الإنتاج العلمي لمركز بحوث علوم الحاسب والمعلومات

VISION 2030
المملكة العربية السعودية
KINGDOM OF SAUDI ARABIA

عمادة البحث العلمي
Deanship of Academic Research

BiCHAT: BiLSTM with deep CNN and hierarchical attention for hate speech detection

Authors   Shakir Khan, Mohd Fazil, Vineet Kumar Sejwal, Mohammed Ali Alshara, Reemiah Muneer Alotaibi, Ashraf Kamal, Abdul Rauf Baig

Abstract: Online social networks(OSNs) face the challenging problem of hate speech, which should be moderated for the growth of OSNs. The machine learning approaches dominate the existing set of approaches for hate speech detection. In this study, we introduce BiCHAT: a novel BiLSTM with deep CNN and Hierarchical ATtention-based deep learning model for tweet representation learning toward hate speech detection. The proposed model takes the tweets as input and passes through a BERT layer followed by an attention-aware deep convolutional layer. The convolutional encoded representation further passes through an attention-aware Bidirectional LSTM network. Finally, the model labels the tweet as hateful or normal through a softmax layer. The proposed model is trained and evaluated over the three benchmark datasets extracted from Twitter and outperforms the state-of-the-art (SOTA) (Khan et al., 2022; Roy et al., 2020; Ding et al., 2019) and baseline methods with an improvement of and in terms of precision, recall, and f-score, respectively. BiCHAT also demonstrates good performance considering training and validation accuracy with an improvement of and , respectively. We also examined the impact of different constituting neural network components on the model. On analysis, we observed that the exclusion of the deep convolutional layer has the highest impact on the performance of the proposed model. We also investigated the efficacy of different embedding techniques, activation function, batch size, and optimization algorithms on the performance of the BiCHAT model.