

ABSTRAK

Kategorisasi dokumen teks merupakan suatu kegiatan yang dilakukan untuk mengelompokkan dokumen berbentuk teks untuk ditempatkan ke dalam kategori tertentu sesuai dengan kesamaan ciri dari topik yang dibicarakan pada dokumen teks tersebut. Metode yang dapat digunakan untuk pengelompokkan teks sangat banyak dan beragam. Namun, dari sekian banyak metode dipilihlah metode yang memiliki kinerja baik. Untuk pemilihan metode dapat dilakukan dengan membandingkan antara metode pengkategorian dokumen teks satu dengan metode yang lainnya. Maka dari itu pada penelitian ini akan membandingkan kinerja antara metode *k-nearest neighbor* dengan *naïve bayes classifier* dalam mengkategorikan artikel berita untuk mendapatkan metode yang baik dalam proses pengkategorian dokumen teks. Pada penelitian ini data yang digunakan untuk pengujian terdiri dari 28 data latih yang tersebar dalam 5 kategori dan 24 data uji. Pengujian dilakukan dengan menggunakan akurasi ketepatan kategori dan waktu proses pengkategorian sebagai parameter untuk menentukan metode yang lebih baik. Hasil pengujian menunjukkan tingkat akurasi metode KNN adalah 62,50% sedangkan *naïve bayes classifier* memiliki akurasi 54,17%. Untuk pengujian waktu yang dibutuhkan untuk mengklasifikasikan 15 dokumen uji rata-rata menggunakan metode *k-nearest neighbour* membutuhkan waktu selama 671,65 detik, sedangkan proses pengkategorian menggunakan metode *naïve bayes* membutuhkan waktu rata-rata selama 1362,83 detik. Penelitian ini menyimpulkan bahwa metode *k-nearest neighbour* lebih efektif dan efisien untuk melakukan proses kategorisasi pada dokumen berita (dokumen teks).

Kata kunci : artikel berita, *k-nearest neighbour*, *naïve bayes classifier*, kategorisasi dokumen.

UIN

UNIVERSITAS ISLAM NEGERI
SUNAN GUNUNG DJATI
BANDUNG