefully validated synthetic data warrant investigation as privacy-preserving strategies, with explicit reporting of privacy-utility and realism trade-offs. The findings of this review may be influenced by geographical bias, as most of the included studies were conducted in Southeast Asia, specifically Indonesia. This may limit the ability to generalize the results to other geographical and cultural contexts.

5. CONCLUSION

This scoping review mapped research on sentiment analysis for customer satisfaction in digital banking from 2015 to 2026. The findings show that research remains overwhelmingly text-based, with no included study integrating text, audio, and video within a single sentiment analysis framework. Traditional machine-learning and hybrid approaches were more common than transformer/LLM-based methods, while the use of large language models, multimodal fusion techniques, and explainable AI remained limited. These findings highlight a substantial research gap rather than an established transition toward multimodal sentiment analysis in digital banking. Future research should therefore focus on developing banking-specific multimodal datasets and models, supporting multilingual and culturally diverse environments, integrating explainable AI, addressing privacy and governance requirements, and evaluating models for real-time applications. Such developments could support a rigorous transition from predominantly text-based sentiment analysis toward more comprehensive multimodal approaches for understanding customer satisfaction in digital banking.

6. DECLARATIONS

Ethical Approval

This article does not involve any studies with human participants or animals conducted by any of the authors. We declare that this manuscript is original and is not currently under consideration for publication elsewhere. We further confirm that the order of authors listed in the manuscript has been approved by all authors.

Conflict of Interest

The authors declare no competing interests.

Authors' Contributions

All authors contributed to the study conception and design. All authors contributed to the literature search, study selection, data extraction, and analysis, and commented on the present version of the manuscript. All authors read and approved the final manuscript.

Funding

No funding was received from any organization for this work.

Availability of Data and Materials

No datasets were used in the present study.

Consent for Publication

The authors confirm that informed consent was obtained for the publication of the data contained in this article.

Consent to Participate

Informed consent was obtained from all authors.

©2026 The Author(s). This article is licensed under CC By 4.0

Declaration of AI-Assisted Technologies

"The authors declare that artificial intelligence (AI) tools were used only for limited purposes, such as language editing, grammar improvement, and formatting assistance. All scientific content, data analysis, interpretations, conclusions, and final manuscript revisions were reviewed and verified by the authors, who take full responsibility for the integrity and originality of the work."

REFERENCES

[1] H. Shuqin and R. C. Raga, "A deep learning model for student sentiment analysis on course reviews," *IEEE Access*, vol. 12, pp. 136 747–136 758, 2024. DOI: 10.1109/ACCESS.2024.3463793

[2] S. M. Dashtaki, M. H. Chagahi, A. Bahadori, B. Moshiri, M. J. Piran, and A. Montazeri, "HSIF: A transformer-based cross-attention framework for cryptocurrency trend forecasting via multimodal sentiment–market fusion," *IEEE Access*, vol. 13, pp. 156 600–156 612, 2025. DOI: 10.1109/ACCESS.2025.3605522

[3] F. Amzar, W. Cholil, I. H. Indriana, Prihandoko, R. Zulfahmi, and R. Purabaya, "Sentiment analysis of seabank application reviews on google play store using naive bayes classifier," in *2024 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, Nov. 2024, pp. 126–132. DOI: 10.1109/ICIMCIS63449.2024.10956937

[4] W. Ahmed, M. A. Al-Sharafi, A. Raza, S. A. S. Al-Zaeemi, M. A. Al-Bashrawi, and Y. K. Dwivedi, "A longitudinal big data approach to theorizing consumers' continuance intention to use loyalty apps," *J. Retail. Consumer Serv.*, vol. 88, p. 104 453, Jan. 2026. DOI: 10.1016/j.jretconser.2025.104453

[5] B. Basu, M. P. Sebastian, and A. K. Kar, "What affects the promoting intention of mobile banking services? insights from mining consumer reviews," *J. Retail. Consumer Serv.*, vol. 77, p. 103 695, Mar. 2024. DOI: 10.1016/j.jretconser.2023.103695

[6] B. R. Liem and D. A. Kristiyanti, "Sentiment analysis approaches for mobile banking app reviews using transfer learning method," in *2024 7th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, Dec. 2024, pp. 830–835. DOI: 10.1109/ISRITI64779.2024.10963599

[7] J. Santoso, N. S. Novanto, Y. R. Franklin, and E. D. Madyatmadja, "Sentiment analysis on google play store user reviews of digital bank applications in indonesia," in *Technology: Toward Business Sustainability*, B. Alareeni and A. Hamdan, Eds., Cham: Springer Nature Switzerland, 2024, pp. 318–327. DOI: 10.1007/978-3-031-54019-6_29

[8] M. Jayanthakumaran, N. Shukla, B. Pradhan, and G. Beydoun, "A systematic review of sentiment analytics in banking headlines," *Decis. Anal. J.*, vol. 15, p. 100 584, Jun. 2025. DOI: 10.1016/j.dajour.2025.100584

[9] C. H. P. Panjaitan, D. Manongga, Hendry, and R. G. Maypopu, "A systematic literature review on multimodal sentiment analysis trends, challenges and opportunities," in *2025 4th International Conference on Creative Communication and Innovative Technology (ICCIT)*, Aug. 2025, pp. 1–7. DOI: 10.1109/ICCIT65724.2025.11167526

[10] A. Ghorbanali and M. K. Sohrabi, "A comprehensive survey on deep learning-based approaches for multimodal sentiment analysis," *Artif. Intell. Rev.*, vol. 56, no. 1, pp. 1479–1512, Oct. 2023. DOI: 10.1007/s10462-023-10555-8

[11] I. César et al., "A systematic review on responsible multimodal sentiment analysis in marketing applications," *IEEE Access*, vol. 12, pp. 111 943–111 961, 2024. DOI: 10.1109/ACCESS.2024.3441514

[12] H. Sethi and R. Gupta, "Artificial intelligence-based sentiment analysis models in the product and service industry: A state-of-the-art review and future directions," *SN Comput. Sci.*, vol. 6, no. 7, p. 774, Aug. 2025. DOI: 10.1007/s42979-025-04317-2

[13] K. Denecke and D. Reichenpfader, "Sentiment analysis of clinical narratives: A scoping review," *J. Biomed. Informatics*, vol. 140, p. 104 336, Apr. 2023. DOI: 10.1016/j.jbi.2023.104336

[14] A. C. Tricco et al., "PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation," *Ann. Intern. Med.*, vol. 169, no. 7, pp. 467–473, 2018. DOI: 10.7326/M18-0850

[15] M. J. Prasetyo and I. M. A. Agastya, "Sentiment analysis of banking application reviews on google play store using support vector machine algorithm," *Sistemasi: J. Sistem Informasi*, vol. 13, no. 6, pp. 2386–2400, Nov. 2024. DOI: 10.32520/stmsi.v13i6.4536

[16] M. Akbar and I. Pratama, "Sentiment analysis of mybca application user reviews using naive bayes, random forest, and decision tree," *SISTEMASI*, vol. 14, p. 2464, Sep. 2025. DOI: 10.32520/stmsi.v14i5.5472

[17] A. A. P. Simarmata and T. B. Sasongko, "Sentiment analysis on brimo application reviews using indobert," *J. Appl. Informatics Comput.*, vol. 9, no. 3, pp. 851–862, Jun. 2025. DOI: 10.30871/jaic.v9i3.8162

[18] G. M. Susanto and T. Mauritsius, "Product backlog recommendations using sentiment analysis with topic modelling and text summarization in the xyz e-wallet application," in *2025 29th International Conference on Information Technology (IT)*, Feb. 2025, pp. 1–6. DOI: 10.1109/IT64745.2025.10930284

[19] A. Ghadiridehkordi, J. Shao, R. Boojihawon, Q. Wang, and H. Li, "Leveraging sentiment analysis via text mining to improve customer satisfaction in uk banks," *Int. J. Bank Mark.*, vol. 43, no. 4, pp. 780–802, 2025. DOI: 10.1108/IJBM-05-2024-0323

[20] E. Seema, R. Suriakala, and P. V, "A machine learning based model for customer satisfaction assessment in digital wallet systems," in *2026 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, Jan. 2026, pp. 1–8. DOI: 10.1109/ICSES66558.2026.11479017

[21] Y. Yudhistira and A. S. Talita, "Analyzing public sentiment towards bsi service disruptions through x: Naive bayes algorithm," *Sink. J. dan Penelitian Tek. Informatika*, vol. 8, no. 3, pp. 1651–1661, 2024. DOI: 10.33395/sinkron.v8i3.13729

[22] N. Latifah, D. H. E. Putra, and F. Nugraha, "Time-series monitoring of sentiment dynamics in reviews of four indonesian e-wallet applications using a hybrid tf-idf and bi-lstm framework," *J. Inf. Syst. Informatics*, vol. 8, no. 2, pp. 1958–1983, 2026. DOI: 10.63158/journalisi.v8i2.1488

[23] S. Aisyah, "Public sentiment analysis on bank danamon and bca services using text mining and machine learning algorithms (logistic regression, svm, and catboost)," *J. Sains Data, Stat. dan Artif. Intell.*, vol. 1, no. 1, pp. 48–59, 2025.

[24] K. A. Safarah, D. I. Inan, R. Juita, and V. A. L. Sirait, "Analysing service quality using sentiment analysis and topic modeling: A case study of the livin mandiri application," *JOISIE (Journal Inf. Syst. Informatics Eng.*, vol. 8, no. 2, pp. 209–220, 2024. DOI: 10.35145/joisie.v8i2.4517

[25] Y. Christian, T. Wibowo, and M. Lyawati, "Sentiment analysis by using naive bayes classification and support vector machine, study case sea bank," *Sink. J. dan Penelitian Tek. Informatika*, vol. 8, no. 1, pp. 258–275, 2024. DOI: 10.33395/sinkron.v9i1.13141

©2026 The Author(s). This article is licensed under CC By 4.0

[26] O. S. Jelni, M. L. Radhitya, G. W. Wardhana, N. W. J. Kusuma, and N. M. M. R. Desmayani, "Sentiment analysis of brimo reviews on google play store using svm and knn," *Indonesian J. Data Sci.*, vol. 6, no. 3, pp. 548–562, 2025. DOI: 10.56705/ijodas.v6i3.365

[27] M. Nabil, A. V. Vitianingsih, S. Kacung, A. L. Maukar, and S. F. A. Wati, "Sentiment analysis of ntb syariah bank application services using the naive bayes and support vector machine methods," *J. Teknologi dan Manajemen Informatika*, vol. 11, no. 2, pp. 177–188, 2025. DOI: 10.26905/jtmi.v11i2.16311

[28] M. Dey, M. Z. Islam, and T. Rana, "Applying text mining to understand customer perception of mobile banking app," in *Handbook of Big Data and Analytics in Accounting and Auditing*, T. Rana, J. Svanberg, P. Öhman, and A. Lowe, Eds., Singapore: Springer Nature Singapore, 2023, pp. 309–333. DOI: 10.1007/978-981-19-4460-4_14

[29] R. Purnamasari and A. Thaha, "Understanding user satisfaction in digital banking: Sentiment analysis and topic modeling of indonesian bank reviews," *Bus. Informatics*, vol. 19, pp. 101–112, Sep. 2025. DOI: 10.17323/2587-814X.2025.3.101.112

[30] 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

[31] R. J. Bhuiyan et al., "Sentiment analysis of customer feedback in the banking sector: A comparative study of machine learning models," *The Am. J. Eng. Technol.*, vol. 6, no. 10, pp. 54–66, 2024. DOI: 10.37547/tajet/volume06issue10-07

[32] D. M. F. Aliffa, D. S. Maylawati, U. Syaripudin, Y. A. Gerhana, A. Wahana, and R. S. Fuadi, "Sentiment analysis of google play store reviews on the top three digital banks in indonesia using bidirectional encoder representations from transformers (bert)," in *2025 11th International Conference on Wireless and Telematics (ICWT)*, IEEE, 2025, pp. 1–5.

[33] R. Alawaji and A. Aloraini, "Sentiment analysis of digital banking reviews using machine learning and large language models," *Electronics*, vol. 14, no. 11, p. 2125, 2025. DOI: 10.3390/electronics14112125

[34] M. M. Ali, M. S. M. Ghori, R. Basri, M. S. Qaseem, M. A. Mohiuddin, and M. A. Ali, "Paytm online wallet payment's sentiment analysis in india using svm and rf model," in *2024 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*, Nov. 2024, pp. 164–169. DOI: 10.1109/3ict64318.2024.10824508

[35] S. Alhagree and G. Al-Gaphari, "Utilizing machine learning based on llm for arabic sentiment analysis in assessing user satisfaction with mobile banking apps: A case study of yemeni banks," *Sana'a Univ. J. Appl. Sci. Technol.*, vol. 3, no. 1, pp. 645–662, 2025. DOI: 10.59628/jast.v3i1.1364

[36] M. A. Ilmansyah, A. V. Vitianingsih, A. L. Maukar, S. F. A. Wati, and A. A. Aziiza, "Sentiment analysis of brimo application user reviews using naive bayes and long short-term memory," *Rabit: J. Teknologi dan Sistem Informasi Univrab*, vol. 11, no. 1, pp. 1381–1397, 2026. DOI: 10.36341/rabit.v11i1.7341

[37] N. R. Ananda, K. R. S. Wiharja, and M. A. Bijaksana, "Sentiment analysis on banking chatbot using graph-based machine learning model," in *2023 International Conference on Data Science and Its Applications (ICoDSA)*, Aug. 2023, pp. 310–315. DOI: 10.1109/ICoDSA58501.2023.10276448

[38] S. Mulyanto, D. Yonia, K. Situmorang, and B. Sutejo, "Sentiment classification of livin' by mandiri reviews in indonesia using lstm for digital banking service improvement," *J. Èkon. Kreat. dan Manajemen Bisnis Digit.*, vol. 4, pp. 141–153, Aug. 2025. DOI: 10.55047/jekombital.v4i1.1011

[39] M. Kilinc, "LSTM-based time series forecasting of user-derived quality signals in mobile banking systems," *Systems*, vol. 13, no. 11, p. 949, 2025. DOI: 10.3390/systems13110949

[40] G. X. Adikara, M. Saputra, and R. Y. Fa'rifah, "Unveiling customer technostress in indonesian mobile banking: A topic modeling and sentiment analysis of user reviews," in *2026 18th International Conference on Electronics, Computer, and Computation (ICECCO)*, IEEE, 2026, pp. 1–6.

[41] M. Karmagatri, C. F. A. Aziz, W. R. Purnama, and I. A. J. Asih, "Uncovering user perceptions toward digital banks in indonesia: A naive bayes sentiment analysis of twitter data," *J. Theor. Appl. Inf. Technol.*, vol. 101, no. 12, pp. 4960–4968, 2023.

[42] Y. S. Balcioglu, "Analyzing customer sentiments and trends in turkish mobile banking apps: A text mining study," *Dumlupinar Üniversitesi Sosyal Bilimler Dergisi*, no. 80, pp. 49–69, 2024. DOI: 10.51290/dpusbe.1391631

[43] U. Kotadiya, A. S. Arora, and T. Yachamaneni, "Ai-powered customer experience management in the credit card industry: Sentiment analysis and adaptive personalization," *Int. J. Emerg. Trends Comput. Sci. Inf. Technol.*, vol. 2, no. 2, pp. 35–44, 2021. DOI: 10.63282/3050-9246.IJETCSIT-V2I2P105

[44] D. Kusuma and C. Fatichah, "Deep learning-based sentiment classification on category service and resolution of consumer complaints in digital banking," *ILKOMNIKA*, vol. 7, no. 3, pp. 385–399, 2025. DOI: 10.28926/ilkomnika.v7i3.825

[45] B. Suleiman, E. D. Ajik, and S. A. Sule, "Leveraging sentiment analysis to optimize customer experience in digital wallets: A case of opay wallet," *Niger. J. Phys.*, vol. 33, no. 4, pp. 147–156, 2024.

[46] M. P. Casanova, A. Meiriza, K. D. Tania, N. R. Oktadini, and D. Kurniawan, "Comparative sentiment analysis of digital banking reviews using knowledge discovery indobert," *J. Appl. Informatics Comput.*, vol. 10, no. 2, pp. 1253–1262, 2026. DOI: 10.30871/jaic.v10i2.12436

[47] K. Mahmutovic, "Analyzing user-generated reviews to identify experience dimensions and their impact on satisfaction with mobile banking applications," *Èkon. vjesnik: Rev. Contemp. Entrepreneurship, Business, Econ. Issues*, vol. 38, no. 2, pp. 283–297, 2025. DOI: 10.51680/ev.38.2.6

[48] V. R. Prasetyo, N. Benarkah, and B. A. H. Rahmad, "Improving the performance of machine learning classifiers in sentiment analysis of jenius application using latent dirichlet allocation in text preprocessing," *J. Tek. Informatika (JUTIF)*, vol. 6, no. 5, pp. 3033–3050, 2025. DOI: 10.52436/1.jutif.2025.6.5.5238

[49] D. Suryadi and K. O. Padlan, "Leveraging topic modeling and sentiment analysis to improve digital bank applications," in *2024 5th International Conference on Data Analytics for Business and Industry (ICDABI)*, IEEE, 2024, pp. 86–90.

[50] R. Mahouachi and H. Mahersia, "Multilingual and tunisian dialect analysis of banking app reviews using LLMs," *Expert Syst. with Appl.*, vol. 290, p. 128 260, 2025. DOI: 10.1016/j.eswa.2025.128260

[51] Bahauddin, M. Ulfah, N. F. B. Angin, and A. R. Fachrur, "Improving the quality of mobile banking services using the integration of text mining, mobile banking service quality, and quality function deployment," *J. Ind. Serv.*, vol. 10, no. 1, pp. 117–124, 2024. DOI: 10.62870/jiss.v10i1.25907

[52] R. K. Pardamean, B. Bunyamin, and H. Hasmawati, "Aspect-based sentiment analysis of blu application reviews using indobert," in *2025 1st International Conference on Data Science and Geoinformatics (ICDSG)*, IEEE, 2025, pp. 37–42.

©2026 The Author(s). This article is licensed under CC By 4.0

[53] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, J. Burstein, C. Doran, and T. Solorio, Eds., Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 4171–4186. DOI: 10.18653/v1/N19-1423

[54] C. N. Karaeng and D. A. Kristiyanti, "Multimodal sentiment analysis approach combining transfer learning and generative ai of e-commerce product reviews," in *2025 4th International Conference on Electronics Representation and Algorithm (ICERA)*, Jun. 2025, pp. 1–6. DOI: 10.1109/ICERA66156.2025.11087376

[55] P. N. S. Kumar, R. Anitha, G. Lavanya, S. M. I. Khalilulla, and T. M. Kumar, "Multimodal sentiment analysis for healthcare," in *2026 International Conference on Machine Learning and Autonomous Systems (ICMLAS)*, Mar. 2026, pp. 871–877. DOI: 10.1109/ICMLAS67792.2026.11483761

[56] L. Kong, J. Yao, J. Shen, and Y. Gu, "A student sentiment analysis method based on multimodal deep learning," *J. Mech. Med. Biol.*, vol. 24, no. 9, p. 2 440 068, Nov. 2024. DOI: 10.1142/S0219519424400682

[57] European Parliament and Council of the European Union, "Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation)," *Off. J. Eur. Union*, vol. L 119, pp. 1–88, May 2016.

[58] National Institute of Standards and Technology, "Nist privacy framework: A tool for improving privacy through enterprise risk management," National Institute of Standards and Technology, NIST Cybersecurity White Paper CSWP 10, version 1.0, Jan. 2020. DOI: 10.6028/NIST.CSWP.01162020

[59] S. Dasari and R. Kaluri, "2P3FL: A novel approach for privacy preserving in financial sectors using flower federated learning," *Comput. Model. Eng. & Sci.*, vol. 140, no. 2, pp. 2035–2051, 2024. DOI: 10.32604/cmes.2024.049152

[60] J. Fu et al., "Differentially private federated learning: A systematic review," *ACM Comput. Surv.*, vol. 58, no. 11, pp. 1–38, 2026, Article no. 271. DOI: 10.1145/3801079

[61] A. Zadeh, R. Zellers, E. Pincus, and L.-P. Morency, "MOSI: Multimodal corpus of sentiment intensity and subjectivity analysis in online opinion videos," *arXiv preprint arXiv:1606.06259*, 2016. DOI: 10.48550/arXiv.1606.06259

[62] A. B. Zadeh, P. P. Liang, S. Poria, E. Cambria, and L.-P. Morency, "Multimodal language analysis in the wild: CMU-MOSEI dataset and interpretable dynamic fusion graph," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2018, pp. 2236–2246. DOI: 10.18653/v1/P18-1208

©2026 The Author(s). This article is licensed under CC By 4.0

Appendix A

Table[4]: Characteristics of the included studies

Authors	Dataset	Model Name	Research Gaps
Giovan Meidy Susanto et al. [18]	2748 documents on the Google Play Store	LSTM	Limited use of sentiment analysis for backlog generation and lack of comparative studies
Amirreza Ghadiridehkordi et al. [19]	23,079 reviews from Trustpilot, a consumer review platform	DistilBERT and SVM	Need for multi-country studies, broader bank coverage, and integration of multiple customer feedback sources
E. Seema et al. [20]	624 responses: hybrid survey and textual data	ML, K-means Clustering, Hierarchical Clustering	Lack of AI-driven customer satisfaction assessment models capable of capturing implicit sentiment
Yudhistira et al. [21]	991 tweets	Naïve Bayes	Single social media platform, short observation period, reliance on NB and lexicon methods, and Indonesian-language data only
Noor Latifah et al. [22]	13,823 labeled reviews	BiLSTM, Logistic Regression	Lack of temporal sentiment monitoring and cross-application temporal analysis
Siti Aisyah et al. [23]	2,000 customer reviews from Google Reviews	LR, SVM, and CatBoost	Need for Transformer/DL models and temporal/time-series sentiment analysis
Khairin Andini Safarah et al. [24]	13,692 reviews from Google Play Store	LSI, Orange Data Mining Tool	Lack of integrated sentiment analysis and topic modeling studies in mobile banking
Yefta Christian et al. [25]	13,692 reviews from the Google Play Store	NBC, SVM	Single-source data- binary classification-small dataset-limited SVM–NB comparison.
Olivia Sutriani Jelni et al. [26]	15,233 reviews from Google Play Store	SVM, KNN	Limited SVM–KNN comparison, restricted data sources, and no transformer or aspect-based analysis
Muh Nabil et al. [27]	814 user reviews from Google Play	Naïve Bayes, SVM, LDA	Limited Islamic banking studies, low use of DL, transformers, and multimodal methods, with reliance on traditional ML
Mouri Dey et al. [28]	7700 reviews	NaïveBayes, SVM,LDA	Limited multimodal studies and need for hybrid NLP methods
Rita Purnamasari et al. [29]	50,000 reviews from Google Play	VADER sentiment analysis	Exploring transformer-based models; lack of standardized criteria; expanding the geographic scope
Adiningtyas et al. [30]	6000 reviews from Google Play Store	Naïve Bayes	Limited data sources and a clear absence of transformer and multimodal approaches
Rowsan Jahan Bhuiyan et al. [31]	Social Media Reviews.	LR; RF; SVM; NB; LSTM	Limited multimodal research, advanced modeling, deployment evaluation, and interpretability
Dzaki Moch Fikri Aliffa et al. [32]	5174 reviews from Google Play Store	IndoBERT-base	Limited multimodal research-advanced modeling-deployment evaluation-and interpretability

©2026 The Author(s). This article is licensed under CC By 4.0

Raghad Alawaji et al. [33]	4922 reviews from Google Play Store	ML, Ensemble: Voting Classifier, LLMs	Limited Arabic studies and datasets, low LLM adoption, and lack of multimodal analysis and XAI.
Md. Mahmood Ali et al. [34]	10,298 reviews from Google Play Store	(SVM), (RF)	No true multimodal analysis, limited Transformers/BERT use, no standard datasets, weak XAI, and limited real-world, multilingual, and multi-source integration
Salah Al-Hagree et al. [35]	3,147 reviews from the Google Play Store.	NB; KNN; SVM; RF; ChatGPT; Bard	limited Arabic studies- underused LLMs-no hybrid ML–LLM human-in-the-loop- and no standard Arabic labeling
Muhammad Alif Ilmansyah et al. [36]	9,272 user reviews from the Google Play Store and Twitter (X)	Naïve Bayes; LSTM	Weak class imbalance handling, limited ML vs DL comparisons, and overreliance on accuracy without macro-level metrics
Naufal Rezky Ananda et al. [37]	40,000 customer text conversations	GCN + Chebyshev-GCN + Transformer-GCN + XGBoost + RF + LGBM	Limited graph-based methods, few GNN studies in banking, weak word-relation modeling, unstable GNN results, and a need for deeper contextual models
Sigit Mulyanto et al. [38]	5,000 reviews from Google Play Store	LSTM	Limited DL (LSTM), reliance on fintech/e-wallet data, lack of standard datasets, weak XAI use, and no true multimodal data
Murat Kilinc et al. [39]	242,504 reviews from Google Play Store	LSTM vs SNaive	Google Play only, short-term forecasting, and only LSTM vs. SNaive
Goesniawan Xena Adikara et al.[40]	100,000 reviews from Google Play Store.	LDA, TextBlob	Transformer models, human validation, multi-platform data, comparative banking studies
Mulyani Karmagatri et al. [41]	Twitter dataset (2,390 tweets)	Naïve Bayes	Lack of advanced DL/Transformer models and multimodal sentiment analysis
Yavuz Selim BALCIOĞLU et al. [42]	139,994 reviews from the Google Play Store	Random Forest, GBM, BERT, LDA	Need cross-country studies, longitudinal analysis, multimodal data, behavioral analytics, churn prediction, AI feature evaluation
Uttam Kotadiya et al. [43]	65,000 data records	NB, SVM, BERT, LSTM, RF, VADER	Need for multimodal data, cross-country validation, stronger ethics, Transformer/LLM use, longitudinal analysis, churn prediction integration, and behavioral-sentiment fusion
Irnayanti Dwi Kusuma et al. [44]	4,889 tweets from Twitter	BiLSTMCNN	Need for Transformers/LLMs, multimodal analysis, balanced datasets, and cross-country context-aware models
Suleiman et al. [45]	160,890 reviews from various platforms	Text Blob + LDA	Need for Transformers/LLMs, multimodal analysis, longitudinal tracking, and cross-platform sentiment studies
Musdalifa Putri Casanova et al. [46]	22,923 reviews from Google Play Store	SVM, CNN, IndoBERT	Limited use of the KDD framework and pattern interpretation
Mahmutovi ć et al. [47]	56,074 reviews from the Google Play Store	mBERT, BERTopic + LR	multimodal UX modeling absence, weak sentiment integration, lack of update/version studies, and regional research scarcity

©2026 The Author(s). This article is licensed under CC By 4.0

Vincentius Riandaru Prasetyo et al. [48]	1,800 reviews from Twitter (X),Google Play Store	NB, DT, (KNN)LR, (SVM)	Need for modern topic modeling comparison, semantic embeddings, multiclass classification, larger datasets, and improved LDA tuning
Dedy Suryadi et al. [49]	7,861 customer reviews from the Google Play Store	LDA, SenticNet 5; POS Tagging; Dependency Tree Parsing	Advanced sentiment models, emotion analysis, multi-source and cross-country datasets
Rim Mahouachi et al. [50]	4,329 reviews from the Google Play Store.	mBERT; RoBERTa, LLMs, VADER, SentiWordNet, AFINN	Limited use of sentiment analysis for product backlog generation and lack of comparisons of sentiment, topic modeling, and summarization techniques
Achmad Bahauddin et al. [51]	100 survey responses + Google Play reviews (number not specified)	SVM, SERVQUAL, QFD	Limited feedback integration, weak AI/multimodal use, and poor long-term UX evaluation
Robby Kishan Pardamean et al. [52]	10,159 reviews from Google Play Store	IndoBERT (fine-tuned + non-fine-tuned)	Limited ABSA, multimodal and Transformer/LLM use, weak neutral modeling, and poor generalization.

©2026 The Author(s). This article is licensed under CC By 4.0