

## INTISARI

### PEINGKASAN TEKS BERBAHASA INDONESIA PADA MULTI DOKUMEN MENGGUNAKAN ALGORITMA GENETIKA

Oleh

Nirwana Hendrastuty

18/433790/PPA/05605

Peringkasan teks secara otomatis merupakan teknik pembuatan ringkasan dari sebuah teks secara otomatis yang mempertahankan poin penting dari dokumen aslinya dengan memanfaatkan aplikasi yang dijalankan pada komputer. Peringkasan teks secara otomatis memiliki dua pendekatan yaitu metode ekstraksi dan metode abstraksi. Metode ekstraksi memilih *subset* dari kalimat untuk dijadikan ringkasan dan metode abstraksi menciptakan sinopsis dari kalimat. Dokumen dibedakan menjadi dua jenis yaitu *single* dokumen dan multi dokumen. Multi dokumen merupakan masukan yang berasal dari banyak dokumen dari satu sumber atau lebih yang memiliki lebih dari satu ide pokok.

Penelitian ini bertujuan melakukan peringkasan teks menggunakan Algoritma Genetika pada multi dokumen menggunakan metode ekstraksi dengan memperhatikan ekstraksi fitur teks di setiap kromosom. Ekstraksi fitur yang digunakan yaitu posisi kalimat, kata kunci positif, kata kunci negatif, kemiripan antar kalimat, kalimat yang mengandung kata entitas, kalimat yang mengandung angka, panjang kalimat, koneksi antar kalimat, jumlah koneksi antar kalimat. Jumlah kromosom yang digunakan adalah setengah dari jumlah aduan masyarakat.

Data yang digunakan yaitu data pengaduan masyarakat terhadap pemerintah DIY dari bulan Februari 2018 sampai Juli 2020. Eksperimen yang dilakukan menggunakan sampling data setiap minggu sebanyak sembilan kali. Dari hasil pengujian di dapatkan rata-rata nilai *Precision* 1, *Recall* 0.71, dan nilai *f-measure* sebesar 0.79.

**Kata Kunci:** Peringkasan Teks Otomatis, Ekstraksi Fitur, Algoritma Genetika, Pemerintah DIY.

## ABSTRACT

### ***AUTOMATIC TEXT SUMMARIZATION IN MULTY-DOCUMENT USING GENETIC ALGORITHM***

By

Nirwana Hendrastuty

18/433790/PPA/05605

*Automatic text summarization is a technique of making a summary of a text automatically that retains the important points of the original document by utilizing an application that is run on a computer. Text summarization automatically has two approaches, namely the extraction method and the abstraction method. The extraction method selects a subset of sentences to be summarized and the abstraction method creates a synopsis of the sentences. Documents are divided into two types, namely single documents and multi documents. Multi-document is input that comes from many documents from one or more sources that have more than one main idea.*

*This study aims to summarize text using Genetic Algorithm on multiple documents using the extraction method by paying attention to the extraction of text features in each chromosome. The feature extraction used is sentence position, positive keywords, negative keywords, similarity between sentences, sentences containing entity words, sentences containing numbers, sentence length, connections between sentences, number of connections between sentences. The number of chromosomes used is half of the number of public complaints .*

*The data used is data on public complaints against the DIY government from February 2018 to July 2020. The experiment was carried out using data sampling nine times every week. From the test results, the average value of Precision 1, Recall is 0.71, and the f-measure value is 0.79.*

***Keyword: Automatic Text Sumarization, Feature Extraction , DIY government, Term Frecuency, Genetic Algorithm.***