PROJECT REPORT

TOXIC COMMENT CLASSIFICATION COMPARISON BETWEEN LSTM, BILSTM, GRU, AND BIGRU

JONATHAN KEVIN GIUSTINO
20.K1.0009

Faculty of Computer Science
Soegijapranata Catholic University
2024

ABSTRACT

One of the biggest problems in modern internet age is toxicity. In this study, our aim is to draw an effective method to classify toxicity in text comment in this case specifically on Wikipedia comment. Old Regular Rule based model attacks, Machine Learning Methods suffer from rules based approaches and thus are incapable of accurate detection of toxicity while maintaining precision at the same time. To overcome this limitation, recurrent neural networks (RNNs), wherein, long short-term memory (LSTM) networks, gated recurrent unit(GRU) have been proposed. In this paper, author compares the LSTM, BiLSTM, GRU, and BiGRU for multi label classification for which is the best model to use towards jigsaw toxicity challenge dataset to classify toxicity. This study will be finding out which of the model is the best for classification and the difference between different type of pre-processing. We'll go ahead to use League of Legends tribunal datasets from kaggle as our base. The results obtained were that, the highest accuracy on the test was attained by BiLSTM model without cleaning with on 87.208% accuracy, 55.205% Precision, 68.540% Recall, and 60.623% F1-Score the result also shows while preprocessing on cleaning improve the resultant metrics by a marginal amount for regular LSTM, GRU, and BiGRU it doesn't always improve the result.

Keyword: Toxicity, LSTM, GRU, BiLSTM, BiGRU