

Abstrak

Dalam penelitian ini, penulis menggunakan metode BERT untuk mengidentifikasi komentar toksik. Penulis membandingkan model yang dihasilkan dari dataset *User Generated Content*(UGC) dan *dataset* UGC yang telah dinormalisasikan. Arsitektur BERT yang telah dipre-train ditambah dengan *output layer* untuk klasifikasi teks. Hasil akhir dari penelitian tugas akhir ini menyatakan bahwa menormalisasikan dataset UGC terlebih dahulu tidak diperlukan d alam pelatihan model BERT.

Kata kunci: klasifikasi teks, deep learning, bert, transformers, attention, klasifikasi