

Paper Title: Interpretable Bangla Sarcasm Detection using BERT and Explainable AI
Paper Link: [2303.12772.pdf \(arxiv.org\)](https://arxiv.org/pdf/2303.12772.pdf)

Motivation:

In "Interpretable Bangla Sarcasm Detection using BERT and Explainable AI" authors aimed to address the challenge of detecting sarcasm in the Bangla language. Sarcasm detection is a great challenge in NLP. So several researches have been done on the English language to detect sarcasm. Meanwhile, not a lot of noticeable studies have been done on this paper for the Bangla language. The sarcasm detection system in this paper was proposed to assist systems in many tasks. Such as sentiment analysis, opinion mining, and marketing tools. So that it can provide a better understanding of sarcasm in the Bangla language. Here optimal models are analyzed and designed to work with the specific challenges and characteristics that can tackle the Bangla language. So that it can better analyze the Bangla language.

Contribution:

"Interpretable Bangla Sarcasm Detection using BERT and Explainable AI" authors made a BERT-based sarcasm detection model which works on the Bangla language. The authors have developed a novel approach to detect sarcasm in the Bangla language using the BERT model. This showed promising results in natural language processing tasks. The research addresses the gap in sarcasm detection between the Bangla and English languages, providing a more comprehensive understanding of sarcasm in various languages. The study has conducted a comprehensive study of various traditional machine learning models with the proposed system, aiming to improve the overall process and reduce the time required for analysis. The proposed model is designed to work with the specific challenges and characteristics of the Bangla language.

Methodology:

The methodology of the paper "Interpretable Bangla Sarcasm Detection using BERT and Explainable AI" is:

1. Data collection: The authors have compiled a dataset of 7800 social media comments, both positive and negative, that have been tagged with sarcasm. like Twitter, YouTube, and additional training and testing resources.
2. Preprocessing: To reduce the size of the vocabulary, tokenization was used to remove stop words, emojis, punctuation, and extra spaces from the text data.

3. Model selection: In order to determine the best framework for sarcasm detection, the study thoroughly examined a number of conventional machine learning models, such as Random Forest, Decision Tree, K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Multinomial Naive Bayes, Logistic Regression, Stochastic Gradient Descent (SGD), and Random Forest.
4. Model development: Using the gathered dataset, the authors suggested a BERT-based sarcasm detection model for the Bangla language, which was trained and checked. Additionally, they used LIME to add explainability to their system.
5. Evaluation: Using the F1-score, accuracy, precision, recall, and area under the ROC curve (AUC-ROC), the model's performance was assessed.

Conclusion:

In summary, the research "Interpretable Bangla Sarcasm Detection using BERT and Explainable AI" suggests a novel method for utilizing BERT and Explainable AI to identify sarcasm in Bangla language usage. On the Bangla sarcasm detection task, it performed notably well and outperformed other algorithms. The introduction of explainable AI aimed to shed light on the model's decision-making procedure. The suggested method is transferable to other languages and has potential uses in social media monitoring, sentiment analysis, and customer feedback analysis, among other areas. While providing a substantial addition to the field of natural language processing, this work also leaves room for further investigation.

Limitations:

The limitations are-

1. Data quality and availability: The annotated Bangla sarcasm dataset's quality and availability determine how well the model performs.
2. Scope of the research: The study may be limited in its scope to a specific type of Bangla text or social media platforms, which could affect the proposed BERT-based model in identifying other types.

Synthesis: The study "Interpretable Bangla Sarcasm Detection using BERT and Explainable AI" describes a BERT-based system that outperforms conventional machine

learning algorithms in sarcasm identification in the Bangla language, achieving 99.60% accuracy. To give the system explainability, the authors additionally present Local Interpretable Model-Agnostic Explanations. The article discusses the difficulties caused by the contrasting meanings that sarcasm implies, as well as the paucity of studies on sarcasm detection in Bangla. The suggested method advances natural language processing for sarcasm detection in conjunction with the recently gathered BanglaSarc dataset. The study's conclusions have a big impact on how NLP-based systems for threat identification, sentiment analysis, and other uses in the Bangla language are developed.