Bangla Abusive Language Detection Using Machine Learning on Radio Message Gateway

[Sumaiya Salim Ritu](https://ieeexplore.ieee.org/author/37088927678), [Joysurya Mondal](https://ieeexplore.ieee.org/author/37088927046), [Md. Moinu Mia](https://ieeexplore.ieee.org/author/37088925389), [Ahmed Al Marouf](https://ieeexplore.ieee.org/author/37085695654)

**Abstract:**

In the era of modern technology, machine learning and natural language processing has been adopted to be applied in several application areas. Natural language processing consists of diversified techniques such as text classification, text summarization, named entity recognition, sentiment analysis. Text classification is considered to be the area of research where the text gets segmented into different category sentences or paragraphs from a single text genre. This paper presents a mechanism for detecting Bangla abusive language from a real-time radio message gateway. Online radio stations nowadays accept communications and voices of their target audience from web-based applications or social media platforms, such as Facebook or Twitter pages. This paper has created a dataset with more than 45000 Bangla sentences, which are labeled as abusive and non-abusive. Sample online radio message gateway has been introduced and machine learning algorithms such as multinomial naive bias (MNB), logistic regression (LR), and random forest (RF) classifiers are utilized to predict the abusive languages. One of the significant prospects of this work would be applied during live radio programs where listeners try to communicate by sending live messages. Our proposed mechanism can check and map the live messages with the dataset and segregate the positive comments or messages only, by filtering the abusive comments. Among the applied classifiers, it has been found that the random forest classifier has performed better than the other two classifiers by leveraging approximately 76% accuracy.

**Conference / Journal Link:**

<https://ieeexplore.ieee.org/document/9489131>

**DOI:**

[10.1109/ICCES51350.2021.9489131](https://doi.org/10.1109/ICCES51350.2021.9489131)