

ABSTRAK

Banyaknya informasi media *online* yang tersebar tentu keakuratan informasi yang disampaikan kepada *user* harus menjadi hal yang diutamakan. Tetapi sering ditemukannya kesalahan penulisan ataupun ejaan pada label berita yang membuat penyebaran informasi khususnya pada portal berita tidak tersampaikan dengan baik. Untuk mengatasi hal tersebut dibutuhkan pengoreksian pada setiap label yang salah sehingga menjadi label yang tepat. Akan tetapi masalah selanjutnya kembali muncul ketika dilakukannya pengoreksian label pada jumlah data yang besar sehingga meningkatkan penggunaan waktu yang besar pada pengoreksian label tersebut. Salah satu cara untuk melakukan pengoreksian pada label adalah dengan membandingkan setiap label pada portal berita dengan kamus sebagai acuan label yang benar. Dan untuk melakukan serta mempercepat proses pengoreksian pada jumlah data yang besar dapat dilakukannya pengelompokan label. Rekomendasi pengoreksian sistem *labeling* portal berita Indonesia yaitu *tribunnews.com* dengan tiga kategori yang sering diakses yaitu berita nasional, olahraga, dan selebriti digunakan sebagai sampel data. Metode yang digunakan untuk melakukan pengoreksian label adalah dengan menggunakan algoritma Damerau Levenshtein Distance yaitu dengan membandingkan setiap label pada data kamus. Dan diakhir dilakukan *Hierarchical Clustering* untuk mengelompokkan setiap label sebagai rekomendasi label. Hasil dari penelitian ini menghasilkan kluster yang menampilkan setiap rekomendasi label pada label yang dikoreksi. Nilai maksimum Damerau Levenshtein Distance pada Hierarchical Distance yang dibatasi menjadi 2 merupakan nilai acuan pengoreksian label yang tepat untuk menampilkan informasi yang relevan dan keakuratan pada informasi. Kesimpulan akhir dari penelitian ini adalah penggunaan algoritma Damerau Levenshtein Distance dapat membantu melakukan pengoreksian label dan penggunaan Hierarchical Clustering untuk menampilkan rekomendasi pengoreksian label dengan jumlah yang besar.

Kata Kunci : *label, Damerau Levenshtein Distance, clustering, Hierarchical Clustering*