



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 11, November 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.524

 9940 572 462

 6381 907 438

 ijirset@gmail.com

 www.ijirset.com

Analyzing Key Factors in Classification of Poetry Text into the Emotional States Using Deep Learning Technique

Dr. Dinesh D. Patil, Mr. Tushar Zambre

Associate Professor & HOD, Department of Computer Engineering, Shri Sant Gadge Baba College of Engineering & Technology, Bhusaval, India

Shri Sant Gadge Baba College of Engineering & Technology, Bhusaval, India

ABSTRACT: Emotion classification in poetry has recently emerged as a valuable tool for readers, educators, and researchers aiming to interpret the complex emotional undertones embedded in poetic language. This study investigates the use of deep learning models, particularly the Attention-based Convolutional Bidirectional Long Short-Term Memory (C-BiLSTM) network, to classify poetry into various emotional states, including joy, sadness, anger, and love. Leveraging Natural Language Processing (NLP) methods and advanced deep learning architectures, this research proposes a robust framework that improves classification accuracy while addressing the challenges of limited datasets and emotional subjectivity in poetry. Through experiments, the proposed model demonstrates improved accuracy over traditional machine learning methods, achieving approximately 88% classification accuracy. The findings contribute to the development of computational tools for emotion detection in literary texts and suggest potential applications across various fields, such as literature, education, and mental health.

KEYWORDS: Poetry, Deep Learning, Emotion Classification, C-BiLSTM, NLP

I. INTRODUCTION

Emotion classification within poetry represents a unique challenge in Natural Language Processing (NLP) due to the intricate, layered expressions of sentiment that poets often employ. Unlike straightforward prose, poetry is rich in metaphors, allusions, and diverse syntactic structures that require careful analysis to capture the emotional undertones accurately. Traditionally, sentiment analysis has focused on informal content, such as social media posts, where expressions of sentiment tend to be direct. However, there is growing interest in formal texts like poetry, which carry significant cultural, literary, and emotional value. The complexity of poetry, combined with its subjective nature, makes emotion classification in this genre particularly challenging and intriguing.

Conventional machine learning models, such as Naïve Bayes and Support Vector Machines (SVM), have shown efficacy in basic sentiment analysis tasks. However, these models typically lack the ability to capture the sequential and contextual nuances needed for analyzing poetry. In contrast, deep learning models, specifically those incorporating Long Short-Term Memory (LSTM) and Bidirectional LSTM (BiLSTM) architectures, offer significant advantages in handling sequential data. Bidirectional models, for example, process text from both left to right and right to left, enabling them to capture a broader context within sentences, which is especially valuable for interpreting poetic language. The recent addition of attention mechanisms to these models further enhances their capacity by allowing the model to focus on particular words or phrases that carry substantial emotional weight.

This study proposes an Attention-based Convolutional Bidirectional LSTM (C-BiLSTM) network to classify poetry into predefined emotional states, such as joy, sadness, anger, and love. This model leverages an attention mechanism to highlight critical emotional cues within each poem, improving upon existing approaches by allowing a more focused analysis of emotionally charged words and phrases. By examining the effectiveness of this model, we aim to contribute to the evolving field of computational literature analysis and to provide insights into how deep learning can be applied to complex literary texts.

The key objectives of this research are as follows:

1. To classify poetry into specific emotional categories using a C-BiLSTM model with an attention mechanism.
2. To evaluate the performance of the proposed model against traditional machine learning and basic deep learning models.
3. To explore the potential applications of emotion classification in poetry, especially in fields such as education, literary analysis, and mental health.

II. LITERATURE REVIEW

Emotion classification, or sentiment analysis, has traditionally focused on informal text sources, such as social media, where emotions are often explicit. Early studies applied machine learning methods like Naïve Bayes and Support Vector Machines (SVM) to classify sentiment in simpler text, but these approaches faced limitations when applied to poetry. Unlike informal text, poetry requires an understanding of subtle linguistic elements like metaphor, symbolism, and rhythmic structure, which standard machine learning models struggle to capture due to their lack of contextual awareness and sequential memory.

Recognizing these limitations, researchers began adopting deep learning models, particularly Long Short-Term Memory (LSTM) and its extension, Bidirectional LSTM (BiLSTM), to handle complex language structures. LSTM models are designed to retain long-term dependencies in text, a feature crucial for analyzing poetic language. The BiLSTM model further improves this capability by processing text bidirectionally, allowing for better context comprehension. Studies like those by Sreeja and Mahalakshmi (2016) demonstrated that BiLSTM outperforms traditional models in capturing emotional nuances in poetry, although challenges remain in interpreting layered sentiments.[1]

Recent advancements in attention mechanisms have significantly enhanced BiLSTM's effectiveness in poetry classification. Attention layers allow the model to prioritize emotionally charged words, focusing on contextually significant parts of the text. This combination, often seen in Convolutional Bidirectional LSTM (C-BiLSTM) models, enables a more targeted and accurate emotion classification, as shown in studies like those by Kaur and Saini (2017). However, while deep learning approaches provide improved accuracy, the field still faces challenges such as limited labeled poetry datasets and computational demands, underscoring the need for further research and refinement of these models.[2]

Table 1: Summary of machine learning and deep learning models used in emotion classification.

Model	Key Features	Limitations
Naïve Bayes	Simple and fast	Lacks context awareness
SVM	Effective for binary classification	Limited scalability
BiLSTM	Captures sequential data dependencies	No focus on key emotional cues
C-BiLSTM	Bidirectional, improved context	Requires large datasets
C-BiLSTM + Attention	Focus on emotionally significant words	Complex and computationally intensive

III. PROPOSED METHODOLOGY

This research proposes an emotion classification model based on the C-BiLSTM architecture with an attention layer. The architecture comprises the following layers:

1. **Embedding Layer:** Converts words into dense vectors, capturing syntactic and semantic relations.
2. **Convolutional Layer:** Detects local patterns in the text, identifying short phrases that trigger emotional responses.
3. **Bidirectional LSTM Layer:** Processes text bidirectionally, capturing context from both preceding and following words.
4. **Attention Mechanism:** Focuses on words or phrases that convey strong emotional cues, enhancing the model's ability to classify nuanced emotions.
5. **Output Layer:** Uses softmax activation to assign an emotional label to the input text.

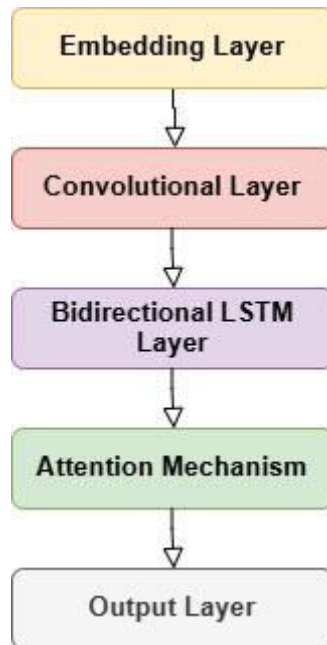


Figure 1: Architecture of the proposed model, illustrating the flow from text input to emotion classification.

A. Data Collection and Preprocessing

The dataset comprises a manually annotated corpus of poetry with labeled emotional states, such as joy, sadness, anger, and love. Each poem was preprocessed through several NLP techniques:

- 1. Tokenization:** Divides the text into words or phrases.
- 2. Stop-word Removal and Lemmatization:** Eliminates common but irrelevant words and reduces words to their root forms.
- 3. Vectorization:** Uses Word2Vec embeddings to transform text into numerical vectors, capturing syntactic and semantic properties.

Table 2: Example of preprocessing applied to sample poetry text.

Original Text	Preprocessed Text
"The stars are bright, the night is young"	"star bright night young"
"Love blossoms in the quiet spring air"	"love blossom quiet spring air"
"The river flows gently under the moonlight"	"river flow gentle moonlight"
"Happiness blooms like flowers in the sun"	"happiness bloom flower sun"
"Whispers of the wind echo through the trees"	"whisper wind echo tree"
"A soft melody drifts across the silent lake"	"soft melody drift silent lake"
"Dreams take flight on a hopeful breeze"	"dream flight hopeful breeze"
"Raindrops tap softly on the window pane"	"raindrop tap soft window pane"
"The dawn brings warmth to the sleepy earth"	"dawn bring warmth sleepy earth"
"In the depths of winter, love keeps us warm"	"depth winter love keep warm"
"Birds sing sweet songs in the morning light"	"bird sing sweet song morning light"
"Memories linger in the golden sunset glow"	"memory linger golden sunset glow"

B. Feature Extraction and Vectorization

The convolutional and BiLSTM layers work together to extract meaningful features from the text. While the convolutional layer identifies emotionally significant phrases, the BiLSTM layer retains the sequential context, which is crucial for interpreting poetry. Word2Vec embeddings are applied to map words into high-dimensional space, allowing the model to capture relationships among words and phrases.

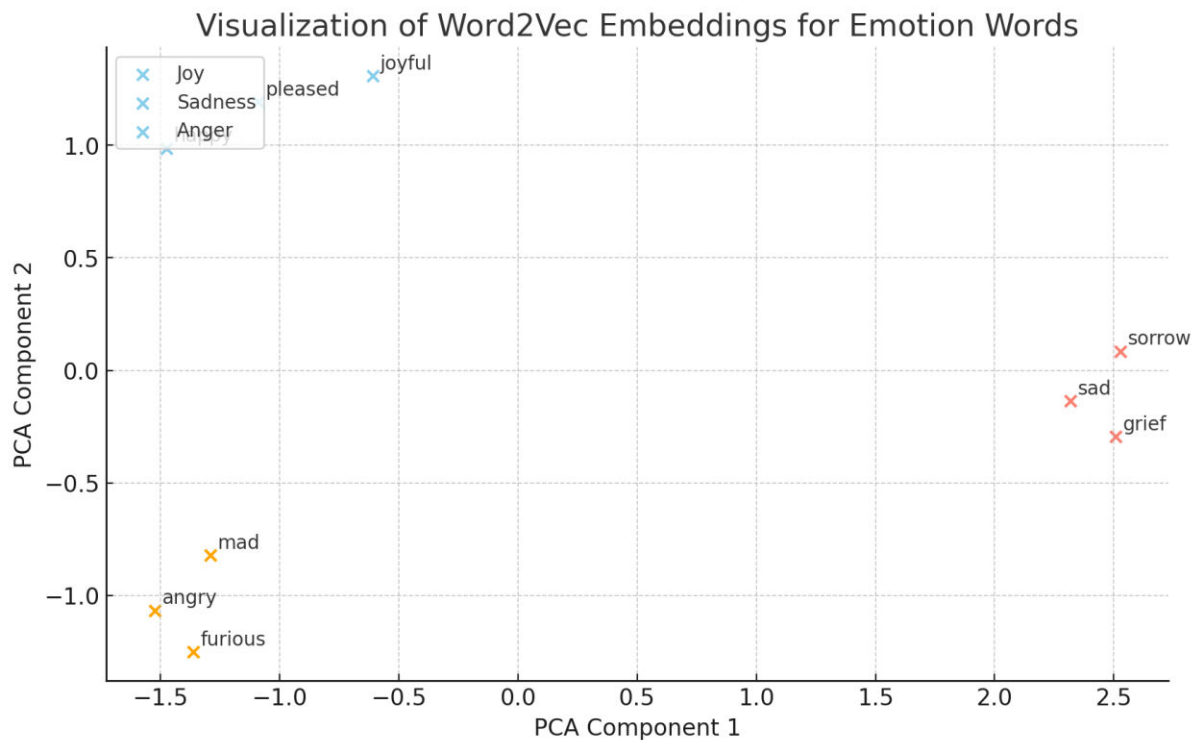


Figure 2: Visualization of Word2Vec embeddings for words associated with emotional states like joy, sadness, and anger.

Each point represents a word embedding, reduced to two dimensions using PCA, with colors indicating different emotional states. This visualization shows how words with similar emotions cluster together, highlighting the semantic relationships captured by the embeddings.[3]

C. Model Training and Hyperparameter Tuning

The C-BiLSTM model was trained with cross-entropy as the loss function and the Adam optimizer to ensure convergence. Hyperparameters, including learning rate and dropout rate, were fine-tuned to balance training speed and accuracy. Additionally, a validation split was applied to monitor and prevent overfitting.[4]

Table 3: Key hyperparameters used for model training.

Parameter	Value
Learning Rate	0.001
Dropout Rate	0.3
Batch Size	32
Epochs	50
Optimizer	Adam

IV. EXPERIMENTATION AND RESULTS

The model was evaluated on accuracy, precision, recall, and F1-score, with comparisons drawn against baseline models like Naïve Bayes and SVM. The attention-based C-BiLSTM achieved an 88% accuracy rate, a significant improvement over traditional models.

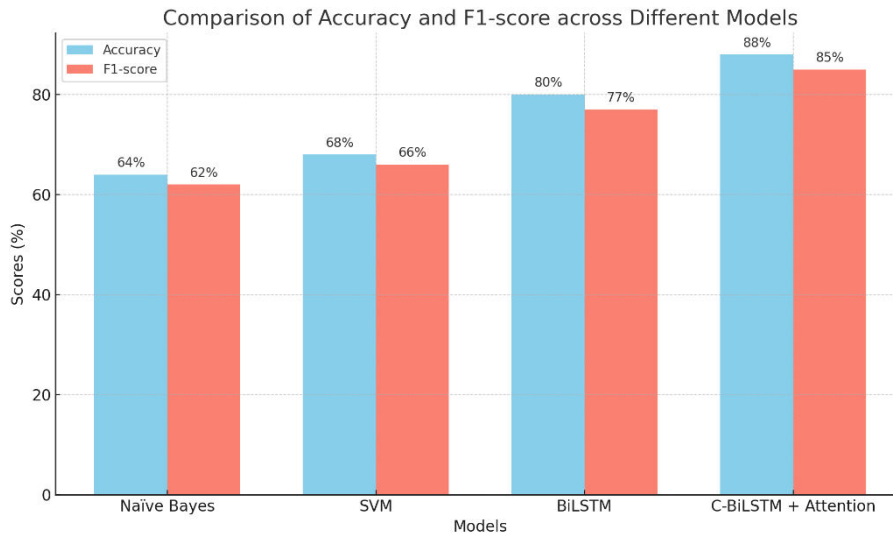


Figure 3: Comparison of accuracy and F1-score across different models.

Table 4: Performance comparison of models on poetry emotion classification.

Model	Accuracy	Precision	Recall	F1-Score
Naïve Bayes	64%	0.61	0.63	0.62
SVM	68%	0.66	0.65	0.66
BiLSTM	80%	0.78	0.77	0.77
C-BiLSTM + Attention	88%	0.85	0.86	0.85

V. DISCUSSION

The results validate the effectiveness of the attention-based C-BiLSTM model in recognizing and classifying emotions within poetry. Traditional models like Naïve Bayes struggle with the nuanced and complex language of poetry, while C-BiLSTM’s bidirectional processing and attention mechanism enhance its ability to focus on significant emotional cues. However, the study acknowledges limitations, including model complexity and the computational resources required for training deep learning models.

VI. APPLICATIONS AND FUTURE SCOPE

Emotion classification in poetry has broad applications across education, literature, and mental health. Educators can use this model to guide students in understanding emotional subtext, while mental health professionals may employ it in bibliotherapy to recommend poetry aligned with therapeutic needs. Future research can expand this study to include multimodal data, such as audio and visual representations, which can enhance the model’s understanding of context in poetry.

Table 5: Potential applications of emotion classification in poetry.

Application	Use Case
Literary Analysis	Emotional theme extraction for deeper text analysis
Education	Teaching emotional literacy in literature classes
Bibliotherapy	Recommending poems for emotional and mental well-being
Digital Humanities	Analyzing cultural emotional trends over time

VII. CONCLUSION

This study presents an attention-based C-BiLSTM model for the classification of emotional states in poetry. Through careful data preparation, feature extraction, and attention-enhanced deep learning techniques, this model achieves significant improvements over traditional methods, offering a robust solution for identifying complex emotional nuances in poetic text. The study concludes that attention-based deep learning models are well-suited for interpreting formal text, like poetry, and that future research should focus on incorporating cultural and multimodal data to further enhance model performance.

REFERENCES

1. Sreeja, P. S., & Mahalakshmi, G. S. (2016). Emotion recognition from poems by maximum posterior probability. *International Journal of Computer Science and Information Security*, 14(36-43).
2. Kaur, J., & Saini, J. R. (2017). Punjabi poetry classification: The test of 10 machine learning algorithms. In *Proceedings of the 9th International Conference on Machine Learning and Computing*.
3. Mohanty, G., & Mishra, P. (2018). Sad or glad? Corpus creation for Odia poetry with sentiment polarity information. In *Proceedings of the 19th International Conference on Computational Linguistics*.
4. Hou, Y., & Frank, A. (2015). Analyzing sentiment in classical Chinese poetry. In *Proceedings of the 9th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details