
Abstrak

Berita adalah salah satu sarana informasi bagi masyarakat umum. Di jaman sekarang dengan jumlah artikel berita yang mencapai 2 juta artikel perhari dapat menyusahkan pengguna dalam menemukan artikel berita yang ingin mereka baca. Agar mempermudah pengguna kebanyakan koran berita Indonesia mengklasifikasikan artikel mereka ke dalam kategori tertentu, tetapi banyak juga *blog*, atau artikel amatir yang belum mengklasifikasikan berita yang mereka edarkan. Oleh karena itu penelitian ini bertujuan mengkategorikan berita berbahasa Indonesia dengan menggunakan metode *weighted k-nearest neighbour*. Pada penelitian ini terdapat beberapa tahap dalam mengklasifikasi berita tersebut yaitu *preprocessing*, *feature extraction*, dan pengklasifikasian menggunakan wK-NN. Penelitian menggunakan metode wK-NN dimana $K = 6$. Pada penelitian ini dilakukan ekstraksi fitur dalam *unigram* dan *bigram* yang menghasilkan akurasi yang tidak jauh berbeda. Sehingga disarankan untuk menggunakan *unigram* karena lebih efisien.

Kata kunci: Klasifikasi, *Preprocessing*, *feature extraction*, *weighted K-nearest neighbour*
