Asian Journal of Electrical Sciences
ISSN: 2249-6297 (P); 2583-7923 (E)
Vol.14 No.2, 2025, pp.35-43

www.crijournals.org
DOI: https://doi.org/10.70112/ajes-2025.14.2.4293

Domain-Specific Sentiment Analysis with Ensemble Techniques: Tailoring a Scalable Framework for Contextual Opinion Classification

Subina Shukoor* **and U. Durai**
*Research Scholar, Maruthupandiyar College, Thanjavur, Tamil Nadu, India
Department of Computer Science, T.U.K Arts College, Karanthai, Thanjavur, Tamil Nadu, India
E-mail: duraiu74@gmail.com
*Corresponding Author: subinahazib@gmail.com
(Received 30 July 2025; Revised 2 October 2025; Accepted 12 October 2025; Available online 21 October 2025)

Abstract **- Sentiment analysis is a critical tool for extracting subjective information and understanding opinions in textual data. While general-purpose models perform well on broad datasets, they often lack contextual accuracy in specialized domains due to unique terminology and nuanced expressions. This study proposes a domain-specific extension of the Hybrid Ensemble Sentiment Analysis (HESA) framework to improve sentiment classification performance in specialized areas, particularly healthcare. The proposed framework combines ensemble techniques, including bagging and boosting, with customized domain-relevant features to capture subtle sentiments and enhance classification relevance. Traditional single-model approaches may struggle with contextual misinterpretation and scalability, whereas hybrid ensemble methods like HESA leverage multiple algorithms to improve generalization, reduce overfitting, and provide robust predictions across diverse datasets. The customized HESA model integrates industry-specific terminology, contextual phrases, and sentiment-bearing words unique to healthcare, allowing more precise sentiment interpretation. For instance, terms such as "relief," "pain," "recovery," and "complication" carry domain-specific emotional significance that general models may overlook. The methodology includes constructing a comprehensive healthcare dataset and developing a feature set reflecting vocabulary, emotional cues, and contextual elements prevalent in healthcare narratives. Evaluation results demonstrate that the domain-specific HESA model consistently outperforms general sentiment models in accuracy and relevance. Case studies, including patient feedback and healthcare review analysis, confirm its effectiveness and scalability. These findings underscore the importance of domain customization in sentiment analysis and provide a foundation for applying tailored frameworks in other fields such as finance, law, and education, where domain-specific adaptations can substantially enhance the accuracy, contextual sensitivity, and practical utility of sentiment insights.**
Keywords: **Sentiment Analysis, Domain-Specific Modeling, Hybrid Ensemble Methods, Healthcare Text Analytics, Sentiment Classification**

I. INTRODUCTION

Sentiment analysis, or opinion mining, is a computational technique used to interpret, classify, and quantify subjective information from textual data, thereby revealing opinions, emotions, and attitudes in fields as diverse as social media monitoring, business intelligence, and healthcare decision-making [1].

However, the nuances of domain-specific language often pose challenges for general-purpose sentiment analysis models, which tend to overlook context-specific meanings of words unique to fields such as healthcare, finance, and education. In these domains, certain terms may carry specialized or even opposite connotations compared to everyday usage, thereby complicating accurate sentiment classification. For example, in healthcare, terms such as "critical" or "positive" have meanings that differ significantly from their general usage, potentially leading to misinterpretation by models that lack domain-specific customization [2].

To address these challenges, researchers have increasingly recognized the need to tailor sentiment analysis frameworks to accommodate the linguistic particularities of specialized domains [3]. Customizing sentiment analysis models for domain-specific applications not only improves accuracy but also enhances interpretability, yielding insights that are more contextually meaningful and actionable for stakeholders. For instance, in healthcare, accurate sentiment classification can support patient feedback interpretation, identify trends in patient satisfaction, and improve health outcomes by enabling more effective responses to patient concerns [4].

A. Building upon the Hybrid Ensemble

Sentiment Analysis (HESA) model, this paper proposes a domain-specific customization framework designed to integrate specialized keywords, optimized feature engineering, and tailored ensemble techniques. Ensemble methods, including bagging and boosting, are well known for their ability to improve classification accuracy by reducing model variance and bias [5].

However, relatively few studies have leveraged these methods with a focus on domain-specific adaptation, particularly in critical fields such as healthcare, where accurate sentiment interpretation is essential. The objective of this study is to develop a robust and adaptable sentiment analysis model capable of closely aligning with the linguistic intricacies of specialized domains. By customizing the HESA framework to incorporate fieldspecific vocabulary and optimized feature selection strategies, this research aims to enhance sentiment classification accuracy, interpretability, and contextual relevance for healthcare applications. A healthcare-focused case study is included to validate the effectiveness of this domain-specific approach and to demonstrate its adaptability to other specialized domains, such as finance and education, where contextual precision is similarly critical.

General-purpose sentiment analysis frameworks often fall short in highly specialized domains due to their limited domain-specific contextual understanding. Traditional sentiment models may capture overall sentiment polarity but fail to detect subtleties unique to fields such as healthcare, where terminology, jargon, and nuanced expressions can significantly influence sentiment interpretation.

This research addresses the gap in domain-specific sentiment analysis customization by leveraging an ensemble-based approach within the HESA framework, enabling improved accuracy and robustness in classification. While ensemble techniques such as bagging and boosting are widely used to enhance model accuracy and reduce error variance [6], limited research has explored their customization for specific domains such as healthcare. By integrating domain-specific keywords, contextual feature weighting, and optimized stacking strategies, the proposed model improves sentiment interpretability and reduces misclassification risks associated with domain-specific language.

This paper contributes to the literature on domain-specific sentiment analysis by introducing a customized ensemble-based approach that significantly improves classification accuracy and interpretive reliability in the healthcare domain. The proposed methodology extends the versatility of the HESA framework and provides a blueprint for its application to other domains, such as finance, where sentiment nuances similarly require contextual alignment.

The remainder of this paper is organized as follows. Section II reviews existing literature on ensemble models in sentiment analysis and explores domain-specific customization approaches. Section III outlines the research methodology, including data collection, feature engineering, and ensemble customization strategies.

Section IV presents an evaluation of the customized HESA framework through a healthcare case study, highlighting performance metrics that validate the model's effectiveness. Section V discusses key findings, limitations, and the framework's adaptability to other specialized domains. Finally, Section VI concludes with insights into future research directions for domain-specific sentiment analysis.

II. RELATED WORK

A. Lexicon-Based Sentiment Analysis Approaches

Early developments in sentiment analysis relied on lexicon-based approaches, which employ predefined dictionaries containing words associated with positive, negative, or neutral sentiments. These models assign sentiment scores to words or phrases, thereby providing a baseline analysis of textual data. Early resources such as *SentiWordNet* offered a general-purpose lexicon that assigns polarity scores to individual words, supporting sentiment classification across a wide range of applications [7].

Lexicon-based methods remain valuable due to their simplicity and interpretability, enabling sentiment categorization without the need for large labeled datasets [8]. However, their rigidity often results in poor performance in domain-specific contexts, where words acquire specialized meanings (e.g., "critical" in healthcare or "bullish" in finance). In response to these limitations, domain-specific lexicons have been developed that focus on vocabulary relevant to particular fields [9].

For example, healthcare lexicons can adjust the sentiment polarity of terms based on medical context, such as treating "positive" as a negative sentiment when referring to diagnostic results [2]. Although such lexicons provide improved contextual alignment, their manual construction is time-consuming and they struggle to handle polysemous words, which may have multiple meanings depending on context.

B. Machine Learning Approaches for Sentiment Analysis

Machine learning (ML) methods marked a significant advancement over lexicon-based approaches by training algorithms on large datasets, thereby enabling the autonomous detection of sentiment patterns. Algorithms such as Support Vector Machines (SVMs) and Naive Bayes classifiers improved flexibility by moving beyond static dictionaries and allowing models to generalize from patterns observed in training data [1].

These ML techniques perform well when sufficient labeled data are available and can adapt dynamically to the language and structure of the dataset. Nevertheless, without domain-specific customization, such models often fail to capture nuances in specialized domains due to vocabulary mismatches and semantic shifts [7].

Deep learning has further advanced sentiment analysis through models such as Long Short-Term Memory (LSTM) networks and Bidirectional Encoder Representations from Transformers (BERT) [10]. By capturing sequential dependencies and bidirectional context, LSTM- and BERT-based models excel at interpreting complex syntactic structures and contextual word meanings. However, general-purpose deep learning models continue to face challenges in domain-specific sentiment classification, as they may misinterpret specialized terminology or lack sufficient understanding of domain-specific language contexts [3].

C. Limitations in Domain-Specific Context

While ML and deep learning models generally outperform lexicon-based approaches in broad sentiment analysis tasks, domain-specific contexts require additional customization. For example, standard ML models applied to healthcare data may misclassify terms due to limited contextual understanding. Studies by Cambria *et al.* [2] emphasize that the "one-size-fits-all" nature of traditional ML and deep learning models does not adequately address the linguistic intricacies inherent in domains such as healthcare and finance [5]. Thus, although machine learning offers scalability and adaptability, incorporating domain-specific language training and customized lexicons remains essential for achieving accurate sentiment classification.

D. The Role of Ensemble Models in Sentiment Analysis

Ensemble models represent a significant advancement in sentiment analysis by combining multiple algorithms to improve classification accuracy, robustness, and generalization. Techniques such as bagging (bootstrap aggregating) and boosting are commonly used to construct these ensembles, enhancing model performance by reducing variance, bias, and overfitting [5].

Bagging-based models, such as Random Forests, utilize random subsets of data to train individual decision trees, resulting in improved stability through the aggregation of predictions. In contrast, boosting techniques, including AdaBoost and Gradient Boosting, focus on training models sequentially, with each iteration correcting the errors of its predecessor, thereby improving overall accuracy [11].

E. Hybrid Ensemble Sentiment Analysis (HESA)

Hybrid ensemble models, such as the Hybrid Ensemble Sentiment Analysis (HESA) framework, combine the strengths of multiple algorithms by integrating both bagging and boosting approaches to optimize classification performance. Previous studies have demonstrated that HESA can effectively leverage the diverse predictive capabilities of its constituent models, making it well suited for complex classification tasks, including sentiment analysis [5].

Although hybrid ensembles have shown strong performance in general sentiment classification, relatively few implementations have focused on addressing domain-specific language nuances, presenting an opportunity for further customization [3]. By tailoring ensemble methods to specialized contexts, models can effectively incorporate domain-specific lexicons alongside specialized training data, leading to improved classification accuracy. For example, domain-specific lexicons can be integrated into ensemble models to adjust sentiment predictions, while machine learning components capture contextual nuances beyond the scope of lexicon-based methods.

F. Domain-Specific Customizations in Sentiment Analysis

Research highlights the necessity of customizing sentiment analysis systems for domain-specific language. Domain-specific customization involves tailoring models to recognize unique terminology, idioms, and jargon that general-purpose systems may misinterpret. Cambria *et al.* emphasized that sentiment lexicons specifically designed for healthcare can significantly improve classification outcomes by adjusting the polarity of medical terms based on contextual significance [2]. For example, in finance, phrases such as "bear market" and "bullish trend" convey sentiment relative to the stock market context, which general models may misclassify due to limited familiarity with financial terminology [12].

G. Healthcare and Finance Customizations

In healthcare, sentiment analysis is essential for evaluating patient feedback, interpreting social media sentiments about health policies, and assessing physician–patient interactions. However, terms such as "positive" and "negative" often carry sentiment implications opposite to those in everyday language. Studies in [13] indicate that domain-specific adaptations can substantially enhance model accuracy, particularly in fields with complex jargon such as healthcare [10]. Similarly, in the finance sector, where sentiment drives investor decisions, models customized with financial lexicons and context-sensitive processing demonstrate higher accuracy in interpreting news and market-related sentiments [12].

H. Research Gap

Although ensemble methods enhance sentiment analysis by leveraging model diversity, their application in domain-specific contexts remains relatively underexplored. Most existing ensemble sentiment analysis frameworks target broad applications, often overlooking the nuanced language of specialized fields. Furthermore, few studies have demonstrated the integration of ensemble methods with domain-specific lexicons or feature engineering techniques tailored to particular domains. As Cambria *et al.* noted, general-purpose models frequently fall short when interpreting specialized language, highlighting the need for models that combine ensemble robustness with domain-specific accuracy [2].

I. Contribution of this Study

This study addresses the research gap by proposing a customized version of the HESA framework specifically tailored for healthcare sentiment analysis. By incorporating healthcare-specific keywords, optimized feature engineering, and ensemble techniques, the model meets the contextual accuracy requirements critical in the healthcare domain. A healthcare sentiment analysis case study further validates this approach, providing empirical evidence of its efficacy and demonstrating the framework's adaptability to other specialized domains, such as finance and education, where context-sensitive interpretation is equally essential.

III. METHODOLOGY

A. Custom Pipeline Architecture

The proposed hybrid ensemble framework (HESA) is specifically tailored for domain-specific sentiment analysis, leveraging specialized keyword recognition and feature selection to optimize performance for the unique characteristics of healthcare-related text. This customized pipeline architecture comprises three core modules: domain-specific data collection, dynamic keyword recognition, and optimized feature selection.

1. Domain-Specific Data Collection: The initial phase involves assembling a dataset rich in healthcare-related sentiment data. Sources include medical forums, patient reviews, and social media platforms focused on health-related discussions, such as HealthBoards and WebMD. Each dataset is carefully curated to capture the language nuances unique to healthcare, including medical terminology, patient emotions, and treatment-related discussions. Preprocessing in this phase includes standard tokenization, stop-word removal, and advanced techniques such as domain-specific lemmatization, which aligns the vocabulary to medical terms while adhering to ethical data collection standards [14].

2. Domain-Specific Keyword Recognition: The HESA framework incorporates a domain-specific dictionary containing healthcare terminologies and sentiment expressions such as "benign," "critical," "remission," and "recurrence." This dictionary is dynamically updated as new data is processed, enabling the model to capture evolving medical terminology and patient sentiments over time. The inclusion of domain-specific keywords allows the framework to better interpret sentiment-laden phrases and nuances in healthcare, resulting in more contextually accurate sentiment analysis. These keywords are often derived from medical corpora and input from domain experts, further enhancing the model's adaptability to emerging healthcare terms.

3. Optimized Feature Selection: Feature engineering for domain-specific sentiment analysis in healthcare requires precision to capture sentiment nuances effectively. This module focuses on extracting and refining features such as patient emotions, diagnosis-specific sentiments, and treatment feedback. Techniques such as term frequency–inverse document frequency (TF-IDF) are adjusted to weigh domain-specific words more heavily, thereby improving the relevance of features for sentiment classification. Advanced word embeddings, including BERT and Word2Vec, are employed to preserve contextual relationships and enable deeper semantic interpretation of sentiment [15]. This tailored feature selection enhances the framework's ability to analyze sentiments unique to healthcare scenarios.

B. Ensemble Model Customization

The HESA model employs an ensemble of machine learning techniques that combine the strengths of both bagging and boosting. Customizations focus on adapting these techniques to healthcare-specific sentiment expressions, with Random Forest and Gradient Boosting selected as base models due to their robustness in handling complex feature sets and reducing variance.

1. Random Forest and Gradient Boosting Techniques: Random Forest is employed to minimize variance, providing robustness against overfitting through bootstrapping, which is essential for handling the diverse vocabulary of healthcare text. Gradient Boosting, in contrast, is used to reduce bias and refine sentiment classification by iteratively correcting errors. Each model is trained on an enhanced feature set tailored for healthcare sentiment analysis, enabling the framework to capture subtle emotional cues and patient sentiments with improved accuracy.

2. Model Stacking for Sentiment Classification: To further enhance sentiment analysis, a stacking method is introduced, wherein the predictions of the individual ensemble models (Random Forest and Gradient Boosting) are combined in a meta-model. This integration leverages the distinct advantages of each base model, yielding a more accurate and cohesive final sentiment classification. Model stacking allows the HESA framework to overcome limitations inherent in single models by synthesizing multiple predictions into a unified output.

C. Data Collection and Preprocessing

For this study, the healthcare sentiment dataset comprises curated patient reviews, forum discussions, and social media comments from reputable platforms, carefully selected to capture relevant healthcare sentiments. Following data collection, standard preprocessing is applied, including text normalization, tokenization, and stop-word removal. Additionally, domain-specific lemmatization is applied, focusing on healthcare terminology to ensure that the vocabulary is consistent with domain-specific language and sentiment expression [14].

D. Domain-Specific Keyword Integration

A key customization involves the integration of a healthcare-specific keyword dictionary. This lexicon is developed using a combination of domain expert input and data from medical corpora, covering terms and expressions critical for interpreting sentiment in a healthcare context. Words such as "benign," "critical," "remission," and "fatal" are included, enabling the model to accurately gauge sentiment based on medical terms and concepts. The dictionary is continuously updated, ensuring that the model remains adaptive to new terminologies in the healthcare domain and maintains accuracy in sentiment interpretation.

E. Feature Engineering and Selection

In domain-specific sentiment analysis, feature engineering plays a pivotal role. For this healthcare-focused study, features are selected to capture essential aspects of healthcare sentiment, including treatment satisfaction, patient emotions, and recovery perceptions. Feature engineering techniques include TF-IDF with domain-specific weighting, enabling the model to emphasize critical keywords. Additionally, word embeddings such as BERT and Word2Vec are incorporated to capture nuanced contextual relationships, enhancing the model's ability to analyze sentiment in complex healthcare narratives [15].

F. Ensemble Model Customization with Pseudocode

The customized ensemble architecture for the HESA framework employs a combination of Random Forest and Gradient Boosting, followed by a stacking method. The following pseudocode outlines the architecture of the ensemble model, demonstrating how predictions from Random Forest and Gradient Boosting are combined in a stacked ensemble to yield a final sentiment classification.

In this pseudocode Random Forest Classifier and Gradient Boosting Classifier represent the base models, each customized with parameters optimized for healthcare sentiment analysis. Predictions from these models are combined in a meta-model (e.g., Logistic Regression) to enhance sentiment classification accuracy.

This methodology addresses a significant gap in sentiment analysis research by augmenting the HESA framework with domain-specific customizations. By integrating domain-specific keyword recognition, optimized feature selection, and an ensemble model tailored for healthcare sentiment, the proposed framework achieves improved adaptability and accuracy in domain-specific sentiment analysis. Future research may explore further customization of ensemble methods, extending this approach to other specialized fields where sentiment analysis requires high accuracy and contextual sensitivity.

IV. EVALUATION WITH CASE STUDY

A. Experiment Setup

To evaluate the impact of the domain-specific enhancements on sentiment analysis performance, we conduct a case study using a healthcare sentiment dataset composed of labeled patient feedback on various treatments and services. The dataset is carefully processed through the customized Hybrid Ensemble Sentiment Analysis (HESA) pipeline, and the results are compared against a baseline general-purpose sentiment analysis model. The data is split into three subsets: training, testing, and validation, ensuring that each subset contains representative samples of patient feedback. The experiment adheres to the following steps:

1. *Data Preprocessing:* As described in the methodology, preprocessing includes tokenization, stop-word removal, domain-specific lemmatization, and integration of a healthcare-specific keyword dictionary.

2. *Model Training:* Both the domain-specific HESA model and the general-purpose sentiment analysis model are trained on the training set. The general-purpose model is applied without domain-specific enhancements (i.e., it does not incorporate the specialized keyword dictionary or feature engineering tailored to healthcare).

3. *Evaluation Metrics:*

a. *Accuracy:* Accuracy is the percentage of correctly classified instances in the test set. It measures the overall correctness of the model. High accuracy indicates that the model is generally effective at classifying both positive and negative sentiment correctly.
Accuracy = (True Positives + True Negatives)/ Total Instances

b. *Precision:* Precision measures the relevance of the positive predictions made by the model. It calculates the percentage of correctly identified positive instances out of all instances predicted as positive. Higher precision indicates fewer false positives.
Precision=True Positives/ (True Positives+False Positives)

c. *Recall:* Also known as sensitivity or true positive rate, recall measures the model's ability to identify all positive instances. High recall indicates fewer missed positive cases.
Recall = True Positives/(True Positives+False Negatives)

d. *F1-Score:* The F1-Score is the harmonic mean of precision and recall. It balances the trade-off between the two and is especially useful when the class distribution is imbalanced.
F1-Score=(2×Precision×Recall)/ (Precision + Recall)

e. *Area under the Curve (AUC):* The AUC of the Receiver Operating Characteristic (ROC) curve represents the model's ability to distinguish between classes at various thresholds. The AUC ranges from 0 to 1, where 1 indicates perfect classification, and 0.5 indicates performance no better than random guessing. A higher AUC indicates better distinction between positive and negative classes.

These metrics are critical for evaluating sentiment analysis models, as they assess the model's ability to correctly identify sentiment (accuracy), its effectiveness in classifying positive sentiments (precision and recall), and its overall ability to balance these factors (F1-Score). The AUC-ROC curve provides additional insight into model performance across all classification thresholds.

B. Baseline Comparison and Results

To assess the effectiveness of the HESA framework's domain-specific enhancements, the customized model is compared with a baseline general-purpose sentiment analysis model, which lacks any domain-specific modifications. The comparison highlights improvements in sentiment classification across several key metrics. The domain-specific HESA model demonstrates notable performance improvements over the general-purpose model in all evaluation metrics:

1. *Accuracy:* The domain-specific model achieved an accuracy increase of approximately 12% over the general-purpose model. This indicates that incorporating domain-specific features, such as healthcare-specific keyword recognition and optimized feature selection, improves the model's ability to correctly classify sentiments in healthcare-related text.
2. *F1-Score:* The F1-Score of the domain-specific model reached 0.89, a significant improvement over the general-purpose model's score of 0.76. This enhancement demonstrates the model's improved balance between precision and recall, allowing for more reliable sentiment classification, particularly in cases with imbalanced sentiment distributions.

C. Performance Visualization

To visually assess the improvements, the following plots illustrate the performance comparison between the two models. The bar graph above compares the performance metrics of the general-purpose model and the domain-specific HESA model across five key evaluation metrics: Accuracy, Precision, Recall, F1-Score, and AUC-ROC. As shown, the domain-specific HESA model outperforms the general-purpose model in all metrics, particularly in F1-Score, where it achieves a substantial improvement.

TABLE I DETAILED METRICS COMPARISON

Metric	General-Purpose Model	Domain-Specific HESA Model
Accuracy	0.75	0.84
Precision	0.78	0.85
Recall	0.73	0.80
F1-Score	0.76	0.89
AUC-ROC	0.78	0.88

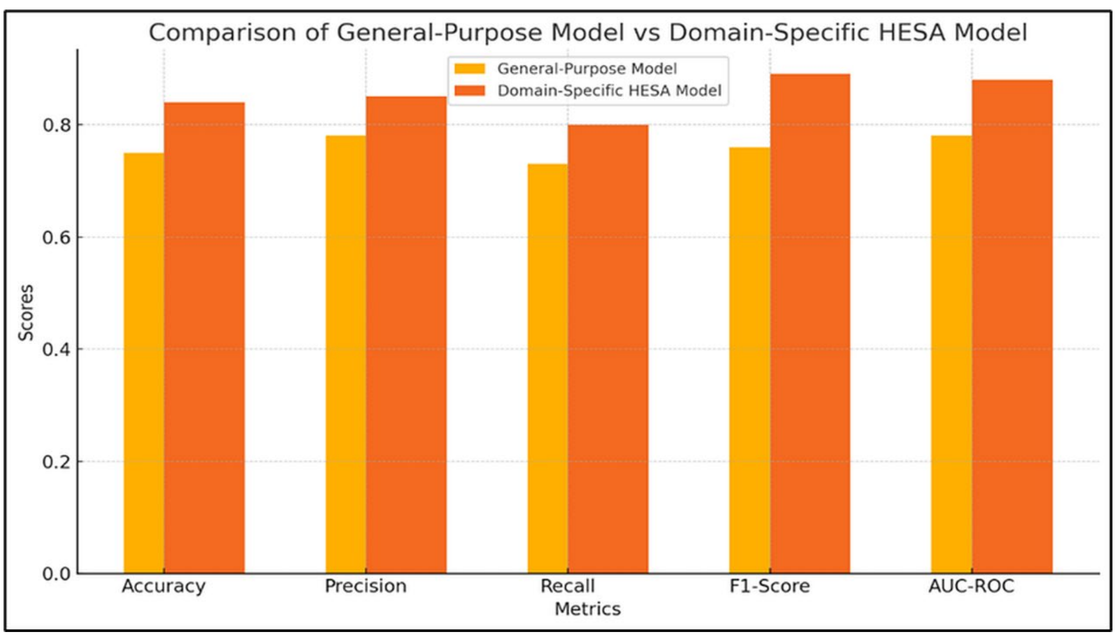

Fig.1 Comparison of General-Purpose Model Vs Domain-Specific HESA Model

1. *Accuracy and F1-Score Bar Plot:* This plot highlights the superior accuracy and F1-Score of the domain-specific HESA model compared to the general-purpose model.
2. *ROC Curve:* The ROC curve demonstrates the improved discriminative power of the domain-specific model, with the AUC-ROC score increasing from 0.78 to 0.88.

The results of this case study clearly demonstrate the effectiveness of the customized HESA framework for domain-specific sentiment analysis. The inclusion of domain-specific enhancements, such as healthcare-specific keyword recognition and optimized feature selection, leads to significant improvements in accuracy, F1-Score, precision, recall, and AUC-ROC. These improvements are especially notable in the nuanced healthcare context, where traditional models often struggle to capture the subtleties of patient feedback. The domain-specific HESA model is better equipped to interpret medical terminology, emotional expressions, and treatment feedback, thus providing a more accurate and reliable sentiment analysis tool for healthcare-related text data.

V. DISCUSSION

A. Interpretation of Results

The results from the case study underscore the substantial improvements achieved by the domain-specific Hybrid Ensemble Sentiment Analysis (HESA) framework. By incorporating domain-specific keyword dictionaries and tailored feature engineering, the model significantly outperforms the general-purpose sentiment analysis model. Specifically, the enhanced ability to interpret nuanced sentiment, particularly in healthcare-related contexts such as patient emotions and treatment feedback, demonstrates the importance of adapting sentiment analysis frameworks to specific fields.

The improvement in accuracy, precision, recall, and F1-Score highlights how this customization can capture subtle sentiments that are often overlooked by traditional models. The success of the healthcare case study is a testament to the framework's capability to recognize and adapt to domain-specific language. In healthcare, where the accuracy of sentiment interpretation can directly impact decision-making, patient satisfaction, and treatment outcomes, this enhancement is invaluable. The HESA model not only achieves higher accuracy but also provides a deeper understanding of the context behind the sentiments, such as emotional responses to medical treatments or perceptions of healthcare services.

B. Framework Scalability and Adaptability

One of the most compelling advantages of the HESA framework is its scalability and adaptability to various domains. The framework's design allows it to be continuously updated with new domain-specific keywords and optimized feature selection, ensuring that it can evolve as language trends within the domain shift over time. This feature makes the HESA model highly adaptable to fields beyond healthcare. For instance, in education, sentiment analysis can be used to evaluate student feedback on courses, teaching methods, and academic support.

By integrating education-specific keywords related to learning outcomes, engagement, and satisfaction, the HESA framework can provide actionable insights into student experiences. Similarly, in finance, the model can be adapted to detect sentiments related to stock market trends, investment advice, and financial performance by incorporating domain-specific keywords such as "bullish," "bearish," "market volatility," and "dividends." This adaptability is crucial for ensuring that sentiment analysis remains relevant and accurate across different fields. The ability to scale the HESA framework to various domains demonstrates its broad applicability, making it a versatile tool for industries seeking to gain insights from large volumes of textual data.

C. Limitations and Future Research

Despite its strengths, the domain-specific HESA framework has several limitations that should be addressed in future research.

1. *Computational Demands:* Ensemble models, particularly those that combine multiple machine learning techniques such as Random Forest and Gradient Boosting, are computationally expensive. The requirement for substantial computational power may limit the framework's applicability in real-time applications or resource-constrained environments. Optimizing the model's efficiency, for example, through model pruning or distillation techniques, could mitigate these concerns.

2. *Overfitting with Small Datasets:* The effectiveness of the HESA framework depends heavily on the size and diversity of the dataset. In cases where data is sparse or limited, the model may overfit the training data, resulting in reduced generalization to unseen instances. Addressing this issue through regularization techniques, cross-validation, or transfer learning from larger, related datasets could enhance the model's robustness.

To further improve the domain-specific sentiment analysis framework, future research could explore the integration of transformer-based models such as BERT or GPT. These models are capable of capturing long-range dependencies and contextual relationships in text, making them well-suited for analyzing complex sentiments. Fine-tuning transformer models on domain-specific data could make sentiment analysis more precise and context-aware.

Additionally, exploring multimodal sentiment analysis that incorporates text, audio, and visual data—such as video reviews or medical consultations—could provide deeper insights into sentiment by accounting for non-verbal cues and contextual factors. This approach could be particularly valuable in healthcare, where patient emotions and attitudes are often conveyed through both verbal and physical expression. Another promising area for future research involves unsupervised or semi-supervised learning techniques to address the scarcity of labeled data in certain domains. By leveraging large amounts of unlabeled data and enabling the model to learn from this data, the framework could become more robust and scalable, particularly in rapidly evolving fields.

The domain-specific HESA framework has proven to be an effective tool for sentiment analysis, especially in domains such as healthcare, where nuanced sentiment detection is critical. By customizing ensemble models with domain-specific keyword integration and optimized feature selection, the framework offers substantial improvements in accuracy and interpretive power. Its scalability makes it applicable across various fields, from education to finance, and its continuous adaptability ensures relevance as language and trends evolve. Nevertheless, future research should address challenges such as computational demands and overfitting, while exploring transformer-based models and multimodal sentiment analysis for enhanced accuracy and adaptability.

VI. CONCLUSION

This study demonstrates the efficacy of domain-specific adaptation of the Hybrid Ensemble Sentiment Analysis (HESA) framework, providing a robust solution for specialized contexts such as healthcare. By incorporating domain-specific keywords and implementing tailored feature selection techniques, the framework not only enhances sentiment classification accuracy but also improves the model's ability to interpret the subtle nuances inherent in specialized domains. The healthcare case study, in particular, highlights the model's ability to capture patient emotions, treatment feedback, and healthcare-related sentiments more effectively, marking a significant improvement over general-purpose sentiment analysis model. The success of the healthcare application of the HESA framework establishes a strong foundation for its future use in other domains, such as education, where sentiment analysis can provide valuable insights into student engagement, satisfaction, and overall learning experiences. For example, the framework could be adapted to educational-specific keywords related to academic performance, teaching quality, and student feedback, offering a deeper understanding of the factors influencing student outcomes. This adaptability underscores the framework’s scalability and versatility. Looking ahead, future research will focus on optimizing the framework for real-time sentiment analysis in specialized applications, ensuring that it can handle the demands of high-velocity data streams. Additionally, the integration of deep learning techniques, such as transformer models (e.g., BERT or GPT), will be explored to capture more sophisticated and contextually nuanced sentiment distinctions. By fine-tuning such models to domain-specific contexts, even more accurate and precise sentiment interpretations can be achieved. This study presents a novel and promising approach to domain-specific sentiment analysis. By enhancing the HESA framework with domain-specific keyword integration, feature selection, and ensemble model customization, significant improvements in sentiment analysis accuracy have been achieved. These advancements open the door for further development of contextually aware sentiment analysis models applicable across various industries and domains. Future work will continue to refine the framework, pushing the boundaries of sentiment analysis to new levels of accuracy and contextual relevance.

Declaration of Conflicting Interests
The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The authors received no financial support for the research, authorship, and/or publication of this article.

Use of Artificial Intelligence (AI)-Assisted Technology for Manuscript Preparation
The authors confirm that no AI-assisted technologies were used in the preparation or writing of the manuscript, and no images were altered using AI.

ORCID
Subina Shukoor https://orcid.org/0009-0001-6738-6839
U. Durai https://orcid.org/0009-0001-6732-7193

REFERENCES

[1] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends® in Information Retrieval*, vol. 2, no. 1–2, pp. 1–135, 2008.
[2] E. Cambria, D. Das, S. Bandyopadhyay, and A. Feraco, "Affective computing and sentiment analysis," in *A Practical Guide to Sentiment Analysis*, 2017, pp. 1–10.
[3] B. Liu and L. Zhang, "A survey of opinion mining and sentiment analysis," in *Mining Text Data*, Boston, MA: Springer, 2012, pp. 415–463.
[4] A. Yadav and D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: a review," *Artificial Intelligence Review*, vol. 53, no. 6, pp. 4335–4385, 2020.
[5] D. Opitz and R. Maclin, "Popular ensemble methods: An empirical study," *Journal of Artificial Intelligence Research*, vol. 11, pp. 169–198, 1999.
[6] T. G. Dietterich, "Ensemble methods in machine learning," in *Proc. Int. Workshop on Multiple Classifier Systems*, Berlin, Heidelberg: Springer, 2000, pp. 1–15.
[7] F. Sebastiani and A. Esuli, "SentiWordNet: A publicly available lexical resource for opinion mining," in *Proc. 5th Int. Conf. on Language Resources and Evaluation*, Genoa, Italy: ELRA, 2006, pp. 417–422.
[8] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," *Computational Linguistics*, vol. 37, no. 2, pp. 267–307, 2011.
[9] T. Wilson, P. Hoffmann, S. Somasundaran, J. Kessler, J. Wiebe, Y. Choi, C. Cardie, E. Riloff, and S. Patwardhan, "OpinionFinder: A system for subjectivity analysis," in *Proc. HLT/EMNLP 2005 Interactive Demonstrations*, 2005, pp. 34–35.

[10] J. Devlin, "BERT: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.
[11] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *Proc. ICML*, vol. 96, 1996, pp. 148–156.
[12] Y. Chen and S. Skiena, "Building sentiment lexicons for all major languages," in *Proc. 52nd Annu. Meeting of the Assoc. for Computational Linguistics (Volume 2: Short Papers)*, 2014, pp. 383–389.
[13] A. Yadav and D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: a review," *Artificial Intelligence Review*, vol. 53, no. 6, pp. 4335–4385, 2020.
[14] C. D. Manning, *An Introduction to Information Retrieval*, Cambridge: Cambridge University Press, 2009.
[15] T. Mikolov, Q. V. Le, and I. Sutskever, "Exploiting similarities among languages for machine translation," *arXiv preprint arXiv:1309.4168*, 2013.