

## ABSTRAK

Saifuddin

(150411100075)

Dosen Pembimbing I : Yoga Dwitya Pramudita, S.Kom.M.Cs.

Dosen Pembimbing II : Ika Oktavia Suzanti S.Kom.M.Cs.

Artikel merupakan sebuah karya tulis yang berisi informasi dari suatu fakta maupun opini yang dipublikasi dan bertujuan untuk menarik minat pembaca sehingga membuat orang-orang datang untuk mencarinya. Pariwisata merupakan salah satu topik artikel yang banyak diminati orang-orang. Membaca artikel membutuhkan waktu yang banyak, apalagi dokumen artikel yang dibaca banyak. Untuk mempersingkat waktu itu dalam membaca artikel dibutuhkanlah sistem peringkasan teks otomatis. Ringkasan teks tersusun dari kalimat-kalimat penting dari dokumen yang panjangnya kurang atau lebih dari setengahnya. Dalam penelitian ini mengusulkan implementasi *TF-IDF* dan *K-Medoid* dalam metode *Ekstraktif* peringkasan teks otomatis dokumen bahasa Indonesia yang diharapkan memberikan kecepatan dan akurasi yang baik dalam peringkasan teks otomatis. Algoritma *TF-IDF* dikenal mampu untuk menghasilkan ringkasan teks berdasarkan bobot yang didapat pada setiap kalimat dalam teks. Metode *K-Medoid* merupakan metode *clustering* yang berfungsi untuk mengelompokkan kalimat-kalimat yang dianggap sistem memiliki makna yang sama. Pada penelitian ini menggunakan data yang berjumlah 33 teks dokumen yang diambil dari sumber web berbeda-beda di internet. Untuk pengujian hasil ringkasan dari penelitian ini dilakukan oleh 3 responden ahli yang akan mengoreksi secara manual hasil ringkasan dari sistem. Dari hasil pengujian yang telah dilakukan mendapatkan nilai akurasi sebesar 65.40 %.

**Kata Kunci:** Artikel, pariwisata, online, Peringkasan Teks Otomatis, Metode *Ekstraktif*, *TF-IDF*, *K-Medoid Clustering*.

## **ABSTRACT**

**Saifuddin**

**(150411100075)**

**Advisor I : Yoga Dwitya Pramudita, S.Kom.M.Cs.**

**Advisor II : Ika Oktavia Suzanti S.Kom.M.Cs.**

*An article is a written work that contains information from a published fact or opinion and aims to attract the interest of the reader so that people come to look for it. Tourism is one of the most popular article topics. Reading articles takes a lot of time, especially if you read a lot of article documents. To shorten that time in reading articles, an automatic text summarization system is needed. Text summaries are composed of important sentences from documents that are less or more than half in length. This study proposes the implementation of TF-IDF and K-Medoid in the Extractive method of automatic text summarization of Indonesian documents which are expected to provide good speed and accuracy in automatic text summarization. The TF-IDF algorithm is known to be able to generate text summaries based on the weights obtained for each sentence in the text. The K-Medoid method is a clustering method that functions to group sentences that the system considers to have the same meaning. In this study using data totaling 33 document texts taken from different web sources on the internet. For testing the summary results of this study, 3 expert respondents will manually correct the summary results from the system. From the results of the tests that have been carried out, the accuracy value is 65.40%.*

**Keywords:** Articles, tourism, online, Automatic Text Summarization, Extractive Methods, TF-IDF, K-Medoid Clustering.

## DAFTAR ISI

|                                              |             |
|----------------------------------------------|-------------|
| <b>HALAMAN JUDUL .....</b>                   | <b>i</b>    |
| <b>HALAMAN PENGESAHAN.....</b>               | <b>ii</b>   |
| <b>HALAMAN PERNYATAAN ORISINALITAS .....</b> | <b>iii</b>  |
| <b>HALAMAN PERSEMBAHAN.....</b>              | <b>iv</b>   |
| <b>KATA PENGANTAR.....</b>                   | <b>v</b>    |
| <b>ABSTRAK .....</b>                         | <b>vi</b>   |
| <b>ABSTRACT .....</b>                        | <b>vii</b>  |
| <b>DAFTAR ISI.....</b>                       | <b>viii</b> |
| <b>DAFTAR GAMBAR.....</b>                    | <b>xi</b>   |
| <b>DAFTAR TABEL .....</b>                    | <b>xii</b>  |
| <b>BAB I.....</b>                            | <b>1</b>    |
| <b>PENDAHULUAN.....</b>                      | <b>1</b>    |
| 1.1 Latar Belakang.....                      | 1           |
| 1.2 Rumusan Masalah.....                     | 2           |
| 1.2.1 Permasalahan .....                     | 2           |
| 1.2.2 Usulan Metode.....                     | 2           |
| 1.2.3 Pertanyaan Penelitian .....            | 2           |
| 1.3 Tujuan & Manfaat.....                    | 3           |
| 1.3.1 Tujuan Penelitian.....                 | 3           |
| 1.3.2 Manfaat.....                           | 3           |
| 1.4 Batasan – Batasan .....                  | 3           |
| 1.5 Sistematika Penulisan .....              | 3           |
| <b>BAB II .....</b>                          | <b>5</b>    |
| <b>KAJIAN PUSTAKA .....</b>                  | <b>5</b>    |
| 2.1 Peringkasan Teks Otomatis .....          | 5           |
| 2.1.1 Teks Preprocessing .....               | 5           |
| 2.1.2 TF-IDF.....                            | 6           |
| 2.1.3 Elbow.....                             | 7           |
| 2.1.4 K-Medoid Clustering.....               | 7           |

|                                                                                              |           |
|----------------------------------------------------------------------------------------------|-----------|
| 2.1.5 Library Python.....                                                                    | 8         |
| 2.2 Penelitian Terkait.....                                                                  | 9         |
| <b>BAB III.....</b>                                                                          | <b>12</b> |
| <b>METODE PENELITIAN.....</b>                                                                | <b>12</b> |
| 3.1 Tahapan Penelitian.....                                                                  | 12        |
| 3.1.1 Studi Pustaka .....                                                                    | 12        |
| 3.1.2 Perancangan Sistem.....                                                                | 12        |
| 3.1.3 Implementasi Sistem .....                                                              | 12        |
| 3.1.4 Uji Coba Sistem.....                                                                   | 13        |
| 3.1.5 Dokumentasi.....                                                                       | 13        |
| 3.2 Arsitektur Sistem .....                                                                  | 13        |
| 3.2.1 Flowchart K-Medoid .....                                                               | 14        |
| 3.3 Data Set.....                                                                            | 15        |
| 3.4 Perhitungan Manual Peringkasan Teks Otomatis Dengan TF-IDF dan K-Medoid clustering ..... | 16        |
| 3.4.1 Perhitungan TF-IDF .....                                                               | 16        |
| 3.4.2 Perhitungan Elbow .....                                                                | 20        |
| 3.4.3 Perhitungan K-Medoid .....                                                             | 22        |
| 3.5 Skenario Pengujian .....                                                                 | 26        |
| <b>BAB IV.....</b>                                                                           | <b>27</b> |
| <b>HASIL DAN PEMBAHASAN .....</b>                                                            | <b>27</b> |
| 4.1 Lingkungan Uji Coba .....                                                                | 27        |
| 4.2 Data Set.....                                                                            | 27        |
| 4.3 Tahapan Proses .....                                                                     | 27        |
| 4.3.1 Pengambilan Data.....                                                                  | 27        |
| 4.3.2 Case Folding.....                                                                      | 28        |
| 4.3.3 Tokenizing.....                                                                        | 29        |
| 4.3.4 Filtering .....                                                                        | 29        |
| 4.3.5 Stemming.....                                                                          | 30        |
| 4.3.6 Stopword Removal .....                                                                 | 30        |
| 4.3.7 TF-IDF.....                                                                            | 30        |

|                                             |           |
|---------------------------------------------|-----------|
| 4.3.8 K-Medoid Clustering .....             | 32        |
| 4.3.9 Elbow .....                           | 34        |
| 4.4 Tampilan Sistem .....                   | 35        |
| 4.4.1 Tampilan Halaman Awal .....           | 35        |
| 4.4.2 Tampilan Hasil Ringkasan Sistem ..... | 36        |
| 4.5 Pengujian Sistem .....                  | 36        |
| <b>BAB V .....</b>                          | <b>40</b> |
| <b>KESIMPULAN .....</b>                     | <b>40</b> |
| 5.1 Kesimpulan .....                        | 40        |
| 5.2 Saran .....                             | 40        |
| <b>DAFTAR PUSTAKA .....</b>                 | <b>41</b> |
| <b>LAMPIRAN .....</b>                       | <b>43</b> |

## DAFTAR GAMBAR

|                                                             |    |
|-------------------------------------------------------------|----|
| Gambar 3.1 Arsitektur Sistem.....                           | 13 |
| Gambar 3.2 Flowchart <i>K-Medoid</i> .....                  | 15 |
| Gambar 3.3 Grafik Perhitungan Elbow Dokumen Ke-1 .....      | 22 |
| Gambar 4.1 Halaman Awal.....                                | 35 |
| Gambar 4.2 Halaman Hasil Ringkasan .....                    | 36 |
| Gambar 4.3 Grafik Pengurangan Kata dan rata-rata nilai..... | 38 |

## DAFTAR TABEL

|                                                             |    |
|-------------------------------------------------------------|----|
| Tabel 3.1 Contoh Dokumen Artikel.....                       | 16 |
| Tabel 3.2 Perhitungan $Tf$ , $Df$ , Dan $Idf$ .....         | 17 |
| Tabel 3.3 Contoh Perhitungan $TF-IDF(W)$ .....              | 18 |
| Tabel 3.4 Hasil $Tf-Idf(W)$ Semua Kalimat .....             | 20 |
| Tabel 3.5 Perhitungan Awal $K-medoid$ .....                 | 23 |
| Tabel 3.6 Perhitungan $K-Medoid$ Iterasi Ke-1 .....         | 23 |
| Tabel 3.7 Perhitungan $K-Medoid$ Iterasi Ke-2 .....         | 24 |
| Tabel 3.8 Total Bobot Kalimat Tiap Kluster .....            | 25 |
| Tabel 3.9 Hasil Ringkasan .....                             | 25 |
| Tabel 3.10 Hasil Perhitungan Elbow Pada Setiap Dokumen..... | 21 |
| Tabel 3.11 Contoh Pengujian Sistem.....                     | 26 |
| Tabel 4. 1 Spesifikasi perangkat .....                      | 27 |
| Tabel 4. 2 Tabel Pengujian Sistem.....                      | 37 |