

Machine Learning Automatic Detection on Twitter Cyberbullying

Authors : Raghad A. Altowairgi

Abstract : With the wide spread of social media platforms, young people tend to use them extensively as the first means of communication due to their ease and modernity. But these platforms often create a fertile ground for bullies to practice their aggressive behavior against their victims. Platform usage cannot be reduced, but intelligent mechanisms can be implemented to reduce the abuse. This is where machine learning comes in. Understanding and classifying text can be helpful in order to minimize the act of cyberbullying. Artificial intelligence techniques have expanded to formulate an applied tool to address the phenomenon of cyberbullying. In this research, machine learning models are built to classify text into two classes; cyberbullying and non-cyberbullying. After preprocessing the data in 4 stages; removing characters that do not provide meaningful information to the models, tokenization, removing stop words, and lowering text. BoW and TF-IDF are used as the main features for the five classifiers, which are; logistic regression, Naïve Bayes, Random Forest, XGboost, and Catboost classifiers. Each of them scores 92%, 90%, 92%, 91%, 86% respectively.

Keywords : cyberbullying, machine learning, Bag-of-Words, term frequency-inverse document frequency, natural language processing, Catboost

Conference Title : ICACML 2023 : International Conference on Adaptive Computation and Machine Learning

Conference Location : Jeddah, Saudi Arabia

Conference Dates : February 20-21, 2023