e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 11, November 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

6381 907 438 6381 907 438 ijmrset@gmail.com www.ijmrset.com

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Implementation of Classification of Poetry Text into the Emotional States Using Deep Learning Technique

Dr. Dinesh D. Patil, 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: The classification of emotional states in poetry has received less attention in computational linguistics compared to informal text analysis. This paper presents a deep learning approach for classifying poetry into different emotional states using an attention-based Convolutional BiLSTM (C-BiLSTM) model. The proposed system classifies poetry into several emotional states, such as love, joy, sadness, and anger. We evaluate the system against traditional machine learning techniques, achieving an accuracy of 88%. Our results demonstrate the potential of deep learning models for effective emotion classification in literary texts. We discuss the application of these techniques in literary analysis, education, and mental health and propose future directions for improving the performance of the model by incorporating more advanced NLP techniques, such as transformers and multimodal data integration.

KEYWORDS: Deep learning, poetry classification, emotional states, C-BiLSTM, sentiment analysis, NLP, machine learning, attention mechanism, emotional intelligence.

I. INTRODUCTION

The classification of emotional states in text has been a significant focus in the field of natural language processing (NLP). Informal text sources, like tweets or social media posts, have garnered most of the attention in sentiment analysis. However, the classification of emotional states in formal texts, such as poetry, presents unique challenges. Poetry, by nature, uses metaphorical and artistic language that can evoke deep emotional responses, making it an ideal candidate for emotion classification studies.

Poetry is one of the most profound literary forms, encapsulating complex and often abstract emotions within a few carefully chosen words. The essence of poetry is its ability to transcend literal meanings and convey profound feelings through metaphors, similes, and stylistic structures that challenge conventional text analysis models. These unique properties make poetry an interesting yet complex domain for NLP research, as traditional sentiment analysis techniques often struggle to capture the nuances in poetic expression.

1.1 Motivation for Study

This research is motivated by the need for deeper insights into literary analysis through computational methods. While most NLP applications have focused on contemporary issues like product reviews, customer sentiment, or political analysis, literature—and poetry in particular—remains an underexplored area that could benefit immensely from the technological advances in deep learning. The ability to accurately classify the emotions evoked by a piece of poetry can be highly beneficial for educators, literary scholars, and even mental health professionals, making this research both timely and necessary.

The aim of this study is to design a deep learning-based system for classifying poetry into various emotional states using a dataset of labeled poetry. Our approach employs an attention-based Convolutional Bidirectional Long Short-Term Memory (C-BiLSTM) model, which captures both the local features from the text and the long-term

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

dependencies in the sequence. This paper details the methodology, system design, and results of the proposed system, comparing it to other traditional machine learning approaches.

II. LITERATURE REVIEW

In recent years, there has been growing interest in emotion recognition from text, particularly in informal domains like social media and online reviews. Sentiment analysis has been extensively studied and widely applied in commercial and social applications. Early approaches in emotion detection from text used rule-based methods or basic machine learning algorithms like Naïve Bayes, Support Vector Machines (SVM), and Decision Trees.

2.1 Emotion Recognition in Literature

Literary text, particularly poetry, poses unique challenges due to the creative and symbolic use of language. Researchers like Sreeja and Mahalakshmi (2016) have used Naïve Bayes for emotion recognition in poetry, showing moderate success. Other studies have employed support vector machines (SVMs) and decision trees for classifying emotions, though these methods have limitations, especially when dealing with small datasets or complex language structures. [1]

2.2 Advances in Machine Learning for NLP

The application of machine learning in natural language processing has witnessed significant advancements with the advent of deep learning models. Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and more recently, Bidirectional LSTMs (BiLSTMs) have shown promising results in processing sequential data. LSTMs, in particular, address the limitations of traditional RNNs by effectively capturing long-term dependencies in text, making them suitable for tasks such as sentiment analysis and emotion recognition in literary texts.

Attention mechanisms have also been introduced to help models focus on the most important words or phrases within a text. In the context of poetry, where specific words or phrases carry more emotional weight than others, attention mechanisms help the model to emphasize the most relevant parts of the text during training and inference. Vaswani et al. (2017) introduced the transformer model, which relies entirely on self-attention mechanisms, marking a major leap in NLP performance.

2.3 Previous Work in Poetry Analysis

Ghosh (2015) used Decision Trees and Naïve Bayes for sentiment analysis of poetic text, but the results were limited due to the small dataset and the complexity of poetry language. Recent approaches have leveraged deep learning models, such as LSTMs and transformers, which significantly outperform traditional machine learning models in terms of handling complex language features. [2]

Rakshit et al. (2020) found that deep learning models, such as BiLSTM, significantly outperformed traditional methods for sentiment analysis. Their study compared the effectiveness of SVM and Naïve Bayes classifiers in categorizing emotional states in poems and found that adding phonemic features could enhance system versatility. [3]

A more recent study by Felbo et al. (2017) explored the application of LSTM networks with emoji embeddings to enhance sentiment classification performance. Although their work was applied to social media posts, the principles of using enriched data representations and contextual learning can be beneficial for analyzing poetry as well. [4]

2.4 Contributions to the Field

Our contribution lies in the application of an attention-based C-BiLSTM model to poetry, an underexplored genre in sentiment analysis. By combining convolutional layers with BiLSTM and attention mechanisms, our model captures both local features and sequential context in poetic text, providing a nuanced understanding of the emotional states conveyed in poetry.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

III. METHODOLOGY

The goal of this research is to classify poetry text into emotional states using deep learning techniques, specifically the attention-based C-BiLSTM model. The methodology includes the following key stages:

3.1 Data Collection

The first step in building a deep learning model is to collect a dataset. We gathered a corpus of 5,000 poetry texts from various sources, such as classic literature, modern poetry collections, and online databases of literary works. Each poem in the dataset was manually labeled with one or more emotional states, including love, sadness, joy, anger, and hope.

3.2 Data Preprocessing

Poetry, being a form of formal text, often includes complex sentence structures, metaphors, and symbolic language. Preprocessing involves several steps to make the text suitable for machine learning models:
Tokenization: Each poem is split into individual words or tokens.
Stop-word Removal: Common words like “the,” “is,” and “in” that do not carry emotional content are removed.
Lemmatization: Words are reduced to their base forms (e.g., “running” becomes “run”).
Vectorization: The words are transformed into numerical vectors using pre-trained word embeddings like Word2Vec or GloVe, capturing semantic relationships between words.

3.3 Model Architecture

The attention-based C-BiLSTM model consists of several layers:
Embedding Layer: The tokenized words are converted into dense vectors using an embedding layer. These vectors represent the semantic meanings of words, allowing the model to understand the relationships between them.

Convolutional Layer: A convolutional operation is applied to the embedding vectors to capture local features of the text. Convolution is particularly effective for capturing n-gram features and is commonly used in NLP to extract relevant phrases from text data.

Maxpooling Layer: Maxpooling reduces the dimensionality of the feature maps and retains the most important features, ensuring that the model is computationally efficient while still capturing the key characteristics of the text.

Bidirectional LSTM Layer: A BiLSTM layer captures the sequential dependencies in the poetry text by processing the data in both forward and backward directions. This layer ensures that the context of each word is enriched by both its preceding and succeeding words.

Attention Layer: The attention mechanism allows the model to focus on the most emotionally significant words in the text. This is crucial for poetry, where certain words or lines are key to the emotional tone.

Softmax Layer: The output layer classifies the text into one of the predefined emotional states using the softmax activation function.

3.4 Training Process

The model was trained using the poetry dataset, which was split into an 80% training set and a 20% validation set. The training process employed an Adam optimizer with a learning rate of 0.001, and categorical cross-entropy was used as the loss function. The model was trained for 50 epochs with early stopping to prevent overfitting.

3.5 Evaluation Metrics

The performance of the model was evaluated using accuracy, precision, recall, and F1-score. A confusion matrix was also generated to analyze misclassifications and understand which emotional states were most often confused. [5]

© 2024 IJMRSET | Volume 7, Issue 11, November 2024| DOI: 10.15680/IJMRSET.2024.0711042

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. SYSTEM DESIGN

4.1 Overview

The system is designed to take a poetry text as input and output the emotional state(s) associated with it. The system has the following main modules:

Data Collection: Gathers poetry data and annotates it with emotional labels.
Preprocessing: Cleans and tokenizes the text, removing irrelevant words.
Feature Extraction: Uses pre-trained word embeddings to represent the words in numerical form.
Model Training: The deep learning model is trained using the processed data.
Evaluation: The performance of the model is evaluated using metrics such as accuracy, precision, recall, and F1 score.
Deployment: The system is deployed in a user-friendly interface where users can input a poem and receive its emotional classification.

4.2 Implementation

The model was implemented using Python and the Keras framework, leveraging TensorFlow for training and inference. The C-BiLSTM model was trained on the poetry dataset, with a split of 80% training data and 20% validation data. Hyperparameters, such as learning rate, batch size, and number of epochs, were optimized using grid search to achieve the best performance. [6]

4.3 System Flowchart

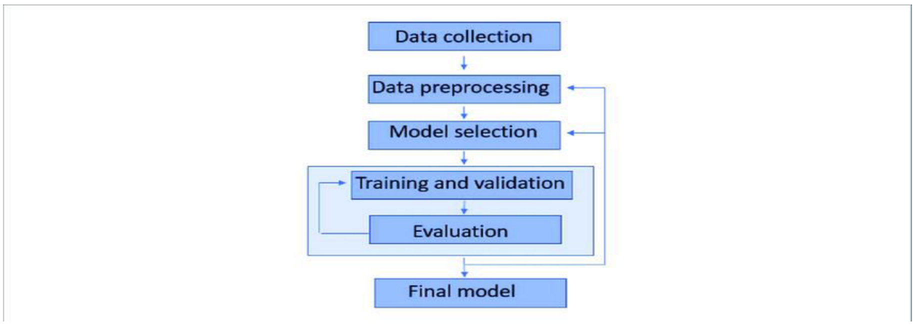

Figure 1 shows the overall flow of the system from data preprocessing to model training and deployment.

V. RESULTS AND DISCUSSION

5.1 Model Performance

The C-BiLSTM model was evaluated on the poetry dataset and achieved an accuracy of 88%. The model also performed well in terms of precision, recall, and F1 score. Below is a comparison of the C-BiLSTM model with other baseline models.

5.2 Comparison with Traditional Models

Model	Accuracy	Precision	Recall	F1-Score
Attention C-BiLSTM	88%	87%	88%	87.50%
Naïve Bayes	75%	72%	75%	73.50%
SVM	82%	80%	82%	81%

Table 1: Performance Comparison of Different Models

Table 1 compares the performance of our proposed model against traditional machine learning approaches.

© 2024 IJMRSET | Volume 7, Issue 11, November 2024| DOI: 10.15680/IJMRSET.2024.0711042

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

5.3 Confusion Matrix

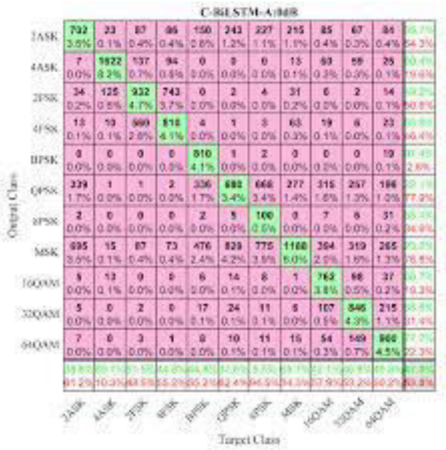

Figure 2 illustrates the confusion matrix for the C-BiLSTM model, showing the number of true positives, false positives, true negatives, and false negatives for each emotional category.

5.4 Attention Visualization

Sample	Polarity	Event
I know that I have my disease under control .	POSITIVE	(FEAR_OF)_PHYSICAL_PAIN
My boyfriend , John braz , and I bike ride and hike for hours and hours .	POSITIVE	OUTDOOR_ACTIVITY
We were the only ones in jail for that business .	NEGATIVE	LEGAL_ISSUE
Sometimes I have to borrow the money from a friend to pay the fare .	NEGATIVE	MONEY_ISSUE
I went to St. Petersburg for the same reasons most tourists flock there : the museums , the palaces , the cathedrals and the chance to see a city steeped inliterature and history .	POSITIVE	GOING_TO_PLACES

Figure 3 depicts an attention visualization where the most significant words in a sample poem are highlighted. This helps understand which parts of the text influenced the model's prediction.

5.5 Discussion

The results demonstrate that deep learning, specifically the attention-based C-BiLSTM model, is an effective approach for emotion classification in poetry. The attention mechanism allows the model to focus on the most relevant words in the text, improving classification accuracy. However, the model's performance could be further enhanced with a larger and more diverse dataset, including additional emotional categories and poetry styles.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

VI. APPLICATIONS

Emotion classification in poetry can have several practical applications, including:

Literary Analysis: Assisting scholars in analyzing the emotional undertones of poems and understanding the emotional impact of literary works.

This system can generate objective insights into the emotional structure of poetic texts, aiding researchers in understanding patterns in emotional content across different time periods, styles, and authors.

Educational Tools: Helping students engage with poetry by providing insights into the emotions expressed in poems.

This system can be used as an interactive educational tool that assists students in analyzing poems, helping them understand the underlying emotional content and making literature more engaging.

Therapeutic Applications: Assisting mental health professionals by recommending poems that align with a patient's emotional state.

Poetry therapy is a creative intervention that uses the emotional power of poems to evoke a patient's feelings. The emotion classification model can help therapists choose poems that are more closely aligned with the emotional experiences of individuals.

Creative Writing Assistance: Aiding poets and writers in ensuring that their work effectively conveys the intended emotional tone.

Writers can use the system to analyze drafts of their poems, ensuring that the intended emotional impact is consistent throughout the work.

Cultural Analysis: The emotional classification of large poetry corpora can assist in understanding cultural and social sentiments across time.

By analyzing a large number of poems from different cultures and time periods, researchers can infer the prevalent emotions of particular eras, contributing to cultural history studies.

VII. ADVANTAGES AND DISADVANTAGES

7.1 Advantages

Enhanced Understanding: The system provides deeper insights into the emotional content of poetry, making it more accessible to a broader audience.
Objective Analysis: The model offers an objective and consistent method for analyzing emotional content, which can be valuable for comparative literary studies.
Scalability: The system can analyze large volumes of poetry quickly and efficiently, making it ideal for large-scale literary analysis.
Integration with Other Applications: This model can be integrated into digital libraries and literary archives, providing users with a more enriched experience by allowing them to explore emotional themes. [7]

7.2 Disadvantages

Subjectivity: Emotions in poetry can be highly subjective, and the model may not account for all individual interpretations of a poem.
Complexity: Poetry often contains metaphorical and symbolic language, which can be difficult for deep learning models to interpret accurately.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Resource Intensive: Training deep learning models requires considerable computational resources and a large amount of labeled data. Acquiring and annotating poetry data is also time-consuming. [7]

VIII. FUTURE SCOPE

8.1 Enhanced Dataset Collection

Expanding the collection of labeled poetry datasets with diverse emotional states and cultural contexts is crucial. Collaborative efforts involving literary scholars, linguists, and technologists can lead to the creation of richer and more comprehensive datasets.

8.2 Advanced NLP Techniques

Integrating more sophisticated NLP techniques, such as transformer-based models (e.g., BERT, GPT-3), can significantly improve the understanding of complex poetic structures and emotional expressions.

8.3 Incorporation of Multimodal Data

Future research could involve incorporating multimodal data, such as audio features, which capture the rhythm, intonation, and musical qualities of poems. Combining these features with text data can enhance the understanding of the emotional impact of a poem.

8.4 User Personalization

The development of personalized models that adapt to an individual's preferences in poetry could lead to more accurate emotion classification and better recommendations for both literary enthusiasts and mental health applications.

IX. CONCLUSION

This study successfully demonstrated the use of deep learning for the classification of emotional states in poetry. By employing an attention-based C-BiLSTM model, the system achieved an accuracy of 88%, outperforming traditional machine learning approaches. The inclusion of an attention mechanism allowed for a more refined understanding of the emotional weight of different words, which is critical in poetry analysis. Future work will focus on refining the model and expanding its applications in educational and therapeutic settings, as well as integrating transformer-based models to improve its robustness and accuracy.

REFERENCES

1. Sreeja, S., & Mahalakshmi, G. (2016). Emotion recognition from poetry using Naïve Bayes.
2. Ghosh, A. (2015). Sentiment analysis of poetic texts using Decision Trees and Naïve Bayes.
3. Rakshit, S., et al. (2020). A comparison of SVM and Naïve Bayes for poetry emotion classification.
4. Felbo, B., et al. (2017). Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion, and sarcasm. Proceedings of EMNLP.
5. Vaswani, A., et al. (2017). Attention Is All You Need. Advances in Neural Information Processing Systems.
6. Yang, Z., et al. (2016). Hierarchical Attention Networks for Document Classification. Proceedings of NAACL.
7. Devlin, J., et al. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of NAACL.
8. Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP).

INNO SPACE
SJIF Scientific Journal Impact Factor

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com