



### **Hak cipta dan penggunaan kembali:**

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

### **Copyright and reuse:**

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

## BAB V

### SIMPULAN DAN SARAN

#### 5.1 Simpulan

Simpulan yang didapat dalam penelitian ini adalah algoritma Rabin-Karp telah berhasil diimplementasikan dengan aplikasi berbasis *desktop* menggunakan bahasa pemrograman Java untuk mendeteksi kesamaan dokumen teks elektronik berbahasa Indonesia dengan bantuan dari proses *preprocessing* dokumen – dokumen yang akan melalui proses *string matching*. Tahapan *preprocessing* dari dokumen meliputi penghilangan karakter unik, proses *lowercase*, proses *tokenization*, dan proses yang terakhir adalah *preprocessing stopwords removal*. Uji coba program dilakukan pada 150 dokumen uji coba. Terdapat 10 dokumen orisinal yang bukan merupakan modifikasi dari dokumen utamanya. Dari 10 dokumen orisinal, program dapat mendeteksi 8 dokumen yang memiliki persentase di bawah 25%, dua dokumen sisanya memiliki persentase yang lebih besar dari 25% yaitu dokumen 100% perbedaan.docx dari dokumen utama original4.docx sebesar 25.08% dan dokumen 100% perbedaan.docx dari dokumen utama original6.docx sebesar 25.64%. Evaluasi terhadap program dilakukan dengan cara *Confusion Matrix*, menghitung nilai – nilai dari *precision*, *recall*, dan *accuracy* dari program sebagai ukuran dari kinerja program. Evaluasi memperoleh hasil *precision* 100% memiliki arti data yang diperoleh tidak ada data diluar dari data yang dicari, nilai *recall* 80% memiliki arti data yang didapat sebesar 80% dari keseluruhan data yang dicari dan nilai *accuracy* 99% yaitu akurasi dari program dalam memperoleh data.

## 5.2 Saran

Berikut merupakan saran dari penulis terhadap penelitian untuk pengembangan atau penelitian aplikasi menggunakan algoritma yang sama atau serupa.

1. Menambahkan algoritma Knuth-Morris-Pratt di dalam melakukan pencocokan string untuk membantu algoritma Rabin-Karp dalam melakukan pencocokan string pada baris dalam *array* 2 dimensi.
2. Penambahan kata dalam proses *stopwords removal* secara spesifik sesuai dengan dokumen yang dibandingkan akan semakin membantu dalam meningkatkan akurasi serta mengurangi ukuran dari data yang akan dilakukan *string matching*. Sebagai contoh, program akan mendeteksi dua dokumen yang membahas mengenai arti dari seni, maka dalam proses *stopwords removal* dapat menambahkan kata “seni” yang kira – kira akan sering disebutkan dalam dokumen, sehingga program otomatis akan menghilangkan kata tersebut dari pendeteksian.

