

Title: Cyberbullying detection from Bangla text on social media

Author Names: Imtiaz Ahmmed Asif, Apurbo Roy, Md. Al Mamun, Md. Tanvir Siddiquee, Sumit Kumar Banshal

Abstract: The potential of social media is expanding as more and more people utilize it. As more individuals use social media, however, bullying in the comment sections of posts by well-known users and of viral material is also on the rise. This number of bullying texts is on the rise and should be eliminated prior to being shown. Using natural language processing and classifier techniques, we identify cyberbullying in this article. We created our data by ourselves. We receive 3500 records, of which 22.1% pertain to bullying and 77.9% do not. After the data were prepared for the classifier model, they were separated into training and testing groups. Multinomial naive Bayes had an accuracy rate of 78.99%, whereas a decision tree classifier had an accuracy rate of 69.48%. The k-nearest neighbor classifier required the shortest time, at 0.0018 seconds, whereas the random forest classifier required the longest time, at 1.44 seconds.

Key words: Cyberbullying detection Social media Natural language processing (NLP) Machine learning

<https://www.taylorfrancis.com/chapters/edit/10.1201/9781003433958-12/cyberbullying-detection-bangla-text-social-media-imtiaz-ahmmed-asif-apurbo-roy-md-al-mamun-md-tanvir-siddiquee-sumit-kumar-banshal>