

Abstrak

Perkembangan teknologi dan kebutuhan akan informasi menyebabkan penumpukan informasi dan data berupa dokumen teks baik dalam bentuk *softcopy* maupun *hardcopy*. Kondisi ini menyebabkan sulitnya menemukan dokumen yang sesuai dengan kebutuhan. Untuk memudahkan dalam menemukan dokumen yang sesuai kebutuhan, diperlukan kata kunci yang menyertai setiap dokumen. Kata kunci merupakan kata yang merepresentasikan isi sebuah dokumen secara keseluruhan. Dengan kata kunci, selain memudahkan dalam menemukan dokumen yang sesuai kebutuhan, juga membantu menemukan dokumen lain yang memiliki keterkaitan topik dengan sebuah dokumen.

Pada tugas akhir ini, diimplementasikan metode *Artificial Immune System* (AIS) berbasis Teori Informasi untuk mengekstrak kata kunci dari sebuah dokumen. Metode ini mengadopsi kemampuan Sistem Imun Alami dalam mengenali zat-zat yang berbahaya yang memasuki tubuh. Dengan metode ini setiap kata dihitung kandungan nilai informasinya untuk menentukan kata kunci. Proses *learning* dilakukan dengan menyimpan informasi setiap kata setiap kali dilakukan training ataupun ekstraksi untuk digunakan pada ekstraksi berikutnya.

Pengujian dilakukan untuk mendapatkan nilai minimal prosentase kandungan informasi terbaik pada sebuah kata dengan menggunakan metode AIS berbasis Teori Informasi berdasarkan parameter *precision*, *recall*, dan *f-measure* dalam memprediksi 10 *term* dengan nilai *tf-idf* yang tertinggi. Selain itu pengujian juga dilakukan untuk mengetahui pengaruh penambahan proses eliminasi *stopwords* terhadap hasil ekstraksi. Berdasarkan pengujian didapatkan minimal prosentase kandungan informasi terbaik pada sebuah kata yaitu minimal 90% dari total keseluruhan entropi dokumen set dan sekitar 50% sampai 90% dari entropi kategori di mana kata tersebut berasal. Semakin rendah nilai minimal prosentasi kandungan informasi pada sebuah kata, semakin besar pengaruh proses eliminasi *stopwords* terhadap peningkatan nilai *precision* hasil ekstraksi. Sebaliknya, semakin tinggi nilai minimal prosentasi kandungan informasi pada sebuah kata, semakin kecil pengaruh proses eliminasi *stopwords* terhadap peningkatan nilai *precision* hasil ekstraksi. Namun penambahan proses eliminasi *stopwords* juga dapat menurunkan nilai *recall* apabila ada di antara 10 *term* dengan nilai *tf-idf* tertinggi yang terdapat pada daftar *stopwords* yang akhirnya menurunkan nilai *f-measure*.

Kata Kunci: *Artificial Immune System* (AIS), ekstraksi kata kunci, kata kunci, Teori Informasi.