TELKOMNIKA Telecommunication Computing Electronics and Control
Vol. 24, No. 4, August 2026, pp. 1224~1240
ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v24i4.27762 ❒ 1224

Sentiment analysis in telecommunications: a systematic review of applications and challenges

Achraf Bouhamidi, Naziha Laaz, Zineb Rachik
LaGeS Laboratory, Hassania School of Public Works, Casablanca, Morocco

Article Info

Article history:

Received Dec 25, 2025
Revised Apr 8, 2026
Accepted May 25, 2026

Keywords:

Customer experience analytics
Machine learning
Natural language processing
Sentiment analysis
Systematic literature review
Telecommunications industry

ABSTRACT

The increasing amount of digital content on online platforms creates both opportunities and challenges for telecommunications operators aiming to improve customer experience. Sentiment analysis (SA), a subfield of natural language processing (NLP), allows for the extraction of subjective information from data generated by users. This paper presents a systematic literature review (SLR), conducted in accordance with Kitchenham's guidelines, examining the applications and challenges of SA within the telecommunications sector. A structured methodology was used to select and analyze over 100 studies published between 2017 and 2025. The review categorizes SA methods and evaluates their application in practical telecom contexts, including customer satisfaction assessment, churn prediction, service monitoring, and reputation management. Findings indicate that machine learning and deep learning models can achieve performance levels up to 97% in specific experimental contexts. However, such results may vary depending on the datasets, evaluation protocols, and application contexts. Key challenges identified include handling informal language, domain dependency, multilingualism, sarcasm detection. Finally, the review highlights future directions such as real-time sentiment tracking, multimodal analysis, and the integration of federated learning for privacy-preserving customer analytics. This review provides a foundational reference for researchers and practitioners aiming to deploy effective sentiment-driven systems in the telecom industry.

Corresponding Author:

Achraf Bouhamidi
LaGeS Laboratory, Hassania School of Public Works
Casablanca 20230, Morocco
Email: bouhamidi.achraf.cedoc@ehtp.ac.ma

1. INTRODUCTION

Over the past two decades, the telecommunications industry has undergone significant transformation, driven by advances in broadband connectivity, mobile networks, and digital service platforms. As competition has intensified and switching costs have decreased, customer experience has emerged as a central differentiator for telecommunications operators [1]. Beyond technical network performance, operators increasingly need to understand how customers perceive service quality, pricing, support interactions, and brand reputation across multiple communication channels, including call centers, emails, chats, forums, and social media. These channels generate large volumes of unstructured and heterogeneous customer feedback, making large-scale manual analysis impractical. Sentiment analysis (SA), a subfield of natural language processing (NLP), has become an essential analytical tool [2]. By detecting sentiment polarity, intensity, and aspect-level opinions, SA enables telecom operators to identify early signs of dissatisfaction, support churn prevention strategies, prioritize customer complaints, and continuously monitor service perception. Recent studies indicate that sentiment-based insights complement traditional technical key performance indicators (KPIs) by capturing qualitative dimensions of customer behavior that are not visible through network metrics alone.

A wide range of SA techniques has been explored in the literature, including lexicon-based approaches, classical machine learning (ML) models such as support vector machines (SVM) and random forests (RF), as well as deep learning (DL) and transformer-based architectures (e.g., convolutional neural network (CNN), long short-term memory (LSTM), bidirectional encoder representations from transformers (BERT), Arabic BERT (AraBERT)). However, the application of these techniques in telecommunications presents domain-specific challenges that are not fully addressed in generic SA studies. Telecom data are often multilingual and dialect-rich, highly informal, and strongly context-dependent. In addition, customer feedback frequently involves technical jargon, sarcasm, and rapidly evolving expressions, while annotated datasets tailored to telecom use cases remain limited.

Although SA has been extensively reviewed in domains like healthcare, e-commerce, education, and finance, there is still no consolidated systematic review that focuses specifically on SA applications, challenges, and future directions within telecommunications operators. Existing surveys tend to address general-purpose SA techniques or specific linguistic settings, without offering a telecom-oriented synthesis that reflects the operational realities and constraints of this sector.
To address this gap, this paper presents a systematic literature review (SLR) of SA in the telecommunications industry. The review addresses three research questions:

- RQ1: What are the potential applications of SA for telecommunications operators?
- RQ2: What are the challenges facing telecommunications operators in the field of SA?
- RQ3: What future research directions and technological trends are emerging?

Following a rigorous review protocol, more than 100 peer-reviewed studies published between 2017 and 2025 were analyzed across five major scientific databases, including IEEE Xplore, ACM Digital Library, Scopus, ScienceDirect, and SpringerLink. The contributions of this paper are threefold. First, it provides a structured synthesis of SA applications in telecom-specific scenarios such as customer service improvement, churn prediction, service quality monitoring, and reputation management. Second, it identifies and analyzes the key linguistic, methodological, and operational challenges that limit the robustness and generalizability of existing approaches. Third, it outlines a forward-looking research roadmap, highlighting emerging directions including large language models (LLM), real-time and multimodal SA, and privacy-preserving learning frameworks.

By adopting a telecommunications-centered perspective, this review aims to support both researchers and practitioners. For researchers, it clarifies open challenges and opportunities for domain-specific NLP and artificial intelligence (AI) in telecom environments. For industry professionals, it provides insights into how SA can be integrated into customer analytics and information and communication technology (ICT) systems to support data-driven decision-making.

The remainder of this paper is organized as follows. Section 2 introduces the background concepts relevant to SA. Section 3 describes the methodology adopted for the systematic review. Section 4 presents and analyzes the results of the review. Section 5 discusses SA applications in telecommunications contexts. Section 6 examines challenges and future research directions. Finally, section 7 concludes the paper.

2. BACKGROUND

SA also known as opinion mining, is a subfield of NLP that is concerned with extracting, identifying, and classifying subjective information from text. In the telecommunications context, SA enables operators to better understand customer opinions expressed through various channels such as tweets, emails, chat logs, and service reviews. To support real-time decision making and customer experience management, SA models need to operate across multiple granularity levels, address nuanced emotional expressions, and adapt to domain-specific language.

2.1. Levels of SA

SA can be performed at various levels of granularity, depending on the scope of analysis: document level, sentence level, and aspect (or entity) level. Document-level SA focuses on classifying the overall sentiment of an entire document (e.g., a full customer email or long complaint form) as positive, negative, or neutral. It assumes that the document centers around a single entity or service, which is not always valid in telecom datasets, where customers may discuss multiple issues in one text. Sentence-level analysis operates on individual sentences, distinguishing between subjective and objective expressions, and then assigning a sentiment polarity. This level is especially relevant in telecom environments, where short, fragmented expressions (e.g., social media posts) dominate customer feedback. Aspect-level (or entity-level) SA provides fine-grained analysis by identifying sentiments associated with specific entities or service attributes (e.g., "internet speed" or "billing process"). This level is particularly useful for telecom operators to pinpoint service dimensions that drive customer satisfaction or dissatisfaction.

2.2. SA tasks

This section outlines the main SA tasks, with a focus on their relevance to the telecommunications domain.

- Subjectivity classification: this task identifies whether a text expresses subjective opinions or objective information. In telecommunications, it is particularly useful for filtering customer feedback from factual service reports.
- Sentiment classification: once subjective content is identified, this task determines the polarity of opinions (e.g., positive or negative). It is widely applied in telecom to assess customer satisfaction from reviews, surveys, and social media data.
- Emotion detection: beyond polarity, this task is not limited to the identification of primary psychological states (joy, sadness, and anger); it tends to reach 6 or 8 scales depending on the emotion model [3]. Emotion detection is gaining importance in telecom support, where understanding the emotional tone can aid chatbot responses.
- Aspect extraction: this task identifies specific service features or components mentioned by users (e.g., network quality, pricing, or customer service), enabling fine-grained analysis of user feedback.
- Aspect-based SA (ABSA): ABSA determines sentiment at the level of specific aspects or attributes of an entity. In telecommunications, it enables operators to evaluate customer opinions on particular services or features, even when aspects are implicitly expressed [4].
- Sentiment intensity analysis: this task measures the strength of expressed opinions, which is essential in telecom scenarios to distinguish between mild dissatisfaction and critical service issues. The degree or level of emotions and feelings often plays a crucial role in understanding the exact feeling within the same class [5].
- Sentiment summarization: this task aggregates opinions from multiple sources to provide a concise overview of customer sentiment, supporting decision-making processes in telecom service management. This approach produces a summary with the relevant information and the orientation of the feeling according to the user request [6].

2.3. Methods for SA

Existing sentiment classification approaches or feeling analysis methods are divided into three categories: lexicon-based methods, ML methods, and hybrid methods. Figure 1 represents these categories.

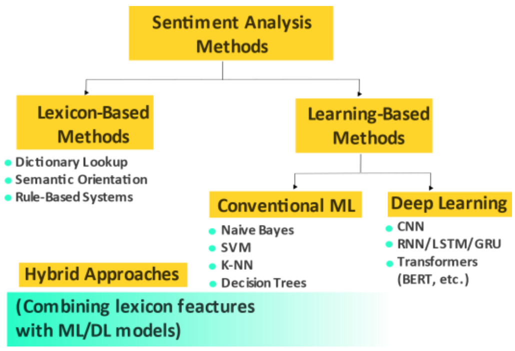

Figure 1. SA methods

2.3.1. Lexicon-based method

Lexicon-based methods rely on predefined sentiment dictionaries in which words or expressions are associated with predefined polarity labels, such as positive or negative. Commonly used lexicons include linguistic inquiry and word count (LIWC), SentiWordNet, and valence aware dictionary and sentiment reasoner (VADER), which have been widely applied in early SA studies and social media analytics. These approaches are generally characterized by their simplicity, interpretability, and ease of implementation [7].

However, in telecommunications-related applications, lexicon-based methods face notable limitations. They are often sensitive to domain dependency, informal language usage, and linguistic variations, and they struggle to capture contextual nuances such as sarcasm or aspect-level sentiment. Moreover, their effectiveness tends to decrease when applied to large-scale or multilingual datasets, which are common in real-world telecom environments. As a result, lexicon-based approaches are increasingly complemented or replaced by ML and DL techniques in recent studies. Consequently, recent research has increasingly focused on ML and DL approaches to overcome these limitations and improve SA performance in complex telecommunications settings.

2.3.2. ML and DL methods

ML techniques have been widely adopted in SA to overcome the limitations of lexicon-based approaches, particularly in handling large-scale and data-driven applications. In the telecommunications domain, supervised learning algorithms such as SVM, Naive Bayes (NB), decision trees (DT), and RF have been extensively used due to their robustness, interpretability, and relatively low computational cost. These methods typically rely on feature engineering techniques to transform textual data into numerical representations [8].

More recently, DL approaches have gained increasing attention for SA in telecommunications, as they enable automatic representation learning from raw textual data. Models based on convolutional and recurrent neural networks (RNN), as well as more advanced architectures, have demonstrated improved performance in capturing complex linguistic patterns, particularly when applied to large-scale and multilingual datasets. However, these models often require substantial training data and computational resources, which may limit their deployment in certain operational settings [9].

Overall, existing studies suggest that traditional ML methods remain competitive for domain-specific and resource-constrained telecom applications, while DL approaches are better suited for complex SA tasks involving high data variability. This evolution highlights a trade-off between model complexity, interpretability, and deployment constraints in real-world telecommunications environments. More recent advances, such as transformer-based and LLM, are discussed as emerging directions rather than core methods, given their limited adoption in current telecom-focused studies.

3. METHOD

This study follows the SLR guidelines proposed by Kitchenham to ensure transparency, reproducibility, and objectivity in the research process. The review was carried out in three main phases: (i) planning the review, (ii) conducting the review, and (iii) reporting the findings, as illustrated in Figure 2.

Figure 2. Phases of the SLR and its activities

3.1. Planning the review

During this planning phase, the objectives of the review were clearly identified in connection with the following events, which describe each step-in detail.

3.1.1. Identifying the need for the review

Although SA has been widely studied in various sectors, its role in telecommunications, particularly in relation to customer experience, remains fragmented in the literature. Existing reviews often focus on general NLP methods, specific languages, or technical model comparisons without addressing the domain-specific challenges faced by telecom operators. This paper presents an SLR of existing studies related to the field of SA in the telecommunications sector. We focus on high-level trends and patterns rather than methodological details, which have been well covered by existing studies. We aim to benefit both newcomers to the field of SA by presenting the basics of the topic and researchers by sharing insights and results that are useful to the community.

3.1.2. Formulating the research questions

Based on the objectives of the review, the following research questions were defined:
- RQ1: What are the possible applications of SA for telecommunications operators?
- RQ2: What are the challenges facing telecommunications operators in the field of SA?
- RQ3: What future research directions and technological trends are emerging?

3.1.3. Selecting data sources

To ensure comprehensive coverage of peer-reviewed research, five major scientific digital libraries were selected: IEEE Xplore, ACM Digital Library, ScienceDirect, SpringerLink, and Scopus. These databases collectively cover journals and conference proceedings in telecommunications, computer science, AI, and NLP. Grey literature and industry reports were not included to maintain methodological consistency and ensure that all selected studies met peer-reviewed academic standards. The literature search covered publications published between 2017 and 2025.

3.2. Conducting the review

The search strategy was defined using the population, intervention, comparison, outcome, context (PICOC) framework. The following query illustrates the general structure used across databases. Search query (example): (“sentiment analysis”) AND (“telecommunication” OR “telecom” OR “mobile network”) AND (“customer experience” OR “CX” OR “customer journey”). The query syntax was adjusted based on the capabilities of each digital library.

3.2.1. Inclusion and exclusion criteria

A three-stage filtering process was applied: (i) initial screening based on title and abstract relevance, (ii) full-text assessment using predefined quality criteria, and (iii) final selection based on relevance to the research questions. Studies were included if they:
- Addressed SA in a telecommunications context,
- Covered at least one of the three dimensions (application, challenge, or future direction),
- Established a link between SA and customer satisfaction or experience.

Studies were excluded if they were not telecom-related, did not address customer-facing aspects, were not written in English, or had inaccessible full texts. Quality assessment was conducted at the full-text stage using mandatory evaluation criteria adapted from SLR best practices. Papers were required to clearly state research objectives, describe related work, explain the methodology and report results relevant to the stated objectives. Studies not meeting at least one mandatory criterion were excluded. Tables 1 and 2 provide summaries of the evaluation criteria and study selection outcomes, respectively.

Table 1. Selection of articles

No	Questions	Inclusion criteria	Exclusion criteria
1	Does this paper deal with the field of telecommunications?	This paper refers to the field of telecommunications (or a similar field)	No mention
2	Does this paper discuss the applications or challenges or future direction of SA?	This paper proposes answers to one of the 3 dimensions	No mention
3	Does this paper establish a connection between the application of SA and customer satisfaction?	This paper establishes this connection	No mention

Table 2. Study selection procedure

Database	Preliminary search result	Screened result	Selected studies in line with inclusion/exclusion criteria
ACM Digital Library	564	46	17
IEEE Xplore	509	41	26
ScienceDirect	1077	34	23
Scopus	1965	77	39
Springer Link	685	145	34

3.2.2. Study selection results

The initial search retrieved 4,800+ records across the five databases. After duplicate removal and abstract screening, 343 studies were retained for full-text assessment. Following quality evaluation and relevance checks, a final set of 139 studies was selected for analysis. Table 2 presents the detailed breakdown by database and selection stage.

3.2.3. Data extraction and synthesis

From each selected study, the following information was systematically extracted: application domain (e.g., churn prediction, service quality monitoring), dominant methods (lexicon-based, ML, DL, hybrid), dataset characteristics (size, source, language), reported performance metrics, identified challenges and limitations, and proposed future research directions. An iterative thematic coding approach was used to categorize applications and challenges. Initial codes were derived inductively from study objectives and refined through successive comparison, leading to higher-level groupings used in the results and discussion sections. To improve consistency, study classification and thematic groupings were cross-checked by the authors, reducing subjectivity and ensuring coherent categorization.

3.2.4. Reporting the findings

The extracted data was analyzed and synthesized to provide structured answers to the three research questions. A visual representation of the distribution of the articles over time and by publication type was included (Figure 3), along with tabular summaries (Table 3) of algorithms used, data characteristics and key insights.

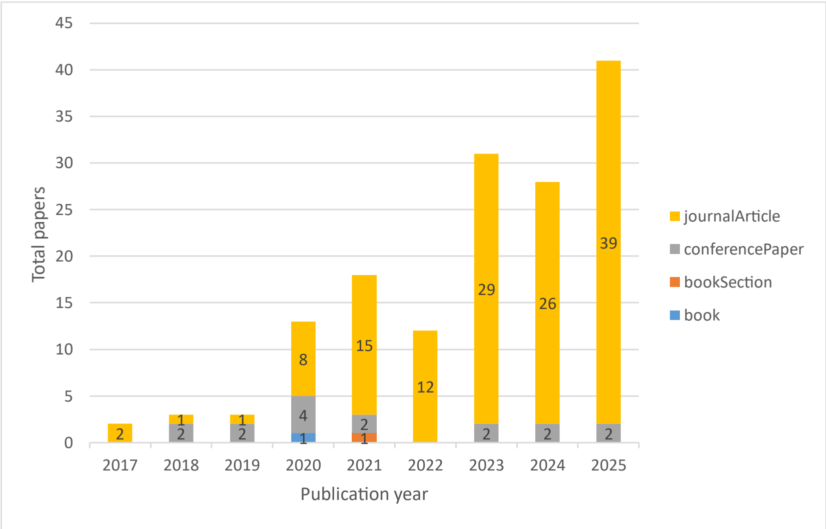

Figure 3. Distribution of the extracted articles from 2017 to 2025

4. RESULTS AND DISCUSSION

The reviewed studies show that SA has progressively become a core analytical tool in the telecommunications domain, supporting a wide range of applications such as customer satisfaction assessment, churn prediction, service quality monitoring, and reputation analysis. Initial contributions largely relied on lexicon-based approaches applied to social media data, mainly to track public sentiment toward operators and services [10], [11]. While these methods enabled large-scale sentiment monitoring and temporal or regional trend visualization, several studies reported important limitations, including domain dependency, difficulties in handling neutral or ambiguous expressions, and weak contextual sensitivity in informal and multilingual settings [12]–[15]. In parallel, SA was increasingly combined with clustering and predictive models to address more advanced use cases, notably churn prediction and customer segmentation. Several studies demonstrated that integrating sentiment features with behavioral or usage data can significantly enhance predictive accuracy and improve customer profiling [16]–[18]. However, despite the frequent reporting of high accuracy values, these results remain difficult to compare directly, as they are obtained under heterogeneous experimental settings, with different datasets, annotation strategies, and evaluation metrics.

More recent work reflects a clear shift toward DL approaches, motivated by the growing linguistic complexity of telecom-related data. Convolutional and RNN architectures, as well as hybrid models, have been shown to better capture contextual information, dialectal variations, and aspect-level sentiment, particularly in large-scale and Arabic-language datasets [19]–[23]. At the same time, several authors point out that these performance gains come with practical trade-offs, including increased computational cost, longer training times, and a higher risk of overfitting when labeled data are limited [9].

Beyond sentiment classification itself, a number of studies extend SA toward predictive and proactive applications. Research on customer growth forecasting, proactive service recovery, and early churn detection illustrates how sentiment dynamics and emotional signals can complement traditional operational indicators [24]–[27]. Taken together, these contributions suggest an ongoing evolution from descriptive sentiment monitoring toward decision-oriented and predictive sentiment analytics in telecommunications.

Most reviewed studies rely on short-term or dataset-specific evaluations, with limited analysis of long-term performance, scalability, or model degradation over time. This limits the assessment of real-world deployment sustainability in large-scale telecom environments. To avoid repetitive study-by-study descriptions, the reviewed works were grouped according to their primary application objectives, and their main characteristics and findings are synthesized in Table 3. This table highlights several cross-cutting patterns. First, customer service improvement and churn-related applications dominate the literature, reflecting the operational priorities of telecommunications operators. Second, classical ML methods remain widely used for structured or moderately sized datasets, while DL and transformer-based approaches are primarily adopted in settings involving large-scale, multilingual, or unstructured data. Finally, Table 3 shows that social media platforms and customer-support channels constitute the most common data sources, underscoring the importance of handling noisy and heterogeneous user-generated content in telecom SA.

Table 3. Synthesized view of SA applications in telecommunications

Application objective	Representative studies	Typical data sources	Dominant methods	Key insights
Customer satisfaction and experience analysis	[14], [22], [28], [29]	Social media posts (Twitter, Facebook), app reviews (Google Play), online comments	ML (SVM, RF, NB), DL (CNN, LSTM, gated recurrent unit/GRU)	High classification performance on domain-specific datasets; effectiveness strongly influenced by language, annotation quality, and service granularity
Churn prediction and retention	[16], [18], [26], [27]	Customer relationship management (CRM) records, call detail records (CDR), user-generated content (UGC), call metadata	ML (RF, LR, AdaBoost), DL (RNN, CNN), Hybrid	Sentimental features significantly enhance churn prediction when combined with behavioral and usage data
Reputation and brand monitoring	[10], [15]	Twitter streams, online forums, word-of-mouth platforms	Lexicon-based (SWN, LIWC, VADER), ML (SVM)	Suitable for large-scale trend monitoring; limited contextual accuracy in informal and multilingual settings
Service quality and network monitoring	[11], [23], [30]	Customer complaints, Twitter posts, service-related feedback	ML (SVM, RF), DL (CNN-BLSTM)	Effective detecting coverage and QoS issues; supports near real-time service monitoring
Telecom-specific linguistic resources	[21], [25], [31]	Annotated telecom corpora (AraCust, Sara-dataset), dialectal tweets	Lexicon construction, DL (AraBERT, CNN, LSTM)	Domain-specific lexicons and corpora substantially improve sentiment classification reliability
Proactive and predictive analytics	[25]–[27]	Social media streams, customer emails, behavioral metadata	DL (RNN, CNN), Hybrid (SA + ML)	Shift from descriptive sentiment monitoring toward anticipatory and decision-oriented analytics

5. Applications of SA in telecommunications

Based on the reviewed studies, SA applications in telecommunications can be structured into four major application domains, each characterized by distinct data sources, analytical objectives, and methodological requirements. Rather than comparing raw performance figures, the following synthesis highlights how reported results vary depending on data characteristics, task complexity, and deployment context.

5.1. Customer support and customer service improvement

A large body of work focuses on improving customer support and service quality by analyzing customer-generated content from social media, forums, and direct communication channels. In this domain, SA is primarily applied at the document or aspect level to identify dissatisfaction related to specific services such as billing, network performance, or customer support interactions [11], [22], [25], [28]. Studies in this group frequently report high classification accuracy; however, these results are often obtained on domain-specific and moderately sized datasets with predefined service categories.

From a methodological perspective, traditional ML models and supervised DL approaches perform well in this setting because customer service data tends to exhibit relatively stable sentiment patterns tied to explicit service aspects. Aspect-based SA, as illustrated in [11], [22], further improves interpretability by linking sentiment scores to concrete operational dimensions. Overall, the effectiveness of SA in customer service improvement depends less on model complexity than on the granularity of service definitions, annotation quality, and the ability to integrate sentiment outputs into operational workflows.

5.2. Churn prediction and customer retention

Churn prediction represents a more complex application of SA, as it involves linking emotional signals to future customer behavior. Traditional churn prediction models rely mainly on network performance metrics, pricing, and usage data [16], [17]. When SA is introduced as an additional signal, typically extracted from social media or online forums, it provides complementary information reflecting customer dissatisfaction before contract termination [10].

In this domain, reported performance improvements are highly dependent on how sentiment features are combined with behavioral and demographic data. ML models such as RF and ensemble methods tend to perform well when sentiment indicators are integrated with structured telecom data [16]–[18]. However, SA alone is generally insufficient for reliable churn prediction, as emotional expressions vary widely across users and contexts. This explains why accuracy figures reported in churn-related studies are less stable and more context-dependent than those observed in customer service classification tasks.

5.3. Social media and service quality monitoring

SA is also widely used for monitoring customer reactions on social networks, particularly during critical events such as network outages or service degradations. In this context, the primary objective is not fine-grained sentiment classification but the rapid detection of abnormal negative sentiment patterns that may indicate service quality issues [23]. Studies in this group emphasize real-time or near real-time processing, often favoring robustness and responsiveness over marginal gains in accuracy. As a result, both lexicon-based and ML approaches are employed, depending on scalability and deployment constraints. Performance metrics reported in these studies should therefore be interpreted cautiously, as they reflect system responsiveness and alerting effectiveness rather than standalone sentiment classification accuracy.

5.4. Reputation and competitive monitoring (G3)

Reputation monitoring focuses on tracking public perception of telecommunications operators and comparing sentiment trends across competitors. In this domain, SA is typically applied at an aggregate level to identify long-term trends rather than individual customer issues [10], [15]. Lexicon-based and lightweight ML approaches remain common due to their scalability and interpretability. However, these methods are particularly sensitive to informal language, sarcasm, and multilingual content, which limits the reliability of fine-grained sentiment interpretation. Consequently, reported accuracy values in reputation monitoring studies are less indicative of model superiority than of their suitability for large-scale trend analysis and benchmarking.

Across all application domains, reported accuracy values ranging from 82% to 97% cannot be directly compared without considering dataset size, language, annotation strategy, and task definition. Customer service and satisfaction studies tend to report higher and more stable performance due to well-defined sentiment targets, whereas churn prediction and reputation monitoring involve higher uncertainty and contextual variability. This synthesis highlights that the effectiveness of SA methods in telecommunications is strongly application-dependent, and that methodological choices should be guided by deployment objectives rather than absolute performance metrics.

6. MAIN CONCLUSIONS AND TRENDS

The telecommunications sector is under rapid transformation driven by the proliferation of digital technologies and the advent of broadband (fiber and 5G/6G). Understanding user sentiment is important for telecom operators to improve the quality of customer experience and maintain customer satisfaction and loyalty. In the field of natural NLP, SA is a crucial discipline that serves as a pillar for decoding these nuanced emotions and opinions [32]. We summarize the conclusions of this study in Figure 4.

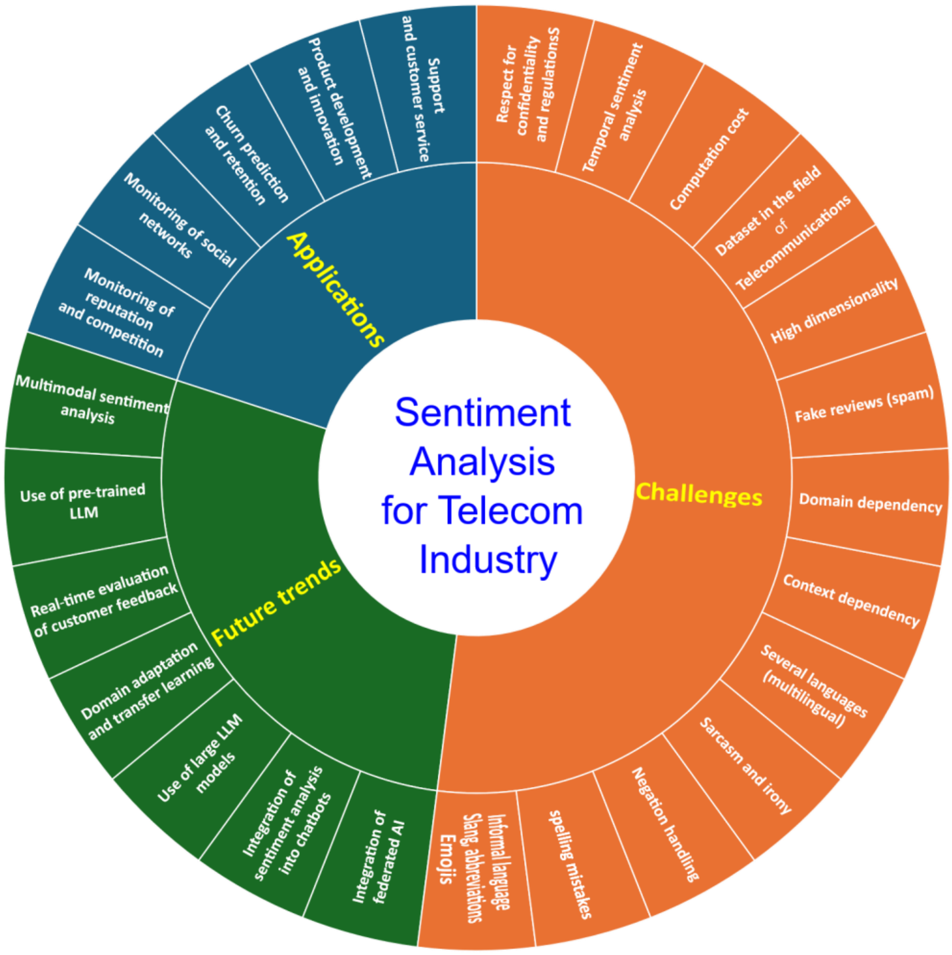

Figure 4. Applications, challenges and future directions of SA in the telecommunications sector

6.1. Discussions

This discussion aims to interpret the variability of SA performance reported in telecommunications studies, rather than to establish direct algorithmic comparisons. In practice, model effectiveness is strongly influenced by data characteristics, including dataset size, linguistic complexity, domain specificity, and task granularity. Telecom-related data range from clean and structured customer surveys to noisy, multilingual, and aspect-dense social media and call-center content, leading to substantial differences in achievable performance. Understanding these contextual factors is essential for explaining why certain methods perform well under specific conditions and for guiding appropriate methodological choices. This analysis therefore provides the conceptual foundation for the challenges discussed in section 6.3.

Telecommunication data present a wide spectrum of linguistic scenarios, from structured satisfaction surveys and support-ticket logs (low-noise, homogeneous) to noisy social media posts and call-center transcripts (highly informal, dialectal, and aspect-dense). Consequently, the efficacy of the model varies substantially with the size of the labeled dataset, the dialect and domain coverage, and the granularity of the task. The distinction is especially relevant because telecom data (like customer emails, billing complaints, surveys, tweets, outage reports, chat transcripts) vary widely in volume and noise level, leading to significant differences in model performance across contexts. Recent evidence in telecom SA studies supports a pragmatic, data-driven guideline for model choice.

6.1.1. Data-driven guideline for model choice

For small, clean and homogeneous telecom corpora (typically under a few thousand labeled documents), classical classifiers such as SVM and RF, when combined with the features of term frequency–inverse document frequency (TF-IDF) and domain lexicon, remain strong baselines due to low variance and robustness to limited data [1], [2]. As the scale of the labeled data and linguistic complexity increases (dialectal Arabic, code-switching, domain jargon), transformer-based encoders (BERT family and domain/dialect-adapted variants) generally outperform classical methods, and when models are adapted to telecom/dialectal distributions [1]. Telecom feedback is naturally aspect-oriented (coverage, billing, speed, support), and ABSA demands both accurate aspect detection and correct polarity assignment per aspect. The ABSA literature shows the best results from hybrid approaches that combine telecom specific lexicons (to identify aspect terms and seed polarity) with contextual encoders (to resolve compositional and contextual polarity), producing more robust per aspect predictions than lexicon only or transformer only pipelines. Lexicons offer reliable anchors and polarity priors, while contextual encoders disambiguate modifiers, handle negation and sarcasm, and disambiguate multi-aspect sentences. These findings recommend a tiered strategy: use SVM/RF for small, curated datasets with limited linguistic variation; adopt transformers with dialect/domain adaptation when data permit; and employ lexicon enhanced transformers for ABSA tasks in telecom scenarios involving multi-aspect customer feedback. This differentiation is crucial for researchers and practitioners designing sentiment pipelines for telecom applications, ensuring that methodological choices are aligned with data characteristics and task complexity.

6.1.2. Limitations in current modeling approaches

Despite advances in transformer-based models, several failure modes remain persistent in telecom SA. First, implicit sentiment driven by sarcasm, irony, and emoji continue to challenge classifiers, as sentiment polarity is often inverted while surface lexical cues remain positive, leading to systematic misclassifications even for large pretrained transformers [1]. Emojis introduce additional ambiguity, as they carry sentiment signals that may override or contradict textual content. Second, code-switching common in customer complaints, mixing French/Arabic or Darija/English reduces embedding stability and produces tokenization errors, causing degraded sentiment predictions unless domain-dialect adaptation is employed. Empirical findings show that sentiment performance decreases sharply on code-switched inputs unless models are pretrained or adapted in multilingual telecom corpora [2]. Domain-specialized or dialectal BERT variants mitigate, but do not fully eliminate, this challenge. Third, models trained in one telecom operator’s corpus transfer poorly to another due to operator-specific terminology, complaint patterns, and product vocabularies, resulting in significant performance drops between domains. Several studies report substantial drops (10-25 F1 points) in cross-domain transfer, even within the same country or language family. This indicates strong domain overfitting and highlights the need for domain adaptation, multi-operator training corpora, or continual learning frameworks. Finally, sentiment in the telecommunications sector is highly sensitive to time-bound events, such as nationwide outages, network upgrades, regulatory announcements, or promotional campaigns. These events produce rapid changes in sentiment polarity, vocabulary distribution, and frequency of complaints.

Telecom datasets exhibit strong temporal drift, with sentiment distributions changing during network outages, billing incidents, or promotional campaigns. Studies in customer-experience analytics show that the accuracy of the model decays significantly after outages or price changes unless periodic retraining or temporal calibration is applied. This drift phenomenon demonstrates that telecom sentiment models must incorporate time-aware components or adaptive retraining schedules. These failure modes reveal structural limitations in current modeling approaches: insufficient sarcasm recognition, incomplete multilingual robustness, narrow domain generalization, and poor adaptability to temporal shocks. Addressing these issues requires techniques such as emoji-aware embeddings, sarcasm-specific pretraining, cross-operator domain adaptation, code-switched language modeling, and continual learning strategies to maintain performance in dynamic telecom contexts. As telecom operators move toward data-driven customer-experience management, SA systems must align with operational, regulatory, and real-time constraints. These constraints have direct implications for how models are trained, deployed, monitored, and evaluated. As customer-experience data are distributed across multiple systems (CRM, digital applications, call-center transcripts, network operations center (NOC) logs) and often contain personally identifiable information, label governance and privacy regulations (national telecom laws) limit centralized data aggregation. Recent research highlights federated learning as a promising approach for telecom contexts, allowing distributed fine-tuning where model updates are shared but customer data remain local [1].

In parallel, telecom operations increasingly require real-time processing, integrating sentiment signals into CRM prioritization workflows and NOC dashboards to detect emerging dissatisfaction during outages or service degradations [3]. As a result, evaluation metrics must expand beyond accuracy, incorporating operational indicators such as time-to-action, alert latency, and churn-savings proxies, which better capture the business impact of sentiment intelligence in telecom environments [4]. These implications underscore the need for architectures that are privacy-preserving, operator-aligned, and operationally measurable. Real-time inference requires lightweight or distilled transformer models, optimized feature stores, streaming ingestion (Kafka/Kinesis), and online retraining to adapt to temporal drift. Together, these strategies operationalize SA as a decision-support engine that reduces churn, accelerates incident response, and improves customer experience.

6.2. Performance of the models

A qualitative subgroup analysis across the reviewed studies highlights clear patterns. Studies based on English-language datasets generally report higher and more stable performance, benefiting from cleaner corpora and richer pretrained resources. In contrast, Arabic and dialectal datasets exhibit lower baseline performance due to informal language, code-switching, and limited annotated data, with transformer-based models such as AraBERT and recurrent architectures showing more consistent gains over classical methods. Similarly, classical models such as SVM and RF remain competitive on small or moderately sized datasets, while DL and transformer-based approaches demonstrate advantages primarily in large-scale or linguistically complex settings.

The performance of lexicon/rule-based models generally falls within a range of 55 to 85%. Classical learning models, on the other hand, generally show performance between 55% and 90%, while DL models tend to achieve higher performance, between 70% and 95%. These ranges are estimates and may be higher or lower depending on various factors, including the quality of the data sets, the chosen model architecture, the pre-processing techniques used, and the quality and coverage of the lexicon used [33]. These ranges should be interpreted as contextual indicators rather than pooled benchmarks, as they depend strongly on dataset size, language, annotation strategy, and task formulation.

In the study by Najadat *et al*. [13], SVM performed better for the classification of Jordanian Facebook posts from the three Jordanian operators with an accuracy of 95.77% compared to NB, DT and K-nearest neighbors (KNN) which have accuracy of 78.5%, 91.98% and 94.37% respectively. In the study by Nahar *et al*. [14], the SVM model performed better than the KNN and NB models with 97.8%, 96.8% and 95.6% accuracy respectively. The study in Abdullah *et al*. [15] also favored the SVM model (95% accuracy) over the DT models (68%) and multinomial naive bayes (MNB) (87%). In the study by Aftan and Shah [19], AraBERT obtained the best prediction accuracy of 94.33% compared to other DL models CNN and RNN. Alshamari [20] gave the advantage to LSTM, which obtained the best learning accuracy (97.03%) in text classification compared to GRU, BiLSTM and CNN-LSTM models. Almutairi and Alotaibi [21], the CNN model had the best accuracy 96.83% against a rate of 94.97% for LSTM. The ensemble model (CNN, LSTM) had an accuracy of 95.81%. Abubaera and Jiddah [34] demonstrated the potential of data augmentation to advance SA for Arabic text data and to solve the problem of under-sampled positive data samples. An improvement of almost 10% in accuracy was observed in a CNN-GRU model before and after the application of this technique. In the study [11], the combined SVM model assisted by SentiCircle produces an F1-measure performance of 96.3%.

These performance differences can be largely explained by model characteristics and data conditions. Classical classifiers benefit from low variance and robustness when training data are limited, making them effective baselines in constrained settings. In contrast, DL and transformer-based models leverage contextual and sequential representations that allow them to better capture linguistic complexity, long-range dependencies, and subtle sentiment cues, albeit at the cost of higher data and computational requirements. Several studies also report relatively low performance, particularly in scenarios involving highly noisy social media data, sarcasm, or strong domain shifts. In such cases, lexicon-based approaches and shallow classifiers often fail to generalize, highlighting the limitations of static sentiment resources and the risks of overfitting in domain-specific datasets.

Overall, the reported performance ranges of lexicon-based, ML, and DL models should not be interpreted as direct benchmarks. The observed differences primarily reflect heterogeneity in datasets, languages, annotation strategies, and application contexts rather than intrinsic model superiority. In telecommunications, simpler models often remain competitive on small or well-curated datasets, while more complex architectures demonstrate advantages when linguistic variability and data scale increase. These findings highlight the need for application-aware evaluation and motivate a deeper examination of the structural challenges that limit robustness and comparability, which are discussed in the following section.

6.3. Challenges of SA in telecommunications

This section addresses RQ2 by synthesizing the main challenges that limit the effectiveness, robustness, and deployment of SA in the telecommunications sector. While many of these challenges are shared with other NLP domains, their impact is amplified in telecommunications due to data heterogeneity, linguistic diversity, and operational constraints [32].

6.3.1. Data-related and linguistic challenges

Telecom-related data exhibit strong variability, ranging from structured satisfaction surveys and customer-support tickets to highly informal social media posts and call-center interactions. Informal language, including slang, abbreviations, acronyms, spelling variations, and emojis, remains a major source of sentiment misclassification, particularly in user-generated content [35], [36]. Although some universal acronyms and patterns can be learned, many regional or evolving expressions remain difficult to capture, reducing model robustness over time.

Multilingualism and code-switching further complicate sentiment interpretation, especially in regions where customers frequently mix languages or dialects within the same message [37]. Empirical studies show that sentiment classification performance degrades significantly on code-switched inputs unless models are adapted using multilingual or dialect-specific corpora [38]. These linguistic challenges partly explain the wide dispersion of reported performance values across studies and highlight why accuracy figures must be interpreted in relation to language and data characteristics.

6.3.2. Contextual understanding and implicit sentiment

Telecommunications sentiment is often implicit rather than explicit, particularly in cases involving sarcasm, irony, or emotionally charged emojis. In such situations, surface lexical cues may contradict the underlying sentiment, leading to systematic errors even for large pretrained models [35], [36]. Negation handling represents a closely related difficulty, as sentiment polarity may be inverted implicitly and across long syntactic spans. Several studies have demonstrated that integrating sarcasm-aware components or joint sentiment–sarcasm modeling can improve classification robustness. For instance, Bi-LSTM-based frameworks that jointly model sentiment and sarcasm have shown measurable performance gains compared to sentiment-only pipelines [39]. Nevertheless, sarcasm detection remains an open challenge due to its reliance on contextual, pragmatic, and cultural cues.

6.3.3. Domain dependency and generalization limits

Another critical challenge concerns domain dependency and cross-domain generalization. Sentiment models trained on data from one telecommunications operator often perform poorly when transferred to another operator, even within the same language or geographical region. This limitation is largely due to differences in product portfolios, complaint patterns, and operator-specific terminology. Cross-domain SA studies report significant performance drops when models are applied outside their training domain, underscoring the difficulty of knowledge transfer in specialized sectors [40]. Although domain adaptation and transfer learning techniques partially mitigate this issue, robust multi-operator generalization remains an unresolved problem in telecom SA.

6.3.4. Temporal dynamics and concept drift

Sentiment in the telecommunications sector is sensitive to time-dependent events, such as network outages, billing incidents, promotional campaigns, or regulatory announcements. These events introduce abrupt shifts in vocabulary usage, sentiment polarity, and complaint frequency. As a result, models trained on historical data may experience rapid performance degradation when applied to new temporal contexts. Temporal SA studies show that sentiment distributions evolve significantly over time and that static models fail to capture these dynamics without periodic retraining or adaptive calibration [33]. This phenomenon of concept drift highlights the need for time-aware modeling strategies in operational telecom environments.

6.3.5. Operational, privacy, and evaluation constraints

Beyond linguistic and modeling challenges, SA in telecommunications is constrained by operational and regulatory requirements. Customer-experience data are distributed across multiple systems, including CRM platforms, digital applications, call-center transcripts, and network operation logs. These data often contain sensitive personal information, and strict privacy regulations limit centralized data aggregation.

Recent work highlights federated learning as a promising approach for telecom contexts, enabling distributed model training while preserving data locality and privacy [1]. At the same time, telecom applications increasingly require real-time or near-real-time SA, particularly for outage detection and customer-support prioritization. Consequently, evaluation metrics must extend beyond traditional accuracy or F1-score to include operational indicators such as alert latency, time-to-action, and business impact measures.

The mitigation strategies outlined in the future research direction (such as domain adaptation, sarcasm-aware modeling, multilingual embeddings, and adaptive learning) directly address the limitations and failure cases identified in the reviewed studies. Overall, the challenges identified in this review demonstrate that limitations in telecom SA are not purely algorithmic. Instead, they arise from the interaction between linguistic complexity, domain specificity, temporal dynamics, and operational constraints. These factors jointly explain the variability of reported performance results and underscore the need for application-aware model design, adaptive learning strategies, and evaluation frameworks aligned with real-world telecommunications requirements.

6.4. Future research directions: a research roadmap for telecom SA

This section addresses RQ3 by moving beyond a descriptive enumeration of emerging trends and proposing a structured research roadmap for SA in the telecommunications sector. The future directions outlined below are directly grounded in the challenges identified in the previous section and reflect both methodological and operational priorities.

A first major research axis concerns the development of multimodal SA frameworks capable of jointly processing text, speech, and visual signals. Multimodal approaches have shown strong potential for improving sentiment detection by capturing complementary emotional cues across modalities [41], [42]. In the telecommunications context, this direction is particularly relevant for call centers, where voice-based emotion recognition can significantly enhance customer support interactions. However, future research must address persistent challenges related to variability in voice tones, accents, background noise, and cultural communication styles, which currently limit the robustness and generalization of multimodal models.

A second priority axis involves the integration of LLMs into telecom SA pipelines. Pre-trained models such as generative pre-trained transformer (GPT), BERT, and LLaMA-based architectures offer powerful contextual representations that can capture complex semantic and syntactic relationships in customer feedback. Rather than treating LLMs as standalone classifiers, future work should investigate how they can be effectively adapted, fine-tuned, or combined with domain-specific resources to support SA tasks under telecom-specific constraints, including multilingualism, domain dependency, and limited labeled data.

A third research direction focuses on real-time and proactive sentiment analytics. The ability to analyze customer feedback streams in real time (whether from social networks, call centers, or digital channels) opens new opportunities for proactive customer service and early intervention. Real-time SA enables operators to adapt communication strategies dynamically and respond to emerging dissatisfaction before it escalates [43]. From a research perspective, this direction raises important questions regarding system latency, streaming architectures, and the trade-off between model complexity and responsiveness.

Closely related to data scarcity and generalization challenges, domain adaptation and transfer learning constitute another key research axis. As labeled telecom-specific datasets remain limited and costly to build, future studies should explore more effective adaptation strategies that leverage knowledge from related domains or across operators while preserving performance and reliability. This includes investigating cross-domain transfer mechanisms and continual learning strategies that reduce retraining costs over time.

Finally, privacy-preserving SA represents a strategic and increasingly unavoidable research direction. Telecommunications operators operate under strict regulatory and confidentiality constraints, which limit centralized data collection and sharing. Federated learning and decentralized training paradigms offer promising solutions by enabling local sentiment modeling while avoiding direct exposure of sensitive customer data. Future research should assess the practical trade-offs of such approaches in terms of performance, communication overhead, and deployment complexity in real-world telecom environments.

Overall, this research roadmap highlights a shift from isolated SA techniques toward adaptive, context-aware, and privacy-conscious sentiment intelligence systems. Addressing these research axes will be essential for translating SA from a descriptive monitoring tool into a scalable, proactive, and decision-support capability aligned with the evolving needs of telecommunications operators.

6.5. Ethics, privacy, and regulation

The application of SA in the telecommunications sector raises ethical and regulatory issues that are more complex than those encountered in generic social media analytics. Telecom operators handle large volumes of sensitive customer data, including personal identifiers, communication records, service usage information, and customer-support interactions. As a result, SA systems in this context cannot be designed independently of legal and regulatory constraints.

From a regulatory standpoint, data protection frameworks such as the general data protection regulation (GDPR) play a central role. In many SA use cases, customer communications constitute personal data, and sentiment inference may implicitly reveal behavioral or emotional attributes. This requires telecom operators to clearly define the legal basis for data processing, ensure transparency toward customers, and strictly control data retention and secondary use. In practice, these requirements significantly limit large-scale centralized data collection and complicate the reuse of customer data for model training, especially when consent is unclear or difficult to manage at scale. In addition to GDPR, data sovereignty and national compliance requirements are particularly influential in the telecommunications domain. Many countries impose sector-specific regulations that require customer data to be stored and processed within national borders or under the oversight of national regulatory authorities. These constraints directly affect architectural choices, especially when SA relies on cloud-based infrastructures or cross-border analytics pipelines. As a result, telecom operators often face trade-offs between analytical efficiency and regulatory compliance.

Ethical considerations extend beyond legal compliance. SA models may introduce biases related to language, dialect, or cultural expression, which is especially problematic in multilingual telecom markets. If such biases are not addressed, automated sentiment scoring may lead to unfair service prioritization or misinterpretation of customer dissatisfaction. Transparency is therefore essential, particularly when sentiment outputs are used to support automated decisions in customer support, churn prevention, or service recovery processes.

Within this regulatory and ethical landscape, privacy-preserving learning approaches, such as federated learning, are increasingly viewed as promising alternatives. By enabling model training across distributed data sources without transferring raw customer data, federated learning aligns well with both privacy regulations and data-sovereignty constraints. However, its adoption also raises new questions related to model governance, auditability, and performance monitoring, which remain largely unexplored in real-world telecom deployments.

Overall, ethical and regulatory constraints should not be treated merely as obstacles, but rather as structuring factors that shape how SA systems are designed and deployed in telecommunications. Explicitly accounting for these constraints is essential for building trustworthy, compliant, and operationally viable AI-driven customer analytics.

7. CONCLUSION

SA in the telecommunications sector stands out because of the complexity and diversity of customer interactions, as well as the growing need to react quickly to user feedback. This review shows that applying SA in this context is not straightforward, as telecom data often combines technical issues, multilingual content, and multiple communication channels. As a result, methods that work well in other domains cannot always be directly applied without adaptation.

At the same time, the findings highlight the real potential of SA as a practical tool for telecom operators. When effectively integrated into operational processes, it can help better understand customer expectations, detect early signs of dissatisfaction or churn, and support service improvement. These insights can be particularly valuable for decision-making across customer service, marketing, and network management.

While this study provides a structured view of current approaches, it also acknowledges certain limitations, particularly regarding data sources and language coverage. Looking ahead, the challenge will be to move toward more scalable, real-time, and trustworthy solutions that can be deployed in real-world environments. In this sense, SA has the potential to become not just an analytical tool, but a key component of customer-centric strategies in the telecommunications industry.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Achraf Bouhamidi	✓	✓		✓	✓	✓	✓	✓	✓	✓				✓
Naziha Laaz	✓	✓						✓		✓	✓	✓	✓	✓
Zineb Rachik	✓	✓		✓	✓			✓		✓	✓		✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

This study does not involve any identifiable personal data or human participants. Therefore, informed consent was not required.

ETHICAL APPROVAL

This study did not involve any experiments on human or animal subjects, and therefore ethical approval was not required.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

[1] Y. Ouyang, L. Wang, A. Yang, T. Gao, L. Wei, and Y. Zhang, "Next Decade of Telecommunications Artificial Intelligence," *CAAI Artificial Intelligence Research*, vol. 1, no. 1, pp. 28–53, Sep. 2022, doi: 10.26599/air.2022.9150003.

[2] A. Bin Rashid and M. A. K. Kausik, "AI revolutionizing industries worldwide: A comprehensive overview of its diverse applications," *Hybrid Advances*, vol. 7, p. 100277, Dec. 2024, doi: 10.1016/j.hybadv.2024.100277.

[3] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text," *Social Network Analysis and Mining*, vol. 11, no. 1, p. 81, Dec. 2021, doi: 10.1007/s13278-021-00776-6.

[4] H. Huan, Z. He, Y. Xie, and Z. Guo, "A Multi-Task Dual-Encoder Framework for Aspect Sentiment Triplet Extraction," *IEEE Access*, vol. 10, pp. 103187–103199, 2022, doi: 10.1109/ACCESS.2022.3210180.

[5] M. S. Akhtar, A. Ekbal, and E. Cambria, "How Intense Are You? Predicting Intensities of Emotions and Sentiments using Stacked Ensemble [Application Notes]," *IEEE Computational Intelligence Magazine*, vol. 15, no. 1, pp. 64–75, Feb. 2020, doi: 10.1109/MCI.2019.2954667.

[6] J. M. Sanchez-Gomez, M. A. Vega-Rodríguez, and C. J. Pérez, "Sentiment-oriented query-focused text summarization addressed with a multi-objective optimization approach," *Applied Soft Computing*, vol. 113, p. 107915, Dec. 2021, doi: 10.1016/j.asoc.2021.107915.

[7] V. Bonta, N. Kumaresh, and N. Janardhan, "A Comprehensive Study on Lexicon Based Approaches for Sentiment Analysis," *Asian Journal of Computer Science and Technology*, vol. 8, no. S2, pp. 1–6, Mar. 2019, doi: 10.51983/ajcst-2019.8.s2.2037.

[8] 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.

[9] F. Zhang, W. Zheng, and Y. Yang, "Graph Convolutional Network with Syntactic Dependency for Aspect-Based Sentiment Analysis," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, p. 37, Feb. 2024, doi: 10.1007/s44196-024-00419-6.

[10] K. H. Wu, P. Y. Chen, and C. Chiu, "A Social Media Based Profiling Approach for Potential Churning Customers: An Example for Telecom Industry," *Journal of Internet Technology*, vol. 23, no. 7, pp. 1565–1571, Dec. 2022, doi: 10.53106/160792642022122307011.

[11] R. Hermansyah and R. Sarno, "Sentiment Analysis Based on Quality Aspects in Effort to Improve Quality of Indhome Product and Services PT Telkom Indonesia Tbk," in *Proceedings of the 3rd International Conference on Business and Management of Technology (ICONBMT 2021)*, 2022. doi: 10.2991/aebmr.k.211226.005.

[12] A. Shoukry and A. Rafea, "Machine learning and semantic orientation ensemble methods for egyptian telecom tweets sentiment analysis," *Journal of Web Engineering*, vol. 19, no. 2, pp. 195–214, Jun. 2020, doi: 10.13052/jwe1540-9589.1924.

[13] H. Najadat, A. Al-Abdi, and Y. Sayaheen, "Model-based sentiment analysis of customer satisfaction for the Jordanian telecommunication companies," in *2018 9th International Conference on Information and Communication Systems, ICICS 2018*, IEEE, Apr. 2018, pp. 233–237. doi: 10.1109/IACS.2018.8355429.

[14] K. M. O. Nahar, A. Jaradat, M. S. Atoum, and F. Ibrahim, "Sentiment analysis and classification of arab jordanian facebook comments for jordanian telecom companies using lexicon-based approach and machine learning," *Jordanian Journal of Computers and Information Technology*, vol. 6, no. 3, pp. 247–262, 2020, doi: 10.5455/jjcit.71-1586289399.

[15] B. Abdullah, N. Alosaimi, and S. Almotiri, "Reputation Measurement based on a Hybrid Sentiment Analysis Approach for Saudi Telecom Companies," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, pp. 929–937, 2021, doi: 10.14569/IJACSA.2021.01206107.

[16] I. Ullah, B. Raza, A. K. Malik, M. Imran, S. U. Islam, and S. W. Kim, "A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector," *IEEE Access*, vol. 7, pp. 60134–60149, 2019, doi: 10.1109/ACCESS.2019.2914999.

[17] X. Xiahou and Y. Harada, "Customer Churn Prediction Using AdaBoost Classifier and BP Neural Network Techniques in the E-Commerce Industry," *American Journal of Industrial and Business Management*, vol. 12, no. 03, pp. 277–293, 2022, doi: 10.4236/ajibm.2022.123015.

[18] D. Subramanian, Ajitha, and S. S. Maidin, "Unveiling Hybrid Model with Naive Bayes, Deep Learning, Logistic Regression for Predicting Customer Churn and Boost Retention," *Journal of Applied Data Sciences*, vol. 6, no. 2, pp. 1379–1391, May 2025, doi: 10.47738/jads.v6i2.675.

[19] S. Aftan and H. Shah, "Using the AraBERT Model for Customer Satisfaction Classification of Telecom Sectors in Saudi Arabia," *Brain Sciences*, vol. 13, no. 1, p. 147, Jan. 2023, doi: 10.3390/brainsci13010147.

[20] M. A. Alshamari, "Evaluating User Satisfaction Using Deep-Learning-Based Sentiment Analysis for Social Media Data in Saudi Arabia's Telecommunication Sector," *Computers*, vol. 12, no. 9, p. 170, Aug. 2023, doi: 10.3390/computers12090170.

[21] S. M. Almutairi and F. M. Alotaibi, "Developing Arabic Sentiment Analysis for Saudi Arabia's Telecommunication Companies using Deep and Ensemble Learning," *Revista Gestão & Tecnologia*, vol. 23, no. 4, pp. 7–25, Dec. 2023, doi: 10.20397/2177-6652/2023.v23i4.2709.

[22] L. Almuqren and A. I. Cristea, "Predicting STC Customers' Satisfaction Using Twitter," *IEEE Transactions on Computational Social Systems*, vol. 10, no. 1, pp. 204–210, Feb. 2023, doi: 10.1109/TCSS.2021.3135719.

[23] S. T. Vieira, R. L. Rosa, D. Z. Rodríguez, M. A. Ramírez, M. Saadi, and L. Wuttisittikulkij, "Q-meter: Quality monitoring system for telecommunication services based on sentiment analysis using deep learning," *Sensors*, vol. 21, no. 5, pp. 1–18, Mar. 2021, doi: 10.3390/s21051880.

[24] A. Pareek, P. Poonam, and S. M. Arora, "Telecom customer churn prediction model : Analysis of machine learning techniques for churn prediction and factor identification in telecom sector," *Journal of Information and Optimization Sciences*, vol. 45, no. 2, pp. 613–630, 2024, doi: 10.47974/jios-1599.

[25] A. Borg and M. Boldt, "Using VADER sentiment and SVM for predicting customer response sentiment," *Expert Systems with Applications*, vol. 162, p. 113746, Dec. 2020, doi: 10.1016/j.eswa.2020.113746.

[26] S. S. Dzreke and S. E. Dzreke, "Preventing complaints before they happen: How AI-driven sentiment analysis enables proactive service recovery," *Advanced Research Journal*, vol. 10, no. 1, pp. 39–55, Aug. 2025, doi: 10.71350/3062192589.

[27] N. Jain, "Machine Learning for Suspicious Behaviour Detection and Churn Prediction in Telecom Customer Call Data," *International Journal of Emerging Research in Engineering and Technology*, vol. 6, no. 3, pp. 67–76, 2025, doi: 10.63282/3050-922x.ijeret-v6i3p108.

[28] A. Alamsyah and E. A. Bernatapi, "Evolving Customer Experience Management in Internet Service Provider Company using Text Analytics," in *Proceeding - 2019 International Conference on ICT for Smart Society: Innovation and Transformation Toward Smart Region, ICISS 2019*, IEEE, Nov. 2019, pp. 1–6. doi: 10.1109/ICISS48059.2019.8969828.

[29] N. Almuhanna and Z. H. Alharbi, "Factors Affecting Customer Satisfaction with The Telecommunication Industry in Saudi Arabia," *TEM Journal*, vol. 12, no. 1, pp. 424–433, Feb. 2023, doi: 10.18421/TEM121-52.

[30] H. Zhou *et al.*, "Large Language Model (LLM) for Telecommunications: A Comprehensive Survey on Principles, Key Techniques, and Opportunities," *IEEE Communications Surveys and Tutorials*, vol. 27, no. 3, pp. 1955–2005, Sep. 2025, doi: 10.1109/COMST.2024.3465447.

[31] L. Almuqren and A. Cristea, "AraCust: a Saudi Telecom Tweets corpus for sentiment analysis," *PeerJ Computer Science*, vol. 7, pp. 1–30, May 2021, doi: 10.7717/PEERJ-CS.510.

[32] Z. Wang, Z. Hu, S. B. Ho, E. Cambria, and A. H. Tan, "MiMuSA—mimicking human language understanding for fine-grained multi-class sentiment analysis," *Neural Computing and Applications*, vol. 35, no. 21, pp. 15907–15921, Jul. 2023, doi: 10.1007/s00521-023-08576-z.

[33] T. A. Al-Qablan, M. H. Mohd Noor, M. A. Al-Betar, and A. T. Khader, "A survey on sentiment analysis and its applications," *Neural Computing and Applications*, vol. 35, no. 29, pp. 21567–21601, Oct. 2023, doi: 10.1007/s00521-023-08941-y.

[34] M. M. Abubaera and S. M. Jiddah, "Natural Language Processing and Sentiment Analysis for Libyan Arabic Language Dataset," *International Journal of All Research Education and Scientific Methods (IJARESM)*, vol. 12, no. 6, pp. 2455–6211, 2024, [Online]. Available: www.ijaresm.com

[35] R. A. Bagate and R. Suguna, "Different Approaches in Sarcasm Detection: A Survey," in *Lecture Notes on Data Engineering and Communications Technologies*, 2020, pp. 425–433. doi: 10.1007/978-3-030-34080-3_48.

[36] S. M. Sarsam, H. Al-Samarraie, A. I. Alzahrani, and B. Wright, "Sarcasm detection using machine learning algorithms in Twitter: A systematic review," *International Journal of Market Research*, vol. 62, no. 5, pp. 578–598, Sep. 2020, doi: 10.1177/1470785320921779.

[37] K. R. Mabokela, T. Celik, and M. Raborife, "Multilingual Sentiment Analysis for Under-Resourced Languages: A Systematic Review of the Landscape," *IEEE Access*, vol. 11, pp. 15996–16020, 2023, doi: 10.1109/ACCESS.2022.3224136.

[38] M. S. U. Miah, M. M. Kabir, T. Bin Sarwar, M. Safran, S. Alfarhood, and M. F. Mridha, "A multimodal approach to cross-lingual sentiment analysis with ensemble of transformer and LLM," *Scientific Reports*, vol. 14, no. 1, p. 9603, Apr. 2024, doi: 10.1038/s41598-024-60210-7.

[39] Y. Y. Tan, C. O. Chow, J. Kanesan, J. H. Chuah, and Y. L. Lim, "Sentiment Analysis and Sarcasm Detection using Deep Multi-Task Learning," *Wireless Personal Communications*, vol. 129, no. 3, pp. 2213–2237, Apr. 2023, doi: 10.1007/s11277-023-10235-4.

[40] R. Sharma and K. Lakhwani, "A Systematic Literature Review on Cross Domain Sentiment Analysis Techniques: PRISMA Approach," *Annals of Emerging Technologies in Computing*, vol. 8, no. 4, pp. 30–55, Oct. 2024, doi: 10.33166/AETiC.2024.04.002.

[41] R. Das and T. D. Singh, "Multimodal Sentiment Analysis: A Survey of Methods, Trends, and Challenges," *ACM Computing Surveys*, vol. 55, no. 13, pp. 1–38, Dec. 2023, doi: 10.1145/3586075.

[42] S. Lai, X. Hu, H. Xu, Z. Ren, and Z. Liu, "Multimodal sentiment analysis: A survey," *Displays*, vol. 80, Jul. 2023, doi: 10.1016/j.displa.2023.102563.

[43] L. A. Maldonado-Canca, A. M. Casado-Molina, J. P. Cabrera-Sánchez, and G. Bermúdez-González, "Beyond the post: an SLR of enterprise artificial intelligence in social media," *Social Network Analysis and Mining*, vol. 14, no. 1, p. 219, Nov. 2024, doi: 10.1007/s13278-024-01382-y.

BIOGRAPHIES OF AUTHORS

Achraf Bouhamidi holds an Engineer's degree in Computer Science (2003) from the Mohammadia School of Engineers (EMI), Morocco, and a Specialized Master's degree in Computer Science and Telecommunications (2005) from CentraleSupélec (formerly École Centrale Paris), France. He is an IT professional with over 18 years of experience at a moroccan telecommunications operator, specializing in information systems, customer relationship management (CRM), and telecommunications. He is currently a Ph.D. student in Morocco, where his research focuses on the integration of artificial intelligence into CRM systems to enable proactive and predictive customer experiences. His research interests include sentiment analysis, machine learning in the telecommunications industry. He can be contacted at email: achraf.bouhamidi@gmail.com.

Naziha Laaz received the Master degree in Software Quality in 2013 and Doctorate of High Graduate Studies degrees in Computer Science in 2019 from the Faculty of Sciences of Ibn Tofail University, Morocco. She is currently professor at the Hassania School of Public Works, Casablanca, Morocco, affiliated with the MIG Department. She is also a researcher in computer science and applications, and a member of the ASYR research team at the LaGes Laboratory. Her main research interests are related to software engineering, model-driven engineering using ontologies, artificial intelligence, interoperability, and blockchain. She can be contacted at email: laaz.naziha@ehtp.ac.ma.

Zineb Rachik received the State Engineering degree in Computer Science in 2011 from the National School of Computer Science and Systems Analysis (ENSIAS), Rabat, Morocco. She obtained the Ph.D. degree in Computer Science in 2015 from the Faculty of Sciences Ben M'Sik, Hassan II University of Casablanca, Morocco. She is currently a Professor at the Hassania School of Public Works (École Hassania des Travaux Publics – EHTP), Casablanca, Morocco, affiliated with the Mathematics, Computer Science and Geomatics (MIG) Department. She is also a researcher in computer science and its applications, and a member of the ASYR research team at the Laboratory of Systems Engineering (LaGeS). Her research interests include software engineering, controllability, prediction, mathematical modeling, and artificial intelligence. She can be contacted at email: rachik.zineb@ehtp.ac.ma.