

ABSTRAK

PERINGKASAN TEKS OTOMATIS PADA ARTIKEL BERITA ONLINE BERBAHASA INDONESIA MENGUNAKAN TEXTTEASER TERMODIFIKASI

oleh

Rokhana Diyah Rusdiati
15/378074/PA/16549

Informasi yang tersedia di Internet semakin bertambah banyak setiap harinya, namun waktu yang dimiliki oleh manusia terbatas untuk bisa mengetahui semua informasi tersebut. Salah satu media untuk mendapatkan informasi adalah melalui berita yang dapat diakses dari situs penyedia berita online. Banyaknya informasi tersebut tidak lantas memudahkan manusia dalam mengetahui apa sedang terjadi baru-baru ini. Membuat ringkasan secara manual bukanlah suatu pekerjaan yang mudah. Oleh sebab itu diperlukan peringkasan teks otomatis yang dapat memberikan ringkasan yang merepresentasikan keseluruhan isi berita yang ada.

Penelitian ini mengimplementasikan TextTeaser, yaitu peringkasan teks otomatis yang dibuat oleh Balbin untuk meringkas artikel dalam bahasa Inggris. TextTeaser tidak menyertakan proses *stemming* dan proses ekstraksi kata kunci memberikan skor yang kecil. Penelitian ini melakukan modifikasi pada TextTeaser, yaitu dengan menambahkan proses *stemming* pada tahap prapemrosesan dan penggunaan algoritma Rapid Automatic Keyword Extraction (RAKE) pada tahap ekstraksi kata kunci. Peringkasan teks otomatis dalam penelitian ini terdiri dari beberapa proses, yaitu prapemrosesan, penghitungan skor kemiripan kalimat dengan judul berita, penghitungan skor panjang kalimat, penghitungan skor posisi kalimat, penghitungan skor kata kunci pada setiap kalimat, dan pemilihan kalimat sebagai hasil akhir ringkasan.

Pengujian dilakukan dengan menghitung nilai *precision*, *recall*, dan *f-measure*, dari ringkasan yang dihasilkan oleh sistem terhadap ringkasan ideal yang dihasilkan oleh manusia menggunakan metode ROUGE-1. Berdasarkan pengujian yang telah dilakukan, diketahui bahwa TextTeaser yang telah dimodifikasi (pada bagian prapemrosesan dengan menambahkan proses *stemming* dan penggunaan algoritma RAKE pada tahap ekstraksi kata kunci) memberikan hasil ringkasan yang lebih mirip dengan ringkasan ideal yang dihasilkan oleh manusia, yaitu dengan nilai *f-measure* sebesar 60,75%.

Kata kunci: Peringkasan Teks Otomatis, TextTeaser, Stemming, RAKE, ROUGE

ABSTRACT

ATOMATIC TEXT SUMMARIZATION FOR ONLINE NEWS ARTICLE IN BAHASA INDONESIA USING MODIFIED TEXTTEASER

by

Rokhana Diyah Rusdiati
15/378074/PA/16549

The information available on the Internet is increasing a lot every day, but human's time is limited to be able to know all the information. One of media that can be used to obtain information is through news that can be accessed from various online news provider sites. The amount of information does not necessarily make it easier for human to know what is happening recently. Therefore, we need automatic text summarizer which can provide summary that represents the entire content of the news.

This research implements TextTeaser, an automatic text summarizer made by Balbin to summarize articles in English. TextTeaser does not include stemming process and the keyword extraction process gives a small score. This research modified TextTeaser by adding stemming process at preprocessing stage, and using Rapid Automatic Keyword Extraction (RAKE) algorithm at keyword extraction stage. The automatic text summarization in this research consists of several processes, namely preprocessing, calculating sentence similarity scores with news title, calculating sentence length scores, calculating sentence position scores, and selecting sentences as summary.

Testing is done by calculating the value of precision, recall, and f-measure, from the summary generated by the system to the ideal summary produced by humans using ROUGE-1 method. Based on the test results, it is known that the modified TextTeaser (in preprocessing stage by adding stemming process and the use of RAKE algorithm in keyword extraction stage) provides summary result that is more similar to the ideal summary produced by humans, with the f-measure value of 60.75%.

Keywords: Automatic Text Summarization, TextTeaser, Stemming, RAKE, ROUGE