

ABSTRAKSI

Plagiarisme atau lebih dikenal sebagai plagiat sudah tidak asing lagi di telinga kita. Apalagi di kalangan para mahasiswa dengan tugas yang begitu banyak, tidak terelakkan lagi tindakan plagiat pasti akan mewarnai dunia pendidikan. Sebenarnya plagiat dapat diatasi dengan membangun sebuah sistem yang mampu mendeteksi kemiripan antar dokumen. Banyak algoritma yang dapat dipakai dalam membangun sistem ini, seperti Rabin Karp, Smith Waterman, Edit Distance, dan lain-lain. Semua algoritma tersebut mempunyai kelebihan masing-masing. Namun algoritma Edit Distance menawarkan proses perbandingan yang lebih cepat dan akurat. Untuk menambah keakuratan sistem diperlukan proses *preprocessing*. Proses *preprocessing* berisi *case folding*, *tokenizing*, *filtering*, dan *stemming*.

Kata Kunci : Plagiarisme, Edit Distance, Preprocessing, String Matching