

ABSTRAK

Karya ilmiah merupakan karya tulis yang berisi suatu pembahasan ilmiah yang dilakukan oleh seorang penulis, untuk memberitahukan sesuatu hal secara logis dan sistematis kepada para pembaca. Dalam penulisan karya ilmiah harus menghindari tindak kecurangan seperti plagiat yaitu pengambilan tulisan atau pemikiran orang lain yang diakui sebagai hasil tulisan atau hasil pemikirannya sendiri. Untuk menghindari hal tersebut maka perlunya dilakukan identifikasi kemiripan karya ilmiah. Salah satu algoritma yang dapat digunakan adalah algoritma k-nearest neighbor yaitu sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Untuk menghitung kedekatan/kemiripannya pada penelitian ini menggunakan 2 rumus yaitu cosine similarity dan euclidean distance. Hasil dari penelitian ini didapatkan dari 10 dokumen seluruhnya yang diuji yaitu 5 dokumen uji sampel yang telah diringkas dibandingkan kemiripannya terhadap 100 dokumen sampel yang telah di training dan disimpan kedalam database, untuk rumus cosine similarity tingkat keberhasilannya mencapai 100% sedangkan untuk rumus euclidean distance hanya 60% dan pada 5 dokumen uji luar sampel yang diidentifikasi kedua rumus tersebut mampu menghasilkan kemiripan yang relevan dengan dokumen yang diuji.

Kata kunci : *Karya Ilmiah, Identifikasi, K-Nearest Neighbor*