

## SARI

Penelitian ini bertujuan untuk menerapkan metode klasifikasi *Naïve Bayes* dan algoritma *Pseudo Labeling* dalam mengelompokkan dokumen medis yang sesuai dengan dokumen yang diinput sehingga menghasilkan kelompok-kelompok yang sesuai. Uji coba dilakukan dengan menggunakan sampel dokumen medis yang diambil dari sebuah media kesehatan elektronik berbasis web. Hasil eksperimen menunjukkan bahwa metode *Naïve Bayes* menggunakan teknik *Semi-Supervised Learning* dapat digunakan secara efektif untuk mengklasifikasikan dokumen medis. Hal ini terlihat dari hasil eksperimen, yaitu dengan porsi *labeled documents* dan *unlabeled documents* yaitu 400:300 mampu mengklasifikasi dokumen dengan tepat mencapai nilai rata-rata akurasi 83.20%. Selain itu, pada hasil eksperimen dengan porsi *labeled documents* dan *unlabeled documents* yaitu 300:400 mencapai nilai rata-rata akurasi 83.40%. Kemudian pada porsi 200:500 dan 100:600 dapat mengklasifikasi dokumen dengan tepat mencapai nilai rata-rata akurasi 81.00%. Hasil penelitian diatas menunjukkan bahwa jumlah *labeled documents* dan *unlabeled documents* sangat menentukan performa untuk melakukan klasifikasi.

Kata kunci: Medis, *Semi-Supervised Learning*, klasifikasi.