|  |  |  |  |
| --- | --- | --- | --- |
| Title | G-BERT: An Efficient Method for Identifying Hate Speech in Bengali Texts on Social Media | | |
| Author(s) Name | Ashfia Jannat Keya, Md Mohsin Kabir, Nusrat Jahan Shammey, MF Mridha, Md Rashedul Islam, Yutaka Watanobe | | |
| Contact Email(s) | firoz.mridha@aiub.edu | | |
| Published Journal Name | IEEE Access | | |
| Type of Publication | Journal | | |
| Volume | 11 | Issue |  |
| Publisher | IEEE | | |
| Publication Date | 2023/7/26 | | |
| ISSN | 2169-3536 | | |
| DOI | [10.1109/ACCESS.2023.3299021](https://doi.org/10.1109/ACCESS.2023.3299021) | | |
| URL | <https://ieeexplore.ieee.org/abstract/document/10194932> | | |
| Other Related Info. |  | | |
|  | | | |

|  |  |
| --- | --- |
| Abstract |  |
| The rapid increase in Internet users has increased online concerns such as hate speech, abusive texts, and harassment. In Bangladesh, hate text in Bengali is frequently used on various social media platforms to condemn and abuse individuals. However, Research on recognizing hate speech in Bengali texts is lacking. The pervasive negative impact of hate speech on individuals’ well-being and the urgent need for effective measures to address hate speech in Bengali texts have created a significant research gap in the Bengali hate speech detection field. This study suggests a technique for identifying hate speech in Bengali social media posts that may harm individuals’ sentiments. Our approach utilizes the Bidirectional Encoder Representations from Transformers (BERT) architecture to extract Bengali text properties, whereas hate speech is categorized using a Gated Recurrent Units (GRU) model with a Softmax activation function. We propose a new model, G-BERT, that combines both models. We compared our model’s performance with several other algorithms and achieved an accuracy, precision, recall, and F1-score of 95.56%, 95.07%, 93.63%, and 92.15%, respectively. Our proposed model outperformed all other classification algorithms tested. Our findings show that the strategy we have suggested is successful in locating hate speech in Bengali texts posted on social media platforms, which can aid in mitigating online hate speech and promoting a more respectful online environment. | |