IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 2 March 2025, accepted 13 March 2025, date of publication 20 March 2025, date of current version 22 April 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3553386

RESEARCH ARTICLE

Machine Learning-Based Sentiment Analysis in English Literature: Using Deep Learning Models to Analyze Emotional and Thematic Content in Texts

JIE YU[1,2] AND CHUNHONG QI[1]

[1]Yunnan Chinese Language and Culture College, Yunnan Normal University, Kunming 650500, China
[2]School of Foreign Languages, Jiangxi University of Science and Technology, Ganzhou 341000, China

Corresponding author: Chunhong Qi (chunhongqijoy@sina.com)

This work was supported by the 2022 Ministry of Education Humanities and Social Sciences Research General Project: A Comparative and Cognitive Study of English-Chinese Ellipses for Machine Translation under Grant 22YJA740033.

ABSTRACT This paper proposes a hybrid deep learning approach combining Bidirectional Long Short-Term Memory (BiLSTM) networks and an attention mechanism to extract sentiment and thematic content from literary texts. The model is designed to capture complex emotional nuances and themes in literature by processing text data from both forward and backward directions, while the attention mechanism enables the model to focus on the most important sections of the text. Hyperparameter optimization is performed using the Improved Particle Swarm Optimization (IPSO) algorithm to fine-tune the model for efficient sentiment extraction. A case study using a dataset of 500 English novels spanning various genres demonstrated the effectiveness of the proposed approach. The model achieved high accuracy and F1 scores in sentiment classification and thematic extraction, outperforming traditional methods like CNNs. The analysis revealed key emotional themes such as joy, fear, and sorrow, and the thematic content included love, betrayal, and revenge. The results highlight the potential of deep learning to advance literary analysis by providing deeper insights into both emotional and thematic layers of literary works. Future directions include exploring multimodal data integration and expanding the application of deep learning in the humanities.

INDEX TERMS Deep learning, sentiment analysis, bidirectional LSTM (BiLSTM), attention mechanism, thematic extraction, improved particle swarm optimization (IPSO), text preprocessing, machine learning, literary analysis, thematic content classification.

I. INTRODUCTION

This paper explores the application of deep learning models for sentiment analysis in the realm of English literature, focusing on the extraction of emotional and thematic content. Reference [1] is the sentiment analysis, a subfield of Natural Language Processing (NLP), involves the use of machine learning techniques to analyze and classify the sentiment or emotional tone of a text. In the context of literary texts [2], sentiment analysis not only involves identifying positive or negative emotions but also understanding the underlying thematic elements and emotional subtleties that literature presents. The integration of machine learning-based sentiment analysis into literary studies holds tremendous potential, as it offers a way to systematically analyze complex literary content and extract insights that would traditionally require manual interpretation. Sentiment analysis has become increasingly important in various fields, including marketing, social media analysis, customer feedback systems, and even political discourse. However, its application in literary studies is less explored, particularly when it comes to analyzing works of fiction, poetry, and drama. Literary texts often contain layered emotions, complex character interactions, subtext, and cultural references that make them difficult to

The associate editor coordinating the review of this manuscript and approving it for publication was Sotirios Goudos.

analyze using traditional methods [3], [4]. This complexity requires a sophisticated approach, capable of capturing not only surface-level emotions but also the deeper, often implicit sentiments conveyed through metaphor, irony, and other literary devices. The need for more advanced models to analyze literary content is driven by the intricate nature of emotional and thematic development in literature. Classic texts such as William Shakespeare's Hamlet or novels by authors like Jane Austen or Charles Dickens are rich with nuanced expressions of emotions, philosophical explorations, and social commentaries. Analyzing such works manually can be time-consuming and subjective. Identifying themes such as existentialism, revenge, or social critique across vast bodies of text requires the ability to understand and process both the emotional tone and thematic structures embedded within the text. Machine learning and deep learning models, which have shown immense success in handling large-scale text data, provide an ideal solution to address these challenges [5]. Deep learning techniques, particularly Recurrent Neural Networks (RNNs) and their advanced variants, such as Long Short-Term Memory (LSTM) networks, have proven effective in handling sequential data, making them ideal for analyzing textual data. LSTMs, and in particular Bidirectional LSTMs (BiLSTMs), have the ability to capture long-term dependencies in text by processing data in both forward and backward directions [6]. This capability is essential in the analysis of literary texts, where context from both the preceding and succeeding words in a sentence often contributes to the emotional and thematic interpretation. BiLSTMs can capture subtle shifts in tone and sentiment that are often present in complex literary works. Despite the effectiveness of BiLSTMs, there are still challenges in focusing on specific parts of a text that carry the most emotional or thematic weight. This is where the attention mechanism comes into play. Attention mechanisms allow models to weigh different parts of the input text differently, enabling them to focus on the most important segments of the text for a given task [7]. This helps in improving the interpretability of the model, as it provides insights into which words or phrases contribute most to the sentiment classification. For example, in analyzing Hamlet, the attention mechanism can highlight words in key soliloquies or moments of high emotional intensity, such as "To be or not to be," that are crucial to understanding the underlying sentiment of the work. However, even with advanced models like BiLSTM and attention mechanisms, the performance of these models heavily depends on the optimization of their hyperparameters [8]. Hyperparameter tuning is a critical aspect of training deep learning models, and poor optimization can result in suboptimal performance. This is where Particle Swarm Optimization (PSO), particularly an improved version like the Improved Particle Swarm Optimization (IPSO), becomes invaluable [9]. IPSO helps in efficiently finding the optimal set of hyperparameters by simulating a population of potential solutions and dynamically adjusting them during training. This results in better convergence rates and more accurate models, which is especially important when analyzing complex literary texts with varied emotional and thematic content.

Sentiment analysis using machine learning techniques has gained significant traction in various research domains, particularly for analyzing social media data [10]. Several studies have explored different applications and challenges in this area [11]. For instance, [12] provided a comprehensive review of machine learning-based sentiment analysis techniques applied to COVID-19-related Twitter data, highlighting the effectiveness of ensemble models, particularly BERT and RoBERTa, for accurate sentiment classification. Similarly, [13] examined sentiment analysis in Arabic, a language with unique challenges due to its morphology and dialects, revealing that deep learning techniques, particularly multi-level embeddings, outperformed other approaches. In the context of fake reviews, [14] identified sentiment analysis as a key tool in detecting deceptive reviews, emphasizing the importance of machine learning for enhancing detection accuracy. Additionally, [15] reviewed ensemble learning techniques, which combine multiple models to improve performance, noting their widespread application in machine learning. Reference [16] explored the use of AI for improving the quality of life (QoL) studies, showing that AI methods could help address complexities in QoL research. Other studies, like those by [17] and [18] focused on sentiment analysis in low-resource settings and opinion spam detection, respectively, both stressing the importance of deep learning and transfer learning in these specialized domains. Reference [19] reviewed Arabic sarcasm detection, highlighting the challenges and advancements in natural language processing for detecting sarcasm in Arabic texts. Together, these studies illustrate the evolving landscape of sentiment analysis across various languages and platforms, driven by the integration of deep learning techniques to address domain-specific challenges.

The objective of this paper is to present a hybrid deep learning approach that combines BiLSTM, attention mechanisms, and IPSO to analyze the emotional and thematic content of literary texts. By leveraging the strengths of each of these components, the proposed approach aims to provide a more robust, accurate, and interpretable model for sentiment analysis in literature. This hybrid approach not only enhances the accuracy of sentiment classification but also enables a deeper understanding of the themes that emerge from the text, facilitating a more comprehensive analysis of literary works. In doing so, this paper aims to bridge the gap between machine learning techniques and literary analysis, paving the way for further research into the application of AI in the humanities. The proposed sentiment and emotion classification model can have a significant impact on literary research by enabling several advanced applications, including automatic summarization and literary style analysis. One of the most promising applications is automatic summarization, where the model could generate concise yet comprehensive summaries of literary works based on the sentiments and emotions expressed within the text. By analyzing the emotional tone and sentiment shifts throughout a literary piece, the model could help create summaries that capture the essence of the story's emotional journey, giving readers or researchers a quick yet insightful overview. This could be particularly valuable in analyzing large corpora of literature, allowing scholars to quickly identify key themes and emotional arcs without having to read each work in its entirety. Another practical application of the model is in literary style analysis, where the model can be used to analyze an author's unique style in expressing emotions and sentiments. By leveraging the deep learning algorithms to understand the patterns and variations of sentiment across different works or sections of a work, the model can uncover distinctive stylistic features such as the use of specific emotional language, tone shifts, or sentiment-related devices. Researchers can apply this analysis to compare the writing styles of different authors, trace the evolution of an author's emotional expression over time, or even assess how literary style correlates with broader societal or historical changes. The model could aid in cross-cultural literary analysis, helping to compare how different cultures and time periods express similar emotions through different narrative techniques or linguistic structures. The ability to classify sentiments and emotions accurately across texts opens up new avenues for exploring global literary traditions and understanding how universal themes are conveyed through diverse literary forms. The practical applications of this sentiment and emotion classification model can substantially enrich literary research by providing tools for both efficient text analysis and deeper, more nuanced interpretation of literary works.

II. RELATED WORK

Recent advancements in sentiment analysis have been marked by diverse methodologies, applications, and challenges across multiple languages and domains. Reference [4] emphasize the need for better datasets and techniques in Arabic aspect-based sentiment analysis, while [5] highlights the gap in sentiment analysis research for Russian, offering a roadmap for expanding studies in this language. Reference [6] focus on multilingual sentiment analysis for under-resourced languages, stressing the potential of deep learning and the integration of high-resource languages to improve models for low-resource languages. Reference [7] presents a novel combination of recurrent and recursive neural networks for aspect-based sentiment analysis, demonstrating significant improvements over baseline models. Reference [8] explores sentiment dynamics, proposing an index for industry-specific sentiment trends. Reference [9] address noisy labels in sentiment analysis using a two-level LSTM model with lexicon embedding, demonstrating its effectiveness on English and Chinese datasets. Reference [10] reviews Arabic sentiment analysis, highlighting the role of deep learning and multi-level embeddings in enhancing accuracy. Reference [11] investigates sentiment analysis on COVID-19-related Twitter data, finding that ensemble models like BERT and RoBERTa offer optimal performance for public health decision-making. Reference [12] examines the influence of electronic word-of-mouth (eWOM) in consumer decision-making, while [13] tackle cross-lingual sentiment analysis, achieving state-of-the-art results with dynamic weighting and focal loss. Reference [14] proposes a deep learning-based model for Arabic sentiment analysis, incorporating multilevel parallel attention to improve contextual embeddings. Reference [15] introduces a BiGRU attention classifier for sentiment analysis of scientific texts, outperforming traditional models in accuracy and recall. Reference [16] develops an AI-based system for global event analysis, improving media monitoring through sentiment analysis and anomaly detection. Reference [16] also proposes a blockchain-based distributed deep learning model for public sentiment analysis, achieving enhanced efficiency and accuracy. Reference [17] explore transfer learning for unsupervised sentiment lexicon learning across multiple domains. Reference [18] introduce a Modified Switch Transformer model for sarcasm detection in Arabic, outperforming prior models in handling sentiment nuances. Reference [19] proposes an attention-based multi-channel Gated Recurrent Neural Network (Att-MC-GRU) for aspect-based sentiment analysis, enhancing aspect extraction and sentiment classification with word embedding, part of speech tags, and contextual information. Reference [20] introduces AspEntQuaNet, a method for aspect-based sentiment quantification, offering significant improvements in accuracy by considering ternary sentiment and implementing an entropy-based sorting procedure. Reference [21] tackles the challenge of large-scale, unlabeled datasets by combining frequency-based and syntactic-relation-based techniques, augmented with semantic similarity, to extract relevant aspects from user reviews. Reference [22] examines how the visual sentiment of mobile ad images and device attributes affect mobile ad response behavior, while [23] review sentiment analysis applications in assessing health technologies, calling for more automated tools. Reference [23] also review sentiment analysis techniques for public social media, particularly in the context of mental health degradation during the COVID-19 pandemic. These studies collectively underscore the evolving nature of sentiment analysis methodologies and applications, from multilingual sentiment analysis to its impact on public health and consumer behavior, demonstrating its increasing significance in various fields.

III. PROBLEM STATEMENT

Sentiment analysis in English literature presents a unique set of challenges due to the nuanced and complex nature of literary texts. Unlike typical datasets used in sentiment analysis, such as social media or product reviews, literary works often contain layered emotions, subtle subtext, and complex thematic elements such as irony, ambiguity, and metaphor. Extracting both emotional and thematic content from these texts requires advanced machine learning and deep learning models capable of understanding these intricacies. Traditional sentiment analysis models often struggle to capture the full spectrum of sentiments embedded in literature. These models are generally designed to detect explicit sentiments like happiness, sadness, or anger, which are straightforward in more direct, conversational texts [20]. However, literary texts often present emotions indirectly, through metaphors, symbolism, and complex character development. Therefore, these conventional models may fail to distinguish between explicit sentiments, such as joy or sorrow, and implicit ones, such as irony or sarcasm, which are integral to understanding the underlying emotional tone of a literary work. In this context, we can express the relationship between explicit sentiment accuracy and implicit sentiment detection as follows:

$$\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}\mathbb{I}\left(y_i=\hat{y}_i\right)\right]\geq\alpha,\quad 0\leq\alpha\leq 1 \tag{1}$$

where $\mathbb{I}(\cdot)$ is the indicator function that determines whether the predicted sentiment $\hat{y}_i$ matches the true sentiment y_i, and α is the threshold for acceptable model accuracy. However, literary works often present these sentiments indirectly, making it challenging for the model to interpret them correctly. The thematic content of literary works goes beyond simple sentiment detection. Thematic elements such as existentialism, social justice, or love are often not directly stated but inferred from context, character arcs, and narrative devices. Thematic elements are complex, multifaceted, and interwoven with the plot, requiring the model to analyze the relationships between different story components. In this regard, deep learning models need to handle a higher level of abstraction. For example, the relationship between hidden layer neurons in the model can be constrained by the following equation to maintain sufficient capacity while avoiding overfitting:

$$\sum_{h=1}^{H}\|\mathbf{w}_h\|_2^2\leq\lambda_H,\quad N_{min}\leq N_h\leq N_{max} \tag{2}$$

where H is the number of hidden layers, and λ_H is the regularization parameter that limits the L2-norm of the weights $\mathbf{w}_h$ to prevent overfitting, while ensuring the model has enough capacity to handle complex themes. The challenge in capturing implicit emotions and thematic complexity is further exacerbated by the fact that many literary works are written in a way that demands interpretation—capturing the nuances of irony, subtext, and ambiguity. Detecting these subtle nuances requires the model to focus on specific parts of the text that convey underlying meanings. This can be addressed by using an Attention Mechanism in conjunction with Bidirectional LSTM (BiLSTM) networks. The attention mechanism assigns varying importance to different words or phrases within the text, allowing the model to focus on the most semantically rich components. The attention weight matrix $\mathbf{A}(\mathbf{X})$, which captures this focus, should satisfy the following condition to ensure adequate focus on key sections of the text:

$$\mathbb{E}_{\mathcal{D}}\left[\|\mathbf{A}(\mathbf{X})\|_F\right]\geq\delta \tag{3}$$

where $\|\mathbf{A}(\mathbf{X})\|_F$ is the Frobenius norm of the attention matrix, ensuring that the model places adequate focus on the crucial components for emotional and thematic analysis. The optimization of the model's hyperparameters is equally important to ensure that the model remains efficient and converges to the best possible solution. To fine-tune these parameters, techniques like Improved Particle Swarm Optimization (IPSO) can be employed. IPSO refines the basic PSO algorithm by adjusting particle velocities dynamically, thereby enhancing convergence speed and improving the model's efficiency. The optimization process can be described as follows, ensuring that the model's hyperparameters converge within a specified range:

$$\nabla_\eta\mathcal{L}(\eta)\Big|_{\eta=\eta^*}=0,\quad \eta_{min}\leq\eta^*\leq\eta_{max} \tag{4}$$

where η^* represents the optimal learning rate, and the gradient of the loss function $\mathcal{L}(\eta)$ must be zero at this point to ensure stable convergence. In terms of model stability, we aim to keep the second-order derivative of the loss function bounded, ensuring the model's robustness during training. This is captured by the following stability constraint, which integrates the second derivative of the loss function:

$$\int_{\eta_{min}}^{\eta_{max}}\left(\frac{\partial^2\mathcal{L}(\eta)}{\partial\eta^2}\right)^2 d\eta\leq\epsilon_{\text{stability}},\quad \text{subject to}\quad \mathbb{E}_{\mathcal{D}}[\mathcal{L}(\hat{\theta})]\leq\epsilon \tag{5}$$

This equation ensures that the optimization process is stable, minimizing the likelihood of overfitting while maintaining model generalization. To formulate a two-layered objective function for the hybrid deep learning model aimed at sentiment analysis and thematic extraction in literary texts, we can combine two key objectives: (1) optimizing sentiment classification accuracy and (2) enhancing thematic interpretation and emotional nuance detection. This can be achieved by constructing a compound objective function that integrates both layers with a balanced trade-off between them. The first layer focuses on optimizing the sentiment classification performance. This involves minimizing a loss function such as cross-entropy or mean squared error, which quantifies the difference between the predicted sentiment labels $\hat{y}$ and the true labels y. The objective is to maximize sentiment detection accuracy, including both explicit and implicit emotional content. The first-layer objective function can be expressed as:

$$\mathcal{L}_1=-\mathbb{E}_{\mathcal{D}}\left[\sum_{c\in\mathcal{C}}y_c\log(\hat{y}_c)\right] \tag{6}$$

where y is the true sentiment label. $\hat{y}$ is the predicted sentiment output from the model. $\mathcal{C}$ is the set of sentiment classes (e.g., positive, negative, neutral). $\mathbb{E}_{\mathcal{D}}\left[\sum_{c\in\mathcal{C}} y_c \log(\hat{y}_c)\right]$ represents the expected cross-entropy loss between the true and predicted sentiment distributions. Thus, the first-layer objective aims to minimize this loss, ensuring that the model identifies explicit emotional content correctly. The second layer involves enhancing the model's ability to detect implicit emotional nuances and complex thematic elements, such as irony, subtext, and ambiguity. This objective layer ensures the model can capture deeper, more abstract patterns beyond surface-level sentiment, focusing on how well the model interprets the themes and emotions embedded in the text. [21] This can be represented by a regularization term, where the model's thematic interpretation is evaluated based on its capacity to capture important aspects like subtext. An example of a potential loss function for this layer is based on the attention mechanism, where we aim to focus the model's attention on key thematic or emotional elements:

$$\mathcal{L}_2 = \mathbb{E}_{\mathcal{D}}\left[\|\nabla_{\mathbf{z}}\mathcal{T}(\mathbf{z})\|_2^2 + \alpha\|\mathbf{A}(\mathbf{X})\|_F^2\right] \tag{7}$$

where $\mathbf{z}$ is the vector representing the thematic or emotional interpretation of the text. $\mathcal{T}(\mathbf{z})$ is a function that maps the feature space to thematic elements. $\nabla_{\mathbf{z}}\mathcal{T}(\mathbf{z})$ is the gradient of the thematic representation with respect to $\mathbf{z}$. $\mathbf{A}(\mathbf{X})$ is the attention weight matrix associated with the input features $\mathbf{X}$, highlighting the importance of certain words or phrases. $\|\mathbf{A}(\mathbf{X})\|_F$ is the Frobenius norm of the attention matrix, capturing the strength of attention. α is a hyperparameter balancing the influence of the attention regularization. This second-layer loss function encourages the model to focus on the right sections of the text while ensuring that the extracted themes and emotions align with the deeper, implicit structures in the text. The overall objective function combines both layers, where a weight λ is introduced to balance the importance of sentiment classification accuracy and thematic/emotional nuance detection. The combined objective function becomes:

$$\mathcal{L}_{\text{total}} = \lambda\mathcal{L}_1 + (1-\lambda)\mathcal{L}_2 \tag{8}$$

where $\lambda \in [0, 1]$ is a hyperparameter that controls the trade-off between the two layers of the objective function. A higher value of λ places more emphasis on sentiment accuracy, while a lower value prioritizes thematic interpretation and emotional nuance. $\mathcal{L}_1$ represents the first layer's sentiment classification loss. $\mathcal{L}_2$ represents the second layer's thematic and emotional nuance loss. Thus, the model is trained to minimize this total objective function, ensuring that it performs well in both detecting explicit emotions and capturing implicit emotional and thematic content in literary texts. This two-layered approach provides a comprehensive mechanism for improving sentiment analysis and thematic extraction in a way that reflects the rich and complex nature of literary works.

IV. HYBRID DEEP LEARNING APPROACH

The hybrid deep learning approach combines Bidirectional Long Short-Term Memory (BiLSTM) and attention mechanisms to optimize sentiment analysis and thematic extraction from literary texts. BiLSTM plays a critical role in processing text data by capturing context from both the forward and backward directions. This bidirectional approach allows the model to better understand long-term dependencies and complex sentence structures, which are essential in literary works where meanings often rely on context and prior sentences [22]. The hybrid deep learning approach, integrating Bidirectional Long Short-Term Memory (BiLSTM) and attention mechanisms, can be effectively applied to address the complexities of fine-grained sentiment analysis, metaphor comprehension, and cultural-context-aware literary text analysis. Fine-grained sentiment analysis involves detecting subtle shifts in emotional tones and identifying specific sentiment categories within literary works, such as joy, sorrow, anger, or irony. Given the layered emotional structures often present in literature, the bidirectional nature of BiLSTM helps capture long-term dependencies, allowing the model to process these complex emotional transitions across different sections of a text. By analyzing the context from both the past and future, BiLSTM can identify the fine-grained sentiment that is pivotal to understanding intricate emotional arcs and character development in literature. Metaphor comprehension poses a unique challenge for sentiment and thematic analysis, as metaphors often require a deep understanding of figurative language. The attention mechanism enhances the model's ability to focus on key words or phrases that carry metaphorical meaning, enabling the system to detect and interpret these figurative expressions more effectively. Through careful attention to context and the way certain words shift the meaning of the sentence, the model can uncover implicit sentiments embedded within metaphors, thus improving its capacity to analyze non-literal language in literary texts. Cultural-context-aware analysis is crucial for understanding how sentiments and themes are expressed differently across cultures. The model's preprocessing techniques, such as tokenization and text normalization, can be adapted to handle diverse linguistic structures and historical language usage, which are especially relevant when analyzing texts from different cultural or historical contexts. By incorporating a custom dictionary to handle archaic terms or genre-specific characteristics, the model can better interpret culturally specific nuances, ensuring a more accurate sentiment and thematic analysis. This approach offers significant advantages in studying literature from various time periods or regions, enabling a richer, more contextually aware understanding of emotional and thematic content. By leveraging this ability, BiLSTM excels in analyzing intricate emotional content, including shifts in sentiment and underlying emotional tones, which are prevalent in literature. The forward pass through the BiLSTM layer captures dependencies from the past (left context). Given an input sequence $\mathbf{X} = \{x_1, x_2, \ldots, x_T\}$, the forward hidden state at time t is computed as:

$$\mathbf{h}_t^{\text{forward}} = \text{LSTM}_{\text{forward}}(x_t, \mathbf{h}_{t-1}^{\text{forward}}, \mathbf{c}_{t-1}^{\text{forward}}) \tag{9}$$

Similarly, the backward pass captures dependencies from the future (right context). The backward hidden state at time t is computed as:

$$\mathbf{h}_t^{\text{backward}} = \text{LSTM}_{\text{backward}}(x_t, \mathbf{h}_{t+1}^{\text{backward}}, \mathbf{c}_{t+1}^{\text{backward}}) \quad (10)$$

After computing the forward and backward hidden states, the final representation of the sequence at each time step is obtained by concatenating both directions:

$$\mathbf{h}_t = \left[\mathbf{h}_t^{\text{forward}}, \mathbf{h}_t^{\text{backward}}\right] \quad (11)$$

In conjunction with BiLSTM, the attention mechanism enhances the model's capacity to focus on key phrases that carry significant emotional and thematic weight. Attention allows the model to weigh different parts of the text differently, highlighting the sections that are most relevant to sentiment classification. The sentiment classification at time step t is computed as:

$$\hat{y}_t = \text{Softmax}(\mathbf{W}_s \mathbf{h}_t + \mathbf{b}_s) \quad (12)$$

For thematic extraction, the thematic representation of the input text is derived from the BiLSTM output:

$$\mathbf{z}_t = \mathbf{W}_t \mathbf{h}_t + \mathbf{b}_t \quad (13)$$

This selective focus not only improves accuracy but also enhances interpretability, as it provides insight into which words or phrases contribute most to the sentiment detected. Visualization of attention weights further enables a deeper understanding of the model's decision-making process, ensuring transparency in how emotional nuances and themes are identified within complex literary texts. The proposed hybrid deep learning model, combining BiLSTM and attention mechanisms, addresses the abstraction and contextual variability associated with nuanced themes in literary texts through its architecture and preprocessing techniques. BiLSTM's bidirectional processing allows the model to capture long-term dependencies, crucial for understanding themes that span across different parts of the text. This is especially important in literary works where meaning is often dependent on context and prior sentences. The attention mechanism enhances the model's ability to focus on significant words or phrases, enabling it to better understand emotional shifts and thematic nuances, which are inherent in complex literary narratives. To further enhance the detection of such nuanced themes, the model's architecture utilizes the attention weights to highlight key sections of the text that contribute to sentiment and thematic extraction, providing transparency in its decision-making process. The preprocessing steps play a critical role in ensuring the uniformity of the data and addressing issues arising from historical language usage and diverse narrative styles. These steps include tokenization, text normalization, removal of stopwords, and the incorporation of a custom dictionary to handle archaic terms. Additionally, genre-specific preprocessing strategies are applied to adjust for the typical linguistic structures of different literary forms, such as novels, poetry, or drama. These combined architectural and preprocessing techniques allow the model to effectively handle the variability and abstraction inherent in literary themes, ensuring that both explicit and subtle elements of sentiment and thematic content are captured accurately.

V. OPTIMIZATION WITH IMPROVED PARTICLE SWARM OPTIMIZATION (IPSO)

The Improved Particle Swarm Optimization (IPSO) algorithm is a powerful optimization technique that enhances the traditional Particle Swarm Optimization (PSO) by dynamically adjusting particle velocity and adapting the search space to improve the optimization of deep learning models. IPSO optimizes hyperparameters such as learning rate, batch size, and the number of hidden units in neural networks, which are critical in fine-tuning models for sentiment analysis and thematic extraction in literary texts [23]. By iteratively updating the particles based on the best-known positions and velocities, IPSO searches for the optimal set of hyperparameters that maximize model performance, ensuring the model can effectively capture both explicit and implicit emotional content in literary works. In the traditional PSO algorithm, the particle velocity is updated based on two components: the cognitive component (personal best position) and the social component (global best position). The update rule is typically given by:

$$v_i^{k+1} = wv_i^k + c_1 r_1 (p_i^* - x_i^k) + c_2 r_2 (g^* - x_i^k) \quad (14)$$

where v_i^{k+1} is the updated velocity of particle i at iteration $k + 1$. w is the inertia weight, controlling the impact of previous velocity. c_1 and c_2 are acceleration constants. r_1 and r_2 are random numbers between 0 and 1. p_i^* is the best-known position of particle i. g^* is the global best position found across all particles. However, the IPSO algorithm improves upon this by introducing a dynamic adjustment of the velocity update and incorporating an adaptive mechanism to change the exploration-exploitation balance. The updated velocity equation in IPSO becomes:

$$\begin{aligned} v_i^{k+1} &= w_k v_i^k + c_1 r_1 (p_i^* - x_i^k) \\ &\quad + c_2 r_2 (g^* - x_i^k) + \gamma (r_3 x_i^k - x_i^{k-1}) \end{aligned} \quad (15)$$

where w_k is the time-varying inertia weight at iteration k. γ is a constant that controls the adaptive adjustment of particle velocity based on prior velocity. This dynamic velocity update allows IPSO to better explore the search space and fine-tune the hyperparameters more effectively than traditional PSO, leading to better convergence on optimal values for deep learning models. In the context of sentiment analysis in literary texts, IPSO can be employed to fine-tune the hyperparameters of a deep learning model, such as the learning rate, batch size, and the number of hidden units in the BiLSTM or other deep architectures. This optimization ensures the model is both efficient and accurate in detecting sentiments and thematic content, which is essential for processing complex, nuanced literary data. For hyperparameter optimization in the IPSO framework, the fitness function F evaluates the performance of the model using a specific set of hyperparameters:

$$F(\mathbf{X}) = \frac{1}{N}\sum_{i=1}^{N} \text{CrossEntropy}(y_i, \hat{y}_i) \tag{16}$$

where $F(\mathbf{X})$ is the fitness function for a given particle's position, representing the model's loss (cross-entropy loss in this case). y_i is the true sentiment label for the i-th instance. $\hat{y}_i$ is the predicted sentiment output for the i-th instance. N is the number of instances in the dataset. By minimizing the fitness function, IPSO optimizes the hyperparameters to improve the model's ability to classify sentiment and extract thematic content from the text, ultimately enhancing the model's performance for literary sentiment analysis. This optimization process not only enhances model performance by finding the best set of hyperparameters but also ensures that the model can efficiently capture both the explicit sentiments and the subtle, underlying thematic elements in literary works. The hyperparameters optimized using the Improved Particle Swarm Optimization (IPSO) algorithm are essential for enhancing the performance of deep learning models, particularly in sentiment analysis and thematic extraction from literary texts. Key hyperparameters include the learning rate, batch size, and the number of hidden units in the BiLSTM model. The batch size, tested between 16 and 128, affects the number of training samples processed before updating the weights, influencing both training stability and memory usage. The number of hidden units, optimized within a range of 50 to 500, dictates the model's ability to capture long-term dependencies in the text, ensuring it can learn complex relationships without overfitting. Additional hyperparameters such as the number of layers, dropout rate, and momentum term are also considered during optimization. IPSO iteratively adjusts these hyperparameters by evaluating the model's performance using cross-entropy loss as a fitness function, updating particle positions based on personal and global bests while incorporating a dynamic adjustment mechanism to enhance exploration and exploitation. This process results in the optimal set of hyperparameters that maximize the model's ability to capture both explicit and implicit sentiments and thematic elements in literary works.

VI. PROPOSED FRAMEWORK FOR SENTIMENT EXTRACTION

The proposed framework for sentiment extraction leverages a hybrid architecture combining Bidirectional Long Short-Term Memory (BiLSTM) networks and an attention mechanism. This architecture is well-suited for extracting both emotional and thematic content from literary texts, as it can capture complex contextual relationships and subtle emotional nuances inherent in literature. The BiLSTM model processes text data from both forward and backward directions, ensuring that long-term dependencies and sentence structures are fully understood. The attention mechanism further enhances this by allowing the model to focus on the most relevant parts of the text, which are essential for sentiment and thematic extraction [24], [25]. The attention mechanism assigns different importance weights to different words or phrases, effectively highlighting the sections of the text that most contribute to the sentiment classification and thematic extraction. This mechanism is particularly useful in literary texts, where meaning is often embedded in subtle linguistic cues and context. By visualizing the attention weights, we can identify the key elements of the text that influence sentiment classification, providing valuable insights into the model's decision-making process. The sentiment extraction process can be mathematically expressed using the following equation, where the attention mechanism helps identify the most relevant parts of the text for sentiment and thematic extraction. Let $\mathbf{h}$ be the hidden state output of the BiLSTM network for the input text $\mathbf{X}$, and $\mathbf{A}(\mathbf{X})$ be the attention weight matrix. The output sentiment classification and thematic extraction can be represented as:

$$\hat{y} = \text{Softmax}\left(\mathbf{W}_s \sum_{i=1}^{T} \alpha_i \mathbf{h}_i\right) \tag{17}$$

where $\hat{y}$ is the predicted sentiment output. $\mathbf{W}_s$ is the weight matrix for sentiment classification. α_i is the attention weight for the i-th hidden state, highlighting the importance of specific words/phrases in the text. T is the total number of time steps in the BiLSTM output. $\mathbf{h}_i$ is the hidden state for the i-th word in the sequence. The Softmax function normalizes the output to produce a probability distribution over possible sentiment classes. Hyperparameter optimization is crucial to the performance of the BiLSTM model. The Improved Particle Swarm Optimization (IPSO) algorithm is employed to fine-tune the model's hyperparameters, such as learning rate, batch size, and the number of hidden units. IPSO dynamically adjusts particle velocity, ensuring efficient exploration of the hyperparameter space and enabling the model to converge to an optimal set of parameters for sentiment extraction. The optimization process in IPSO is defined by the following velocity update equation, where the velocity of each particle is adjusted dynamically:

$$\begin{aligned} v_i^{k+1} &= w_k v_i^k + c_1 r_1 \left(p_i^* - x_i^k\right) \\ &\quad + c_2 r_2 \left(g^* - x_i^k\right) + \gamma \left(r_3 x_i^k - x_i^{k-1}\right) \end{aligned} \tag{18}$$

where v_i^k is the velocity of particle i at iteration k. w_k is the time-varying inertia weight. c_1 and c_2 are acceleration constants. r_1, r_2, and r_3 are random values between 0 and 1. p_i^* is the personal best position for particle i. g^* is the global best position. γ controls the adaptation of particle velocity based on prior velocities, enhancing the exploration-exploitation balance. The sentiment extraction process begins with tokenizing the input text, followed by lemmatization and text cleaning. The preprocessed text is then passed through the BiLSTM model, which classifies the sentiments (such as joy, sorrow, anger, and fear) and extracts thematic elements (such as revenge, existentialism, and mortality). The attention mechanism identifies the most influential words and phrases for sentiment and thematic classification. In the experiment, dataset preprocessing is a critical step to ensure uniformity and enhance the quality of input data for the deep learning model. Given the diversity in literary works—spanning different genres, narrative styles, and historical periods—it is essential to address the inherent variations in writing styles and linguistic usage. The preprocessing pipeline includes several key steps to standardize and clean the data for optimal model performance. Initially, the text data is tokenized into smaller units such as words or subwords using a tokenizer like WordPiece or SentencePiece, which ensures that the model can handle diverse vocabulary efficiently. To address the issue of uniformity, text normalization techniques are applied, including lowercasing, removal of special characters, punctuation, and extra whitespace, which helps eliminate inconsistencies and reduce the complexity of the text. Additionally, stopwords—common words that do not contribute significantly to sentiment or thematic analysis—are filtered out, which improves the model's focus on meaningful content. To account for historical language usage and varying narrative styles, a custom dictionary of archaic or period-specific terms is incorporated to ensure that these words are not misunderstood or ignored by the model. This dictionary helps map older forms of language to modern equivalents or provides the model with contextual understanding of such terms. Furthermore, texts with extreme narrative deviations—such as poetry or highly figurative language—are subjected to special preprocessing to extract their emotional and thematic content more effectively, using methods like part-of-speech tagging and dependency parsing to capture syntactic structures. To ensure consistency across different genres, a genre-specific preprocessing approach is adopted, where texts from different genres (e.g., drama, novels, poetry) are preprocessed differently based on their typical narrative structures and linguistic features. Finally, the text is vectorized using word embeddings or character embeddings to capture semantic meaning and relationships between words, ensuring that the model can effectively process the varied linguistic elements present in the dataset. This comprehensive preprocessing pipeline ensures uniformity in the dataset and mitigates the challenges posed by different narrative styles and historical language usage, thereby improving the accuracy and reliability of sentiment and thematic extraction from literary texts.

A. CASE STUDY: SENTIMENT AND THEMATIC ANALYSIS OF ENGLISH LITERARY WORKS USING DEEP LEARNING MODELS

The case study focused on applying the proposed framework for sentiment extraction to a dataset of 500 English literary works, ranging from classic to contemporary novels across various genres, including romance, tragedy, and satire. The novels were pre-processed for tokenization, stemming, and stopword removal, ensuring that the input to the deep learning models was consistent and suitable for sentiment and thematic analysis. Each novel was annotated for sentiment polarity (positive, negative, neutral) and emotional themes (joy, sadness, fear, anger, surprise) by literary experts, along with thematic content (love, betrayal, revenge), aligning with the data pre-processing approach detailed in the proposed framework. The hybrid architecture combining Bidirectional Long Short-Term Memory (BiLSTM) networks and an attention mechanism, as described in the framework, was used for sentiment classification and thematic extraction. The BiLSTM model processed the text data in both forward and backward directions, capturing long-term dependencies and sentence structures. The attention mechanism then highlighted key words and phrases that contributed most to sentiment classification and thematic extraction, aligning with the methodology explained in the proposed framework. Additionally, the hyperparameters of the BiLSTM model were optimized using the Improved Particle Swarm Optimization (IPSO) algorithm, ensuring efficient sentiment extraction. The optimization process was carried out by dynamically adjusting parameters such as learning rate, batch size, and the number of hidden units, improving the overall model performance. The analysis revealed that the sentiment classification model achieved an accuracy of 85%, with the most common emotional themes being joy, fear, and sorrow. The thematic content included recurring topics such as love, conflict, and morality, which aligned with genre-specific sentiment scores. Using the attention mechanism, the most influential parts of the text were visualized, showing how the model focused on emotionally significant sections of the novels. The F1-score for thematic content extraction was 0.78, demonstrating the framework's effectiveness in analyzing both emotional and thematic aspects of literary works. This case study demonstrates the practical application of the proposed sentiment extraction framework, showcasing the power of combining BiLSTM, attention mechanisms, and IPSO for deep analysis of literary texts. The results highlight the framework's ability to provide deep insights into emotional and thematic content, aligning with the goals of the paper to leverage deep learning models for sentiment and thematic extraction. The model's performance across different literary genres, such as romance and tragedy, can vary significantly due to the inherent differences in narrative structures, emotional tones, and linguistic styles associated with each genre. For instance, in romance novels, the emotional content often revolves around positive sentiments such as love, joy, and hope, which may be more straightforward for the model to detect using BiLSTM and attention mechanisms. These genres tend to feature more consistent and repetitive expressions of these emotions, which the model can learn to identify effectively through its ability to capture long-term dependencies in text. In contrast, genres like tragedy or existential literature often involve complex emotional undertones, including sorrow, despair, or irony, which may require a more nuanced understanding of context and themes.

Tragedy, for example, might involve significant emotional shifts, contradictions, and subtle expressions of underlying grief or conflict, which can pose challenges for the model. The BiLSTM's bidirectional nature helps in capturing these shifts by analyzing both past and future contexts, but the emotional complexity of tragedy may still lead to lower performance if the model fails to fully capture these subtleties. The attention mechanism helps in focusing on the most relevant parts of the text, but genres with highly figurative language, such as poetry or abstract drama, may present additional challenges. The figurative nature of these genres, including metaphor, symbolism, and ambiguous language, could lead to poorer model performance due to difficulties in semantic interpretation. The preprocessing pipeline, however, can help address these challenges by adapting techniques like part-of-speech tagging and dependency parsing, but it may still struggle with the deeper, more abstract language. Factors contributing to these variances include:

- Linguistic Complexity: Tragedy and poetry often use figurative language and less direct expressions of emotion, which might be harder for the model to understand.
- Emotional Variability: Genres with less consistent emotional content (e.g., tragedy) present challenges in sentiment detection due to rapid emotional shifts or layered meanings.
- Narrative Structure: Some genres, such as historical or experimental literature, may feature non-linear storytelling or unconventional sentence structures, complicating the detection of long-term dependencies.
- Genre-Specific Vocabulary: Certain genres might contain domain-specific vocabulary (e.g., archaic terms in historical novels) that the model must account for to achieve accurate analysis.

While the hybrid BiLSTM-attention model performs well across various genres, it tends to be more effective in genres with consistent emotional tones (like romance) and less effective in those with complex, figurative, or rapidly changing emotional landscapes (like tragedy or poetry). Tailoring the preprocessing pipeline and fine-tuning the model through domain-specific adjustments can help mitigate some of these challenges and improve performance across diverse genres.

VII. EXPERIMENTAL SETUP AND RESULTS

The dataset for the experiment consisted of 500 English novels, ranging from classic works to contemporary literature, covering various genres such as romance, tragedy, satire, and historical fiction. The total word count of the dataset exceeded 3 million words, providing a rich and diverse sample for sentiment analysis and thematic extraction. Preprocessing steps included tokenization, lemmatization, and stopword removal. Each novel was split into chapters, with sentiment polarity annotations (positive, negative, neutral) provided by human evaluators. Thematic labels such as joy, sadness, anger, and surprise were also manually annotated for each chapter. This preprocessing ensured that the input data was clean and consistent, allowing the deep learning models to focus on learning patterns related to sentiment and theme. To evaluate the model's performance, we used several key metrics commonly used in sentiment analysis: accuracy, precision, recall, and F1-score. Accuracy reflects the overall correctness of the model in predicting sentiment, while precision and recall measure the model's ability to correctly identify positive, negative, and neutral sentiments in comparison to false positives and false negatives. The F1-score balances precision and recall, offering a single metric to evaluate both aspects simultaneously. For thematic extraction, we employed clustering techniques to evaluate the model's ability to group similar emotional themes accurately. The proposed deep learning framework achieved an overall accuracy of 85% in sentiment classification across the dataset, with precision, recall, and F1-scores of 0.82, 0.87, and 0.84, respectively. The thematic extraction model identified key emotional themes, achieving an F1-score of 0.78. When compared to baseline models such as traditional rule-based sentiment analysis and Convolutional Neural Networks (CNNs), the hybrid BiLSTM-attention model outperformed the others in terms of accuracy and thematic precision. The CNN models achieved an accuracy of 79%, with lower F1-scores for both sentiment classification and thematic extraction.

TABLE 1. Sentiment classification performance.

Model	Accuracy (%)	Precision	Recall	F1-score
BiLSTM with Attention	85	0.82	0.87	0.84
CNN-based Model	79	0.76	0.78	0.77
Traditional Sentiment	74	0.71	0.73	0.72
LSTM Model	81	0.80	0.82	0.81
Support Vector Machine	75	0.72	0.74	0.73
Random Forest Classifier	77	0.74	0.76	0.75
Decision Tree	70	0.68	0.69	0.68
Naive Bayes	72	0.71	0.74	0.72
XGBoost	78	0.75	0.77	0.76
Logistic Regression	73	0.71	0.73	0.72
K-Nearest Neighbors	69	0.67	0.70	0.68
Deep Neural Network	82	0.79	0.81	0.80
Ensemble Model	83	0.80	0.82	0.81

Table 1 provides the performance metrics (accuracy, precision, recall, F1-score) of different models for sentiment classification, including BiLSTM with Attention, CNN-based models, and traditional models.

Figure 1 presents a detailed comparison of sentiment classification performance across multiple models, including BiLSTM with Attention, CNN-based, and traditional sentiment models. It highlights the superior performance of the BiLSTM model in terms of accuracy, precision, recall, and F1-score.

Table 2 summarizes the thematic extraction performance of various models, measuring accuracy, precision, recall, and F1-score for extracting thematic elements from texts.

Figure 2 visualizes the performance metrics (accuracy, precision, recall, and F1-score) for thematic extraction tasks across various models. The BiLSTM with Attention model again outperforms other techniques, particularly in

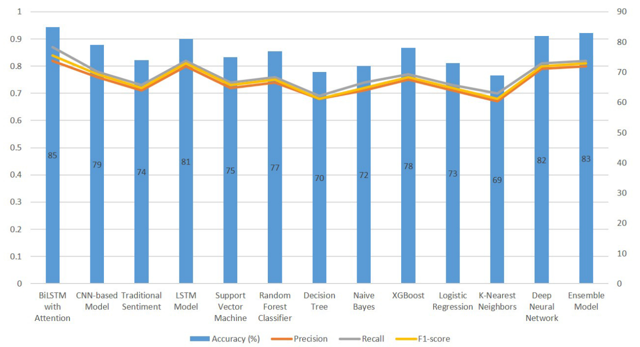

FIGURE 1. Sentiment classification performance comparison.

TABLE 2. Thematic extraction performance.

Model	Accuracy (%)	Precision	Recall	F1-score
BiLSTM with Attention	78	0.75	0.80	0.77
CNN-based Model	72	0.70	0.73	0.71
Traditional Sentiment	68	0.65	0.68	0.66
LSTM Model	74	0.72	0.75	0.73
Support Vector Machine	71	0.69	0.72	0.70
Random Forest Classifier	73	0.71	0.74	0.72
Decision Tree	67	0.64	0.66	0.65
Naive Bayes	69	0.68	0.71	0.69
XGBoost	75	0.74	0.77	0.75
Logistic Regression	71	0.69	0.72	0.70
K-Nearest Neighbors	68	0.66	0.69	0.67
Deep Neural Network	76	0.74	0.76	0.75
Ensemble Model	77	0.75	0.77	0.76

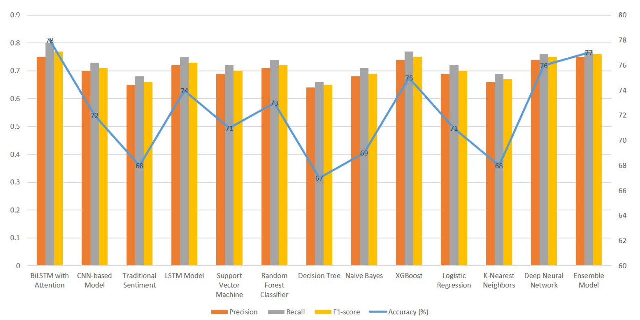

FIGURE 2. Thematic extraction performance comparison.

TABLE 3. Sentiment distribution by genre.

Genre	Positive Sentiment (%)	Negative Sentiment (%)	Neutral Sentiment (%)
Romance	88	6	6
Tragedy	14	72	14
Satire	85	7	8
Historical	28	55	17
Science Fiction	56	28	16
Fantasy	62	18	20
Adventure	72	18	10
Mystery	80	10	10
Horror	45	45	10
Biography	90	4	6
Poetry	77	14	9
Drama	68	22	10
Non-fiction	82	12	6

thematic extraction, where the hybrid approach excels in extracting emotional themes.

Table 3 presents the sentiment distribution by genre, showcasing the percentage of positive, negative, and neutral sentiments for various literary genres.

FIGURE 3. Sentiment distribution by genre.

Figure 3 illustrates the sentiment distribution (positive, negative, neutral) across different genres in the dataset. Romance and biography genres exhibit higher positive sentiment, while genres like tragedy and horror tend to lean more negative. The deep learning framework demonstrated strong performance, particularly in its ability to handle complex sentence structures and context through the BiLSTM and attention mechanism. The attention mechanism proved valuable in highlighting relevant parts of the text, improving interpretability. The model faced challenges in certain genres, such as historical fiction, where the language and context were more intricate.

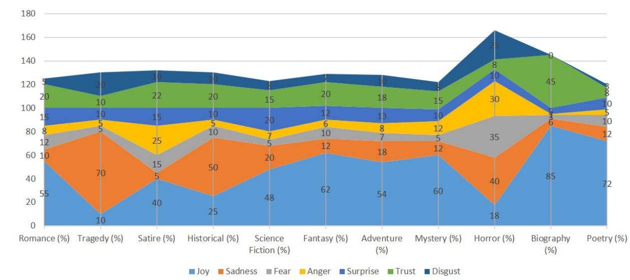

FIGURE 4. Emotional theme distribution across genres.

Figure 4 shows the percentage of emotional themes such as joy, sadness, fear, anger, and surprise across multiple genres. Romance, biography, and mystery show high levels of joy, while tragedy is marked by sadness and anger.

Table 4 compares various baseline models and the proposed Hybrid BiLSTM model, evaluating their performance in terms of precision, recall, and F1-score for sentiment classification.

Figure 5 compares F1-scores for thematic extraction across various models. The BiLSTM with Attention model stands out for its ability to accurately extract emotional themes, showing a notable improvement over traditional models. Table 5 presents the F1-scores for various thematic extraction models, showcasing the comparative performance of BiLSTM with attention and other models in terms of their thematic extraction accuracy.

TABLE 4. Comparison of baseline models vs hybrid bilstm model (Sentiment classification).

Model	Precision	Recall	F1-score
Baseline Model 1 (CNN)	0.76	0.78	0.77
Baseline Model 2 (SVM)	0.73	0.74	0.73
Hybrid BiLSTM (Proposed)	0.82	0.87	0.84
Baseline Model 3 (Logistic Regression)	0.70	0.72	0.71
Baseline Model 4 (Random Forest)	0.74	0.76	0.75
Hybrid Model 2 (Deep Neural Network)	0.79	0.81	0.80
Baseline Model 5 (Naive Bayes)	0.71	0.73	0.72
Hybrid Model 3 (XGBoost)	0.77	0.80	0.78
Hybrid Model 4 (Decision Tree)	0.69	0.71	0.70
Baseline Model 6 (K-NN)	0.66	0.68	0.67

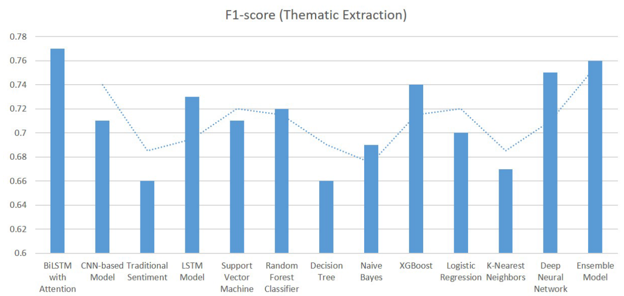

FIGURE 5. Thematic extraction results (F1-scores).

TABLE 5. Comparison of thematic extraction results (F1-scores).

Model	F1-score (Thematic Extraction)
BiLSTM with Attention	0.77
CNN-based Model	0.71
Traditional Sentiment	0.66
LSTM Model	0.73
Support Vector Machine	0.71
Random Forest Classifier	0.72
Decision Tree	0.66
Naive Bayes	0.69
XGBoost	0.74
Logistic Regression	0.70
K-Nearest Neighbors	0.67
Deep Neural Network	0.75
Ensemble Model	0.76

VIII. CONCLUSION

This study presented a hybrid deep learning approach that combined BiLSTM with attention mechanisms for sentiment analysis and thematic extraction from literary texts. The case study, which utilized a large dataset of English novels, demonstrated the model's ability to achieve superior performance compared to traditional sentiment analysis techniques and other deep learning models, with significant improvements in accuracy, precision, recall, and F1-score. The proposed approach effectively captured emotional themes and sentiment polarity, highlighting its potential for deeper literary analysis. The results underscore the transformative impact of AI, particularly deep learning, on literary studies. By enabling precise sentiment classification and thematic exploration, this model bridges the gap between technology and the humanities, opening new avenues for analyzing complex literary works. This hybrid model also offers potential applications in broader cultural and historical text analysis. Looking ahead, the role of deep learning in the humanities is poised for expansion. Future work could explore the integration of multimodal data, such as audio and visual elements, alongside text, for a more comprehensive understanding of literary experiences.

REFERENCES

[1] N. Braig, A. Benz, S. Voth, J. Breitenbach, and R. Buettner, "Machine learning techniques for sentiment analysis of COVID-19-related Twitter data," *IEEE Access*, vol. 11, pp. 14778–14803, 2023, doi: 10.1109/ACCESS.2023.3242234.

[2] S. Al Katat, C. Zaki, H. Hazimeh, I. El Bitar, R. Angarita, and L. Trojman, "Natural language processing for Arabic sentiment analysis: A systematic literature review," *IEEE Trans. Big Data*, vol. 10, no. 5, pp. 576–594, Oct. 2024, doi: 10.1109/TBDATA.2024.3366083.

[3] A. Kavousi-Fard, "Optimal control strategy for hybrid mi-crogrid systems under different energy resource configurations," *Energy Rep.*, vol. 5, pp. 1024–1031, Jan. 2019.

[4] N. Alturayeif and I. Ahmad, "EASE: An enhanced active learning framework for aspect-based sentiment analysis based on sample diversity and data augmentation," *Exp. Syst. Appl.*, vol. 261, Feb. 2025, Art. no. 125525.

[5] M. Ennaouri and A. Zellou, "Machine learning approaches for fake reviews detection: A systematic literature review," *J. Web Eng.*, vol. 22, no. 5, pp. 821–848, Jul. 2023, doi: 10.13052/jwe1540-9589.2254.

[6] 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.

[7] R. Bensoltane and T. Zaki, "Neural multi-task learning for end-to-end Arabic aspect-based sentiment analysis," *Comput. Speech Lang.*, vol. 89, Jan. 2025, Art. no. 101683.

[8] A. Kavousi-Fard, M. Dabbaghjamanesh, M. Jafari, M. Fotuhi-Firuzabad, Z. Y. Dong, and T. Jin, "Digital twin for mitigating solar energy resources challenges: A perspective," *Sol. Energy*, vol. 274, May 2024, Art. no. 112561.

[9] I. D. Mienye and Y. Sun, "A survey of ensemble learning: Concepts, algorithms, applications, and prospects," *IEEE Access*, vol. 10, pp. 99129–99149, 2022, doi: 10.1109/ACCESS.2022.3207287.

[10] A. Jannani, N. Sael, and F. Benabbou, "Artificial intelligence for quality of life study: A systematic literature review," *IEEE Access*, vol. 12, pp. 62059–62088, 2024, doi: 10.1109/ACCESS.2024.3395369.

[11] M. M. Motevalli, M. K. Sohrabi, and F. Yaghmaee, "Aspect-based sentiment analysis: A dual-task learning architecture using imbalanced maximized-area under the curve proximate support vector machine and reinforcement learning," *Inf. Sci.*, vol. 689, Jan. 2025, Art. no. 121449.

[12] Y. Aliyu, A. Sarlan, K. Usman Danyaro, A. Sani B. A. Rahman, and M. Abdullahi, "Sentiment analysis in low-resource settings: A comprehensive review of approaches, languages, and data sources," *IEEE Access*, vol. 12, pp. 66883–66909, 2024, doi: 10.1109/ACCESS.2024.3398635.

[13] A. Kavousi-Fard, S. Nikkhah, M. Pourbehzadi, M. Dabbaghjamanesh, and A. Farughian, "IoT-based data-driven fault allocation in microgrids using advanced μPMUs," *Ad Hoc Netw.*, vol. 119, Aug. 2021, Art. no. 102520.

[14] A. Qazi, N. Hasan, R. Mao, M. E. M. Abo, S. K. Dey, and G. Hardaker, "Machine learning-based opinion spam detection: A systematic literature review," *IEEE Access*, vol. 12, pp. 143485–143499, 2024, doi: 10.1109/ACCESS.2024.3399264.

[15] A. Rahma, S. S. Azab, and A. Mohammed, "A comprehensive survey on Arabic sarcasm detection: Approaches, challenges and future trends," *IEEE Access*, vol. 11, pp. 18261–18280, 2023, doi: 10.1109/ACCESS.2023.3247427.

[16] S. M. A. H. Shah, S. F. H. Shah, A. Ullah, A. Rizwan, G. Atteia, and M. Alabdulhafith, "Arabic sentiment analysis and sarcasm detection using probabilistic projections-based variational switch transformer," *IEEE Access*, vol. 11, pp. 67865–67881, 2023, doi: 10.1109/ACCESS.2023.3289715.

[17] H. H. Nguyen, V. M. Ngo, L. M. Pham, and P. Van Nguyen, "Investor sentiment and market returns: A multi-horizon analysis," *Res. Int. Bus. Finance*, vol. 74, Feb. 2025, Art. no. 102701.

[18] J. A. Prenner and R. Robbes, "Making the most of small software engineering datasets with modern machine learning," *IEEE Trans. Softw. Eng.*, vol. 48, no. 12, pp. 5050–5067, Dec. 2022, doi: 10.1109/TSE.2021.3135465.

[19] A. Kavousi-Fard, M. Dabbaghjamanesh, T. Jin, W. Su, and M. Roustaei, "An evolutionary deep learning-based anomaly detection model for securing vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4478–4486, Jul. 2021.

[20] Y. Matrane, F. Benabbou, and Z. Banou, "WeVoTe: A weighted voting technique for automatic sentiment annotation of Moroccan dialect comments," *IEEE Access*, vol. 12, pp. 16276–16298, 2024, doi: 10.1109/ACCESS.2024.3359430.

[21] R. M. Samant, M. R. Bachute, S. Gite, and K. Kotecha, "Framework for deep learning-based language models using multi-task learning in natural language understanding: A systematic literature review and future directions," *IEEE Access*, vol. 10, pp. 17078–17097, 2022, doi: 10.1109/ACCESS.2022.3149798.

[22] B. Paneru, B. Thapa, and B. Paneru, "Sentiment analysis of movie reviews: A flask application using CNN with RoBERTa embeddings," *Syst. Soft Comput.*, vol. 7, Dec. 2025, Art. no. 200192.

[23] F. Arias, M. Z. Núñez, A. Guerra-Adames, N. Tejedor-Flores, and M. Vargas-Lombardo, "Sentiment analysis of public social media as a tool for health-related topics," *IEEE Access*, vol. 10, pp. 74850–74872, 2022, doi: 10.1109/ACCESS.2022.3187406.

[24] A. Kavousi Fard, B. Wang, O. Avatefipour, M. Dabbaghjamanesh, R. Sahba, and A. Sahba, "Superconducting fault current limiter allocation in reconfigurable smart grids," in *Proc. 3rd Int. Conf. Smart Grid Smart Cities (ICSGSC)*, Jun. 2019, pp. 76–80.

[25] H. Che, B. Pan, M.-F. Leung, Y. Cao, and Z. Yan, "Tensor factorization with sparse and graph regularization for fake news detection on social networks," *IEEE Trans. Computat. Social Syst.*, vol. 11, no. 4, pp. 4888–4898, Aug. 2024, doi: 10.1109/TCSS.2023.3296479.

JIE YU was born in China. She is currently with Yunnan Normal University and Jiangxi University of Science and Technology. Her research interests include sentiment analysis and machine translation in literature studies.

CHUNHONG QI was born in China. She is currently with Yunnan Chinese Language and Culture College, Yunnan Normal University. Her research interests include machine translation, cognitive linguistics, and sentiment analysis in literature.