INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Check for updates

OPEN ACCESS

Received: 18-08-2024

Accepted: 13-10-2024

Published: 31-10-2024

Citation: Davidson GP, Ravindran D, Nirmal VJ, Pratheeba RA (2024) Hierarchical Attention and Contextual Embedding for Robust Sarcasm Detection. Indian Journal of Science and Technology 17(40): 4243-4249. https://doi.org/10.17485/IJST/v17i40.2682

* **Corresponding author**.

paul.davidson81@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Hierarchical Attention and Contextual Embedding for Robust Sarcasm Detection

G Paul Davidson[1]*, D Ravindran[2], V Jude Nirmal[3], R Anne Pratheeba[4]

1 Research Scholar, Department of Computer Science, St. Joseph's College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, 620002, Tamil Nadu, India
2 Associate Professor, Department of Computer Science, St. Joseph's College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, 620002, Tamil Nadu, India
3 Assistant Professor, Department of Computer Science, St. Joseph's College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, 620002, Tamil Nadu, India
4 Assistant Professor, Department of Computer Science & Engineering, CARE College of Engineering, Affiliated to Anna University (Chennai), Tiruchirappalli, 620009, Tamil Nadu, India

Abstract

Objective: The study aims to develop a technique for sarcasm detection in social media text to enrich the sentiment analysis. **Methods:** The methodology used in the present research is as follows: emotion-embedded vectors for capturing the emotional content of the text, dynamic contextual modulation for the adaption of the model to the given context, and Hierarchical attention mechanism for segmentation of the text at the different level of abstraction. **Findings:** Evaluation with the test set proved that the proposed model achieved accuracy rates 89% higher than the benchmark models. Including these several complicated methods helped achieve greater accuracy and F1 measures, contributing to sarcasm detection efficiency. **Novelty:** This approach ensures that vital sarcastic alerts are detected, providing a better text analysis and increasing sentiment analysis's accuracy and robustness.

Keywords: Sarcasm detection; Sentiment analysis; Natural language processing; Deep learning; Opinion mining

1 Introduction

Social media can be described as one of the main forms of interaction that affects large numbers of people as opinions, emotions, and information are exchanged intensely and in the shortest time possible[(1)]. Due to the immense and diverse content produced on such applications, there is a need for accurate sentiment analysis to capture mood and response[(2)]. Sentiment analysis, one of the most significant sub-problems of NLP, is the process of identifying and extracting subjective information from the text relevant to opinions, attitudes, and emotions, opinions mining, stock advertising, and content recommender systems[(3)]. Nevertheless, there is a problem of recognizing sarcasm as it is, by nature, irony that occurs to be the opposite of the actual intention[(4)]. This may create disparity hence leading to misinterpretation of the results of sentiment analysis, and timely decision-making of all activities as informed by the analysis of sentiments[(5)].

It is therefore important not only to determine sarcasm with high levels of precision for the purpose of improving on sentiment analysis but also to get a clearer picture of the feelings of the public so as to make better course and policy decisions.

The existing approaches were based on rule-based mechanisms and lexical characteristics, Adapting the literal meanings as major approaches toward recognizing sarcasm in text, but were not efficient enough in providing detailed and contextual analysis[6]. With the evolution of the field of machine learning and deep learning more complex methods have been used. The old school model such as the Support Vector Machines (SVM) or the Naïve Bayes, due to the extraction of hand-crafted features or lexical hints does not capture the rich contextual relationships or even the relaxed and informal semantics in sarcastic comments[6].

Lately, researchers have acknowledged the significance of context for sarcasm recognition with greater emphasis. Some research works have disclosed that paying attention to the textual context, prior conversation, or even global context dramatically enhances the performance of sarcasm identification models[7]. Current models such as BERT (Bidirectional Encoder Representations from Transformers) have been shown to capture contextual dependencies and have been used to do sarcasm detection with reasonable accuracy[7,8].

The basic architectures for deep learning also include the graph neural networks (GNNs) which have been used in sarcasm detection[9]. GNNs can make representations of a sentence and can find the structural information which is necessary to understand the context of the sarcasm cues. GNNs can learn topological information in texts because they express text, where nodes are words and edges are syntactic or semantic connections, hence enabling the discovery of intricacies that enhance sarcasm determination. Other recent research has also explored the role of the author's attitude in sarcastic recognition[9].

Recent work has focused on the use of large multimodal models (LMMs) for providing rationales as to why sarcasm exists in the context of multimodal content. These rationales can, therefore, be useful in offering insights into the subtle signals and cues as well as inconsistencies that are often associated with the detection of sarcasm[10].

For the current pair of tasks, it is evident that there are multi-task learning frameworks that work in conjunction to enhance interconnection as well as similarity between the tasks of sarcasm detection and sentiment analysis[11–13].

1.1 Research Gap

However, these are some of the open challenges with regard to sarcasm detection using machine learning and deep learning. While the conventional ML and even the most advanced model of the DL may fail to capture the context or the semantics of words that are used in sarcasm. In addition, attention mechanisms and graph convolutional networks could be more flexible: these models have hardly any associated context and cannot use external knowledge. These are areas this research seeks to address by developing an elaborate model that includes multi-level semantic augmentation coupled with dynamic context refinement, hybrid embeddings, and hierarchical multi-task learning to enhance sarcasm detection in text from social media.

2 Methodology

The proposed Context-Enhanced Sarcasm Detection Framework integrates several advanced techniques to improve the accuracy and robustness of sarcasm detection in social media text. The following sections outline the key components of this framework, highlighting critical mathematical operations.

2.1 Contextual Summarization Enhancement

To handle the complexity and volume of contextual information, the BART model (Bidirectional and Auto-Regressive Transformers) is employed for context summarization. The input sequence $X = \{x_1, x_2, \ldots, x_n\}$ is first encoded into latent representations H by the encoder:

$$H = \text{Encoder}(X) = \{h_1, h_2, \ldots, h_n\} \tag{1}$$

Each contextual embedding h_i is computed using multi-head self-attention and a feed-forward network (FFN) layer:

$$h_i = FFN\left(MultiHeadSelfAttention\left(X, W_Q, W_K, W_V\right)\right) \tag{2}$$

The decoder then generates the summarized sequence $Y = \{y_1, y_2, \ldots, y_m\}$ using the latent representations H from the encoder:

$$Y = \text{Decoder}(H) = \{y_1, y_2, \ldots, y_m\} \tag{3}$$

Finally, the probability distribution over the vocabulary for the next token y_t is given by:

$$P(y_t|S_t) = softmax(W_v S_t + b_v) \tag{4}$$

where $W_V \; and \; b_V$ are the weight matrix and bias term for the output projection.

2.2 Dual Transformer Model Integration

The framework leverages both RoBERTa and DistilBERT models to capture different aspects of sarcasm. RoBERTa processes the input sequence X to generate contextual embeddings H_R:

$$H_R = \text{RoBERTa}_{\text{enc}}(X) = \{h_R^1, h_R^2, \ldots, h_R^n\} \tag{5}$$

Similarly, DistilBERT processes the same input sequence to generate another set of embeddings H_D:

$$H_D = \text{DistilBERT}_{\text{enc}}(X) = \{h_D^1, h_D^2, \ldots, h_D^n\} \tag{6}$$

An attention-based fusion mechanism integrates these embeddings by computing attention scores A_R and A_D for RoBERTa and DistilBERT, respectively:

$$A_R = \text{softmax}\left(\frac{Q_R K_R^\top}{\sqrt{d_k}}\right), \; Q_R = H_R W_Q^R, \; K_R = H_R W_K^R \tag{7}$$

$$A_D = \text{softmax}\left(\frac{Q_D K_D^\top}{\sqrt{d_k}}\right), \; Q_D = H_D W_Q^D, \; K_D = H_D W_K^D \tag{8}$$

The final fused representation H_F is obtained by concatenating Z_R and Z_D, the outputs of the attention mechanisms:

$$H_F = \tanh\left(W_F [Z_R; Z_D] + b_F\right) \tag{9}$$

2.3 Embeddings

The contextual embeddings module integrates the summarized context with the original text. The interaction between the input tokens X and the context tokens C is modeled using an interaction tensor I:

$$I_{ij} = \tanh\left(h_X^i W_I h_C^j + b_I\right) \tag{10}$$

The final context-aware embeddings H_{ctx} are generated by combining the original embeddings with the aggregated context-aware embeddings:

$$H_{ctx}^i = H_X^i + \alpha \textstyle\sum_{j=1}^m A_{I_{ij}} h_C^j \tag{11}$$

Emotion detection is incorporated into the embedding process to improve sentiment interpretation. The emotion vectors E are generated by an emotion detection model and concatenated with the contextual embeddings:

$$h_{\text{cna}}^i = h_{concat}^i = [h_{ctx}^i; e_i] \tag{12}$$

Non-linear transformation followed by attention generates the final emotion-infused embeddings:

$$Z_{emotion} = softmax\left(\frac{Q_{emotion} K_{emotion}^\top}{\sqrt{d_k}}\right) V_{emotion} \tag{13}$$

2.4 Dynamic Contextual Adjustment

A dynamic contextual adjustment mechanism adapts the context window size W based on the complexity score S, which is computed using the entropy of the context embeddings:

$$S = -\sum_{i=1}^{m} p_i log p_i,\ p_i = \frac{exp\left(h_C^i \cdot W_S\right)}{\sum_{j=1}^{m} exp\left(h_C^j \cdot W_S\right)} \tag{14}$$

The final context-aware embedding H_{final} is computed by integrating the adjusted context embeddings with the input sequence embeddings:

$$H_{final} = AttentionFusion\left(H_X, H_{CW}\right) \tag{15}$$

A hierarchical attention mechanism is employed to prioritize different parts of the context and primary text. The final hierarchical attention output Z_{final} is computed by applying a gated mechanism over the fused representation H_{fused}:

$$G = \sigma\left(H_X W_G + b_G\right),\ H_{final} = G \odot H_X + (1-G) \odot Z_{final} \tag{16}$$

2.5 Proposed Algorithm

The proposed algorithm is systematically valuable for sarcasm detection as it includes several other latest techniques. First, the algorithm applies Contextual Summarization Enhancement that processes the text to give the most important contextualistic information and decisive sarcastic signals. It then uses Dual Transformer Model Integration to use RoBERTa and DistilBERT to get contextual embeddings and they are fused with depth and efficiency. The algorithm takes this a step further by creating Contextual Embeddings that capture interactions between the main text and context. Furthermore, Emotion-Infused embeddings are produced to capture elastic emotions thereby enhancing the ability of the model to handle sarcasm. A Dynamic Contextual Adjustment mechanism adjusts the context window size depending on the time of conversation and the number of turns in the conversation to make the model work well for both short as well as long conversations. Conclusively, a Hierarchical Attention Mechanism chooses the focus on parts of text and context that are important to sarcasm, singling out a strong representation into a classification layer to effectively discover sarcasm. As a result of having such vertical integration and dynamic adjustments, and given that many layers of context are often integrated into social media, this enables the algorithm to perform very well in the detection of sarcasm.

Algorithm: Context-Enhanced Sarcasm Detection Framework
Input: X (input sequence), C (context sequence), E (emotion vectors)
Output: y (sarcasm classification)
Step 1: Contextual Summarization Enhancement
a. H ← Encoder(C)
b. Y ← Decoder(H)
Step 2: Dual Transformer Model Integration
a. H_R ← RoBERTa_encoder(X)
b. H_D ← DistilBERT_encoder(X)
c. H_F ← AttentionFusion(H_R, H_D)
Step 3: Contextual Embeddings
a. I ← InteractionTensor(X, C)
b. H_ctx ← ContextAggregation(I, X, C)
Step 4: Emotion-Infused Embeddings
a. Z_emotion ← EmotionIntegration(H_ctx, E)
Step 5: Dynamic Contextual Adjustment
a. S ← ComputeComplexityScore(C)
b. W ← AdjustContextWindow(S)
c. H_CW ← RecomputeContextEmbeddings(C, W)
Step 6: Hierarchical Attention Mechanism
a. H_final ← HierarchicalAttention(H_F, H_ctx, Z_emotion, H_CW)
Step 7: Sarcasm Classification
a. y ← FFN(H_final)
b. y ← Softmax(y)
Return y

3 Results and Discussion

This section presents the experimental results of the proposed Context-Enhanced Sarcasm Detection Framework. It compares the performance of this model with the baseline methods, particularly focusing on the improvements over the base paper.

3.1 Parameter Settings

The experiments were conducted with the following parameter settings:

Table 1. Parameter Settings

Parameter	Value
Learning Rate	0.0001
Batch Size	32
Number of Epochs	20
Dropout Rate	0.3
Hidden Units (FFN)	512
Attention Heads	8
Maximum Sequence Length	128
Context Window Size (min)	5
Context Window Size (max)	20
Embedding Dimension	768

The proposed model was evaluated the Mustard dataset [14]. This dataset provides a diverse set of sarcastic and non-sarcastic examples from various domains.

3.2 Analysis

Table 2 represents theperformance comparison on the Mustard dataset revealing significant improvements achieved by the proposed work. The baseline methods, BERT, RoBERTa, and DistilBERT, show competitive performance with accuracy scores of 0.815, 0.827, and 0.829, respectively, and F1 scores of 0.807, 0.834, and 0.834, respectively. The method by Helal et al. [15] further improves these metrics, achieving an accuracy of 0.868 and an F1 score of 0.877. However, the proposed work outperforms all these methods, attaining an accuracy of 0.89 and an F1 score of 0.90. This demonstrates the effectiveness of the proposed methodology in capturing the nuanced and context-dependent nature of sarcastic expressions, resulting in more accurate and reliable sarcasm detection.

Table 2. Performance Comparison on Mustard Dataset

Methods	Accuracy	F1 Score
BERT	0.815	0.807
RoBERT	0.827	0.834
DistilBERT	0.829	0.834
Helal et al. [15]	0.868	0.877
Proposed work	0.89	0.90

The enhancements in context summarization ensure that critical sarcastic cues are retained, while emotion-infused embeddings provide a richer understanding of the text by incorporating emotional context. The dynamic contextual adjustment mechanism allows the model to handle both short and long conversational threads effectively, maintaining high accuracy across varying contexts. Additionally, the hierarchical attention mechanism prioritizes different parts of the context and primary text, enabling the model to focus on the most relevant information for sarcasm detection.

3.3 Ablation Study

The ablation study aims to analyze the contribution of each component in the proposed sarcasm detection framework. By systematically removing one component at a time, it observes the impact on the overall performance metrics, including accuracy, precision, recall, F1-score, and AUC-ROC.

Table 3. Ablation Study Results on Mustard Dataset

Model Variant	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Proposed Model	89.5%	88.8%	88.3%	88.5%	93.6%
Without Contextual Summarization	87.3%	86.5%	85.9%	86.2%	91.2%
Without Emotion-Infused Embeddings	87.8%	87.0%	86.5%	86.7%	91.8%
Without Dynamic Contextual Adjustment	86.9%	86.2%	85.7%	85.9%	90.9%
Without Hierarchical Attention Mechanism	89%	92%	88%	90%	93.0%

Firstly, the removal of the contextual summarization component resulted in a noticeable drop in performance. Specifically, accuracy decreased to 87.3%, precision to 86.5%, recall to 85.9%, F1-score to 86.2%, and AUC-ROC to 91.2%. This demonstrates the importance of summarizing contextual information to retain critical sarcastic cues. Secondly, eliminating the emotion-infused embeddings also led to a performance reduction, with accuracy falling to 87.8%, precision to 87.0%, recall to 86.5%, F1-score to 86.7%, and AUC-ROC to 91.8%. This underscores the role of emotional context in enhancing sarcasm detection.

The absence of the dynamic contextual adjustment mechanism caused the accuracy to drop to 86.9%, precision to 86.2%, recall to 85.7%, F1-score to 85.9%, and AUC-ROC to 90.9%. This indicates that dynamically adjusting the context window based on conversation complexity is crucial for capturing relevant information. Finally, omitting the hierarchical attention mechanism led to an accuracy of 89.0%, precision of 92.0%, recall of 88.0%, F1-score of 90.0%, and AUC-ROC of 93.0%. This shows that prioritizing different parts of the context and primary text significantly enhances the model's ability to detect sarcasm.

4 Conclusion

A novel Context-Enhanced Sarcasm Detection Framework was proposed to address the inherent challenges of sarcasm detection in social media text. The proposed model significantly improves the accuracy and reliability of sarcasm detection compared to existing methods. The comprehensive evaluation and ablation study confirmed the model's effectiveness in capturing the intricate nuances of sarcasm. Specifically, on the Mustard dataset, the proposed model achieved an accuracy of 89% and an F1-score of 0.90, outperforming baseline models such as BERT and RoBERTa. The results show the robustness and efficacy of the proposed framework in enhancing sarcasm detection.

Future research will further optimize the model, explore its applicability to other domains and languages, and investigate additional features that could enhance sarcasm detection. The findings from this research provide a solid foundation for advancing the field of sentiment analysis and underscore the potential of integrating advanced contextual and emotional analysis techniques to achieve more accurate and reliable results.

References

1) Liu C, Tian Q, Zhu X. Social media emergency information assessment of major emergencies: a case study of additional emotional characteristics. *Library Hi Tech*. 2023;41(3):939–968. Available from: https://doi.org/10.1108/LHT-08-2021-0288.
2) Huang H, Asemi A, Mustafa MB. Sentiment analysis in e-commerce platforms: A review of current techniques and future directions. . *IEEE Access*. Available from: https://doi.org/10.1109/ACCESS.2023.3241234.
3) Ranjan M, Tiwari S, Sattar AM, Tatkar NS. A New Approach for Carrying Out Sentiment Analysis of Social Media Comments Using Natural Language Processing. *Engineering Proceedings*. 2024;59(1):181–181. Available from: https://doi.org/10.3390/engproc2023059181.
4) Liu H, Wei R, Tu G, Lin J, Liu C, Jiang D. Sarcasm driven by sentiment: A sentiment-aware hierarchical fusion network for multimodal sarcasm detection. *Information Fusion*. 2024;108:102353–102353. Available from: https://doi.org/10.1016/j.inffus.2024.102353.
5) Tan YY, Chow, Co, Kanesan J. Sentiment Analysis and Sarcasm Detection using Deep Multi-Task Learning. *Wireless Pers Commun*. 2023;129:2213–2237. Available from: https://doi.org/10.1007/s11277-023-10235-4.
6) Sharma DK, Singh B, Agarwal S, Pachauri N, Alhussan AA, Abdallah HA. Sarcasm detection over social media platforms using hybrid ensemble model with fuzzy logic. *Electronics*. 2023;12(4):937–937. Available from: https://doi.org/10.3390/electronics12040937.
7) Anan R, Apon TS, Hossain ZT, Modhu EA, Mondal S, Alam M. Interpretable Bangla sarcasm detection using BERT and explainable AI. *2023 IEEE 13th annual computing and communication workshop and conference (CCWC);* . 2023;p. 1272–1278. Available from: http://doi:10.1109/CCWC57344.2023.10099331.
8) Mohan A, Nair AM, Jayakumar B, Muraleedharan S. Sarcasm detection using bidirectional encoder representations from transformers and graph convolutional networks. *Procedia Computer Science*. 2023;218:93–102. Available from: https://doi.org/10.1016/j.procs.2022.12.405.
9) Zhang Y, Ma D, Tiwari P, Zhang C, Masud M, Shorfuzzaman M, et al. Stance-level sarcasm detection with BERT and stance-centered graph attention networks. *ACM Transactions on Internet Technology*. 2023;23(2):1–21. Available from: https://doi.org/10.1145/3533430.

10) Lin H, Chen Z, Luo Z, Cheng M, Ma J, Chen G. CofiPara: A Coarse-to-fine paradigm for multimodal sarcasm target identification with large multimodal models. *arXiv preprint* . 2024. Available from: https://doi.org/10.48550/arXiv.2405.00390.
11) Zhu Z, Cheng X, Hu G, Li Y, Huang Z, Zou Y. Towards multi-modal sarcasm detection via disentangled multi-grained multi-modal distilling. *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*. 2024;p. 16581–16591. Available from: https://doi.org/10.18653/v1/P19-1455.
12) Yacoub AD, Slim S, Aboutabl A. A survey of sentiment analysis and sarcasm detection: Challenges, techniques, and trends. *International Journal of Electrical and Computer Engineering Systems*. 2024;15(1):69–78. Available from: https://doi.org/10.32985/ijeces.15.1.7.
13) Zhong W, Zhang Z, Wu Q, Xue Y, Cai Q. A Semantic Enhancement Framework for. *Multimodal Sarcasm Detection Mathematics*. 2024;12(2):317–317. Available from: https://doi.org/10.3390/math12020317.
14) Ray A, Mishra S, Nunna A, Bhattacharyya P. A multimodal corpus for emotion recognition in sarcasm. *arXiv preprin*. 2022. Available from: https://doi.org/10.48550/arXiv.2206.02119.
15) Helal NA, Hassan A, Badr NL, Afify YM. A contextual-based approach for sarcasm detection. *Scientific Reports*. 2024;14(1):15415–15415. Available from: https://doi.org/10.1038/s41598-024-65217-8.