

ABSTRAK

Pada masa ini penerapan bahasa yang banyak muncul di media sosial adalah penerapan bahasa slang sebagai karakter perkembangan bahasa saat ini. Fenomena ini layak untuk dibahas karena dapat mengetahui seberapa banyak bahasa slang digunakan pada media sosial. Sumber dataset untuk penelitian ini adalah bentuk verbal yang terdapat pada media sosial pribadi penulis seperti Instagram dan Tiktok yang diperoleh dengan metode web scraping sebanyak 2.000 sampel dan data tersebut akan dibagi menjadi dua kategori yakni kategori bahasa slang dan bukan bahasa slang. Penelitian ini bertujuan untuk membandingkan dua algoritma klasifikasi yaitu Naïve Bayes Dan Decision Tree untuk melihat algoritma mana yang lebih efektif dalam melakukan klasifikasi seberapa banyak pengguna media sosial menggunakan bahasa slang dalam berkomentar berdasarkan dataset yang sudah kami kumpulkan, sehingga diperoleh hasil untuk melihat seberapa tinggi persentase penggunaan bahasa slang masyarakat indonesia dalam berkomentar di media sosial.

Kata kunci bahasa slang, klasifikasi, decision tree, Naïve Bayes