

ABSTRACT

Hatespeech is a form of verbal communication with purpose to insult, to provoke or to incite the victim. With the help of social media such as Twitter, it become easier to spread or find hate speech. So, in order to reduce online hate speech we made a classification system to detect hate speech from Twitter. This system is using LightGBM, which was the development of Gradient Boosting Decision Tree (GBDT) Algorithm. GBDT was often used for this type of classification but the outcome was less satisfactory. LightGBM use Gradient-based One Side Sampling (GOSS) and Exclusive Feature Bundling (EFB). With this two technique implemented in GBDT, we hope the outcome will be more accurate with less computation time.

Keyword: Hate Speech, Twitter, LightGBM