

## ABSTRAK

Artikel merupakan suatu karya tulis yang mempunyai sifat faktual serta terdapat pendapat atau ide seseorang mengenai masalah tertentu. Artikel memiliki tujuan untuk mendidik, mempengaruhi, memberitahu, meyakinkan, atau dapat menghibur dengan cara mempublikasikannya. Dengan banyaknya persebaran berita saat ini, maka akan sulit untuk menentukan suatu berita tersebut tergolong dalam berita yang benar adanya atau berita hoax. Begitupun dengan berita mengenai kesehatan. Tidak dapat dipungkiri ketika seseorang mendapat berita mengenai kesehatan pasti akan membacanya dan kemudian akan melennya mentah-mentah. Informasi yang ada di berita mengenai kesehatan tersebut pasti akan mempengaruhi pemikiran seseorang mengenai kondisi tubuhnya. Oleh karena itu dibutuhkan suatu sistem cerdas untuk mengklasifikasikan artikel kesehatan tersebut untuk mengidentifikasi apakah termasuk dalam kategori hoax atau fakta.

Algoritma Kombinasi *K-Nearest Neighbor* dan *Naïve Bayes* merupakan algoritma klasifikasi yang menggabungkan algoritma *K-Nearest Neighbor* dengan algoritma *Naïve Bayes*. Penggabungan kedua metode ini berguna untuk mempercepat kinerja *K-Nearest Neighbor* sehingga tidak perlu menghitung keseluruhan data, tetapi menghitung dari probabilitas yang mungkin. Sehingga dengan metode kombinasi tersebut dapat memperoleh hasil yang lebih baik. Untuk proses yang dilakukan dalam klasifikasi yaitu mengumpulkan data, melakukan tahap *preprocessing* (*cleansing, case folding, tokenizing, stopword removal* dan *stemming*), pembobotan fitur menggunakan metode *Term Frequency Invers Document Frequency*, melakukan tahap klasifikasi (menghitung probabilitas dengan *Naïve Bayes* dan dilanjutkan klasifikasi dengan *K-Nearest Neighbor*). Pada tahap pengujian menggunakan tabel *confusion matrix* untuk mencari nilai akurasi, presisi, dan *recall*.

Dari pengujian tersebut menghasilkan nilai akurasi optimum untuk klasifikasi artikel kesehatan sebesar 88%, untuk nilai presisi optimum sebesar 83%, sedangkan untuk nilai *recall* optimum sebesar 100%. Dari pengujian tersebut dihasilkan nilai akurasi yang cukup tinggi sehingga algoritma tersebut cocok diterapkan pada klasifikasi teks.

**Kata Kunci:** Klasifikasi Teks, *Text Mining*, Kombinasi *K-Nearest Neighbor* dan *Naïve Bayes*, *Confusion Matrix*