JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

TONE TRACER: DECODING THE HIDDEN MEANING

A Context-Aware Transformer-Based Framework for Sarcasm Detection

[1]Riddhisha Srivastava, [2]Anurag Shrivastava, [3]Sachin Tripathi, [4]Anshika Mishra, [5]Ishaan Tripathi

[1]B.Tech Student, [2]Professor, [3]B.Tech Student, [4]B.Tech Student, [5]B.Tech Student
[1]Engineering (All Branches),
[1]Babu Banarasi Das Northern India Institute of Technology, Lucknow, India

Abstract: Since sarcasm involves saying the opposite of what one means, it is still a problem for computers, classifying as one of the main obstacles for most sentiment analysis programs, especially when addressing social media sites. However, the more frequent approach for typical sarcasm identification algorithms is to rely almost entirely on surface text, although recent approaches with deep neural networks have only found a measure of success, as most have typically processed a sentence-by-sentence basis. Even the most successful algorithms regarding performance, which make use of transformers, fail at the conversational aspect of the task.

In this work, we introduce **Tone Tracer**, a sarcasm detection system that is context aware, using transformer encoders in such a way that the system is supplied with the segments of the targeted speech and the accompanying context in order to detect the contradictions implicit in the tone shifts that are commonly found in the tone of sarcastic speech. Evaluation on benchmark datasets indicates that Tone Tracer performs better than the existing machine learning models, deep learning models, and the transformer models; therefore, sarcasm detection is better considered at the discourse level rather than at the sentence level.

***Index Terms* - Sarcasm Detection, Context-Aware Learning, Transformer Architecture, Natural Language Processing, Conversational Analysis.**

I. INTRODUCTION

Sarcasm is very common in communication and even more prevalent in social networking sites, discussion boards, and conversational interfaces. It helps in communicating negative opinions or criticisms indirectly. Although computers can easily detect sarcasm with the help of domain and background knowledge, they cannot easily determine the sarcastic intentions of the statements since they are not explicitly stated.

Early sentiment analysis tools considered the polarity of the sentiment as the direct representation of the user's opinion. Nevertheless, Bhattacharyya et al. **[1]** proved that the existence of sarcasm is one of the factors responsible for the incorrect classification of the sentiment. Moreover, recent studies proved that sarcasm tends to utilize the concept of incongruity of context, where the actual message is revealed only after inspecting the context of the surrounding conversations **[6]**. Even though deep learning models enhanced the understanding of language **[2][4]**, many sarcasm detection models inspect the sentences individually.

To address this issue, we present the concept of **Tone Tracer**, a solution that takes the conversation context into account when detecting sarcasm. Our solution is based on previous works done by **[5][6]** on context-based sarcasm representation, which have been extended through a transformer architecture.

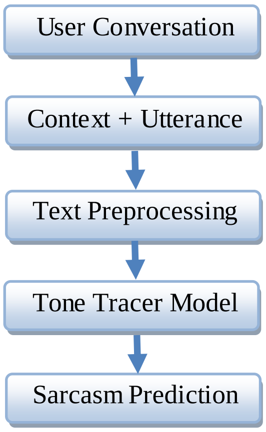


Figure 1 Overall Workflow of Sarcasm Detection

II. RELATED WORK

Research in sarcasm detection has undergone substantial advancement in the past decade. Current state-of-the-art methods for sarcasm detection can be broadly divided into traditional text-based methods, deep learning methods, transformer methods, context-aware methods, and multimodal methods.

2.1 Traditional and Text-Based Approaches

Early sarcasm detection methods relied on a bunch of handcrafted features related to the language, such as out-of-place sentiment polarity, punctuation patterns, capitalization, quotation marks, and emoticons. The features were then used to train classical machine learning models like Support Vector Machines and Naïve Bayes. Bhattacharyya et al. **[1]** summarize a number of these works and point out their deficiencies in generalization and robustness.

Joshi et al. **[6]** further emphasized that sarcasm is closely related to contextual incongruity, which is not modelled by traditional sentence-level methods. These are interpretable models; however, they do not capture deeper semantic and pragmatic meaning.

2.2 Deep Learning Approaches for Sarcasm Detection

With the use of deep learning, the necessity for manual feature engineering reduced. Representation learning was automatically done by models such as CNNs and LSTMs from the text data itself. In this regard, Joshi et al. **[3]** have also shown that neural networks are much better at capturing sentiment contrast than the traditional methods.

However, Tay et al. **[9]** pointed out that most of the deep learning models still treat sarcasm detection as a sentence-level task. Thus, they cannot handle discourse level dependent sarcasm that does not rely on internal sentence patterns.

2.3 Transformer-Based Models

The transformer architectures introduced contextualized word representations with the help of self-attention mechanisms. Devlin et al. **[2]** came up with BERT that remarkably outperformed previous results in a wide range of NLP tasks. Mishra et al. **[4]** explored transformer models for figurative language processing, which includes sarcasm detection.

Despite these, transformer-based sarcasm detectors ignore the conversational history. Dong et al. **[5]** demonstrated that sentence-level transformers often misclassify sarcasm when the sarcastic intent depends on the prior turns in a dialogue.

2.4 Context-Aware Sarcasm Detection

Context-aware approaches tend to take the conversation history into account explicitly. It was shown in **[5]**, for example, that the modeling of preceding statements improves the accuracy of detection. Similarly, it was found in **[6]**, for instance, that sarcasm arises through contrast in a discourse rather than a single statement.

However, it is challenged by noisy contextual information and unnecessary dialogue turns. Tone Tracer proposes a solution for this problem through the application of transformer self-attention to identify relevant contextual information.

2.5 Multimodal Sarcasm Detection

The multimodal methods incorporate text along with images, speech, or videos. They are able to capture the contradictions between the modes. The multimodal sarcasm detection methods are introduced by Cai et al. **[7]** based on cross-modal attention. The survey by Kumar et al. **[11]** on IJCAI presents a detailed description of such methods.

Despite the great results, the requirements for large annotated data and complex architectures make scaling a problem for multimodal modeling. The scoping to design better text-based and contextually aware solutions like Tone Tracer emerges here.

2.6 Research Gap

Based on existing literature **[1–11]**, the following gaps are identified:

- Over-reliance on sentence-level analysis
- Limited discourse-level reasoning
- High computational cost of multimodal systems
- Poor generalization across domains

Tone Tracer aims to bridge these gaps by offering a scalable, text-based, context-aware transformer framework.

III. Task Definition

Sarcasm detection is posed as a supervised classification problem with a binary output. For a conversation context C and a targeted utterance U, it tries to predict the likelihood of U being a sarcastic message. Sarcasm detection differs from senti ment analysis as it aims to identify implied meanings and not explicit words of sentiment, as cited by Bhattacharyya et al. **[1]** and Joshi et al. **[6]**.

IV. Dataset Description

The corpora used in this research are harvested from Twitter and online discussion boards, where the use of sarcasm is a frequent phenomenon. By using hashtags related to sarcasm for distant supervision, a large amount of data is collected **[1]**, while the drawbacks of label quality are overcome in Text-based Sarcasm Detection on Social Networks **[10]**.

The conversational context is formed by considering a fixed number of preceding statements in the conversation, adhering to the best practices for the design of the context in the field of context-aware NLP studies **[5]**. The datasets contain imbalanced classes, colloquial language, and diverse topics.

V. Proposed Methodology: Tone Tracer

5.1 System Overview

Tone Tracer jointly encodes conversational context and target utterances using a transformer-based architecture inspired by BERT **[2]** and context-aware sarcasm modeling strategies **[5]**.

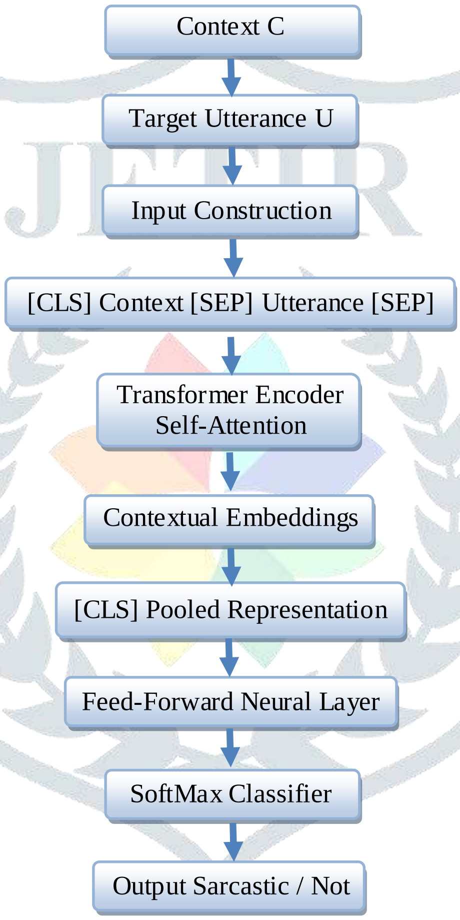
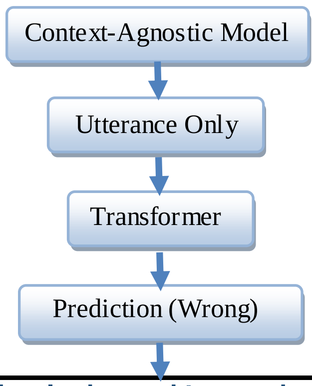


Figure 2 Architecture of Tone Tracer

5.2 Context-Aware Encoding

The self-attention mechanism also enables tokens in the context to affect the interpretation of the utterance. This is similar to findings done by Dong et al. **[5]**, regarding the underlying contradictions and tone shifts.



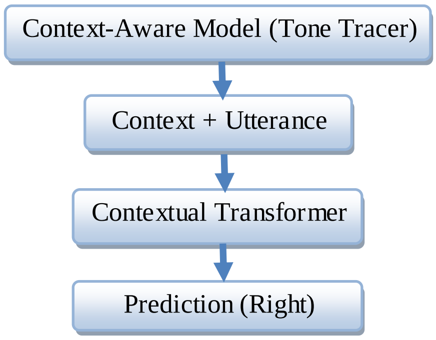


Figure 3 Context-Aware vs Context-Agnostic Models

VI. Experimental Setup

The splits of the dataset into the validation and test components, as well as the training set, are done through stratified sampling. The results of the Tone Tracer algorithm will be compared to the results of the conventional machine learning models **[1]**, deep models **[3]**, and the sentence-level transformer models **[2]**.

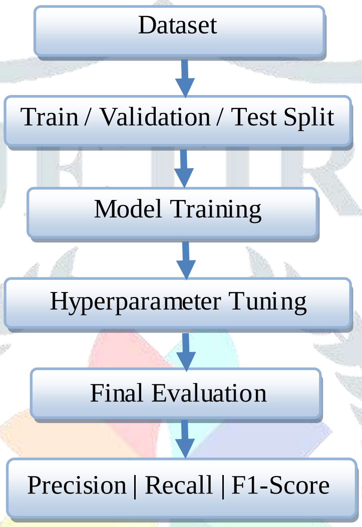


Figure 4 Experimental Evaluation Pipeline

VII. Evaluation Metrics

Performance is evaluated using precision, recall, macro-averaged F1-score, and accuracy. The F1-score is emphasized due to class imbalance, following evaluation practices in prior sarcasm detection research **[1][9]**.

VIII. Results And Analysis

The experimental outcomes verify Tone Tracer leads to better performance compared to baseline models, mostly in the recognition of sarcastic examples regarding the recall metric. Such evidence assures improved handling of subtle cases of sarcasm with implicit features, in accordance with the findings of previous works **[5][8]**.

Common Error Categories

- Cultural References
- World Knowledge
- Subtle Irony
- Ambiguous Context

IX. Discussion

It is clear from the experiment that sarcasm is a discourse-level phenomenon in the sense that sentence-level models tend to be misled by instances of sarcasm, whereas modeling with an understanding of the context proves to be beneficial for the task. Upon analysis of the error report, observations are as follows: Cultural Refs/World Knowledge – as predicted by Kumar et al. **[8]**.

X. Limitations

The Tone Tracer tool depends purely on text input. It does not include visual and audio aspects, which are discussed under multimodal methods **[7][11]**. The transformer-based methods consume a lot of computer processing capacity. The marking of sarcasm also has a subjective component.

XI. Applications

Tone Tracer has many potential usages, such as for sentiment analysis systems, social media monitors, chatbots, as well as for content moderators. Greater precision for sarcasm detection allows for a better understanding of intentions.

XII. Future Work

Future research may focus on integrating multimodal signals **[7][11]**, incorporating commonsense reasoning **[8]**, improving domain generalization, and developing lightweight architectures for real-time deployment.

Future Extensions of Tone Tracer

- Multimodal Inputs
- Commonsense Knowledge
- Domain Adaptation
- Lightweight Deployment

XIII. Conclusions

This paper presents the Tone Tracer, a context-aware transformer-based framework for the identification of sarcasm. Explicitly modeling the conversational context, the framework effectively decodes hidden sarcastic meaning missed by sentence-level models. Results establish that contextual reasoning forms an important step toward the correct identification of sarcasm and set a strong base for future research.

References

[1] Bhattacharyya, P., Kumar, A., & Banerjee, S. (2017). *A survey on sarcasm detection*. ACM Computing Surveys, 50(5), 1–38.

[2] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of deep bidirectional transformers for language understanding*. NAACL-HLT.

[3] Joshi, A., Sharma, V., & Bhattacharyya, P. (2016). *Harnessing context incongruity for sarcasm detection*. ACL.

[4] Mishra, S., Dandapat, P., & Ekbal, A. (2020). *Figurative language identification using contextualized word representations*. FigLang.

[5] Dong, Y., Chen, J., & Yang, X. (2020). *Modeling conversational context for sarcasm detection*. AAAI.

[6] Joshi, A., Sharma, V., & Bhattacharyya, P. (2015). *Towards sarcasm detection using contextual information*. ACL.

[7] Cai, H., Sun, Y., & Wang, X. (2023). *Multimodal sarcasm detection with cross-modal attention*. ACL.

[8] Kumar, R., Singh, A., & Mehta, S. (2024). *Contextual reasoning for sarcasm detection*. NAACL.

[9] Tay, Y., Tuan, L. A., & Hui, S. C. (2022). *Sarcasm detection using neural representation learning*. IJECES.

[10] *Text-based Sarcasm Detection on Social Networks*. (2023). IJCA.

[11] *Multimodal Sarcasm Detection: A Survey*. (2024). IJCAI.

[12] Shrivastava, A., Mishra, A., Tripathi, I., Srivastava, R., & Tripathi, S. (2025). *Tone Tracer: Decoding the Hidden Meaning*. BBDNIIT, Lucknow, India.