

**PERINGKASAN TEKS BAHASA INDONESIA MENGGUNAKAN  
*MODIFIED DISCRETE DIFFERENTIAL EVOLUTION ALGORITHM***

**Jurnal**

Diajukan kepada Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Negeri Yogyakarta

untuk Memenuhi Sebagian Persyaratan

guna Memperoleh Gelar Sarjana Sains



Oleh:

Heny Setyawan

12305141034

**PROGRAM STUDI MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS NEGERI YOGYAKARTA**

**2016**

## PERSETUJUAN

Skripsi yang berjudul “**PERINGKASAN TEKS BAHASA INDONESIA MENGGUNAKAN *MODIFIED DISCRETE DIFFERENTIAL EVOLUTION ALGORITM***” yang disusun oleh Heny Setyawan, NIM. 12305141034 ini telah disetujui oleh dosen pembimbing dan direview oleh dosen penguji untuk memenuhi sebagian persyaratan guna memperoleh Gelar Sarjana Sains.



Yogyakarta, 13 Juni 2016

Direview  
Dosen Penguji

**Kuswari H., S.Si., M.Kom.**  
NIP. 19760414 200501 2 002

Disetujui  
Dosen Pembimbing

**Nur Hadi W., M.Eng.**  
NIP. 19780119 200312 1 002

# PERINGKASAN TEKS BAHASA INDONESIA MENGGUNAKAN *MODIFIED DISCRETE DIFFERENTIAL EVOLUTION ALGORITHM*

## *INDONESIAN TEXT SUMMARIZATION USE MODIFIED DISCRETE DIFFERENTIAL EVOLUTION ALGORITHM*

Oleh: Heny Setyawan<sup>1)</sup>, Nur Hadi Waryanto<sup>2)</sup>

Program Studi Matematika, Jurusan Pendidikan Matematika, FMIPA UNY

<sup>1)</sup>henysetyawan@gmail.com, <sup>2)</sup>nurhadiw@gmail.com

### Abstrak

Perkembangan kehidupan modern yang pesat membuat informasi yang beredar semakin banyak. Sebagian besar informasi yang beredar berbentuk teks. Orang perlu membaca untuk memperoleh informasi. Akan tetapi, membaca teks panjang membutuhkan waktu yang lama. Sistem peringkasan teks merupakan solusi karena dapat menghasilkan ringkasan dari suatu teks. Salah satu metode yang dapat digunakan pada peringkasan teks adalah *Modified Discrete Differential Evolution* (MDDE). Perancangan sistem menggunakan metode *Waterfall* dengan tahapan analisis, perancangan, implementasi, dan pengujian. Sistem dibuat menggunakan bahasa pemrograman PHP dan didesain secara *responsive*. Pengujian yang dilakukan pada sistem adalah evaluasi ringkasan dan pengujian *Betha*. Berdasarkan evaluasi ringkasan, sistem mempunyai tingkat akurasi 51.33% pada teks bahasa Indonesia. Akurasi sistem berdasarkan jenis teks adalah: narasi=18.89%, deskripsi=48.22%, persuasi=31.55%, eksposisi=55.56%, dan argumentasi=66.66%. Berdasarkan pengujian *Betha*, sistem membantu memahami isi teks, menghasilkan ringkasan yang relevan, aman, dan mudah digunakan.

Kata kunci: sistem, sistem peringkasan teks, MDDE

### Abstract

*The rapid development of modern life increases the information's circulation. Most of the information are in text form. People need to read the text in order to get the information. However, reading a long text is a time-consuming process. Automatic text summarization is one of the solution to this problem because it can produce summarization. Modified Discrete Differential Evolution (MDDE) is a method that can be used in text summarization. The development of text summarization system is based on Waterfall's model. Waterfall's model consist of analysis, design, implementation, and test. The system is developed using PHP language and designed to be responsive. The text summarization system is tested with two different method, which are summarization evaluation and betha testing. The result of summarization evaluation shows that the system has accuracy of 51.33% in Indonesian text. Text summarization accuracy based on text type are as follows: narration=18.89%, description=48.22%, persuasion=31.55%, exposition=55.56%, and argumentation=66.66%. Based on beta testing result, it can be concluded that text summarization system help to understand the text contents, produce relevant summarization, safe, and easy to use.*

Keywords: system, text summarization system, MDDE

### PENDAHULUAN

Perkembangan kehidupan modern yang pesat membuat informasi yang beredar juga semakin banyak. Hal ini didukung dengan perkembangan teknologi yang memudahkan orang untuk mendapatkan informasi. Salah satu teknologi yang membantu menyebarkan informasi adalah internet. Orang dapat mencari dan mendapatkan informasi yang diinginkan menggunakan internet. Bahkan hampir tidak ada

batasan informasi yang dapat diperoleh dari internet.

Informasi yang tersebar sebagian besar berupa teks, sehingga orang harus membacanya untuk mengetahui informasi tersebut. Akan tetapi orang tidak selalu mempunyai banyak waktu untuk membaca. Hal ini biasanya disebabkan oleh kesibukan orang yang tidak memungkinkan untuk membaca teks yang cukup panjang. Oleh sebab itu, dibutuhkan alat untuk meringkas teks secara

otomatis untuk menyiasati masalah tersebut. Dengan membaca ringkasan teks tersebut maka akan menghemat waktu memperoleh informasi.

Peringkasan teks adalah proses mengambil kalimat penting dari suatu teks yang menggambarkan inti teks tersebut. Menurut Niladri (2007:448), peringkasan teks otomatis merupakan penelitian yang penting dan menarik pada bidang *natural language processing*. Menurut Alguliev & Aliguliyev (2009:128), proses peringkasan teks dapat dibagi menjadi tiga tahap, yakni analisis, transformasi, dan sintesis.

*Modified Discrete Differential Evolution* (MDDE) merupakan pengembangan metode *Differential Evolution*. Metode ini dapat meningkatkan akurasi ringkasan sampai 10% (Alguliev & Aliguliyev, 2009:136). Selain itu, metode MDDE menggunakan bilangan bulat pada setiap elemennya. Hal ini dapat memudahkan proses klasterisasi kalimat sebagai dasar peringkasan teks.

Berdasarkan hal-hal tersebut muncul gagasan untuk membuat sistem peringkasan teks bahasa Indonesia dengan menggunakan metode MDDE. Sistem ini diharapkan dapat memudahkan orang dalam mencari informasi. Selain itu, penggunaan sistem ini diharapkan dapat mempercepat waktu memperoleh informasi.

## HASIL PENELITIAN DAN PEMBAHASAN

Sistem adalah sekumpulan elemen-elemen yang berinteraksi untuk mencapai suatu tujuan tertentu (Jogiyanto, 2005:2). Sistem peringkasan teks merupakan suatu sistem dengan tujuan menampilkan ringkasan dari teks yang dimasukkan pengguna. Perancangan sistem peringkasan teks menggunakan model *Waterfall*. Ada empat tahap pada model *Waterfall* yaitu Analisis Sistem, Perancangan Sistem, Implementasi Sistem, dan Pengujian Sistem (Pressman, 2001:28). Berikut penjelasan masing-masing tahap:

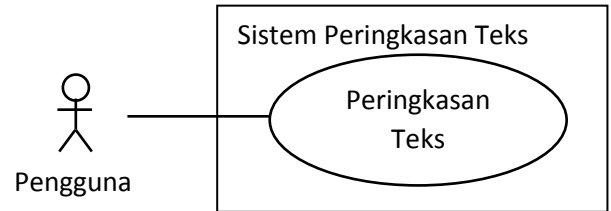
### Analisis Sistem

Ada satu prosedur utama pada sistem peringkasan teks, yakni prosedur peringkasan. Prosedur peringkasan merupakan prosedur saat pengguna memasukkan teks ke dalam sistem untuk diringkas. Sistem yang akan dibuat tidak

menggunakan pembagian dan pembatasan hak akses.

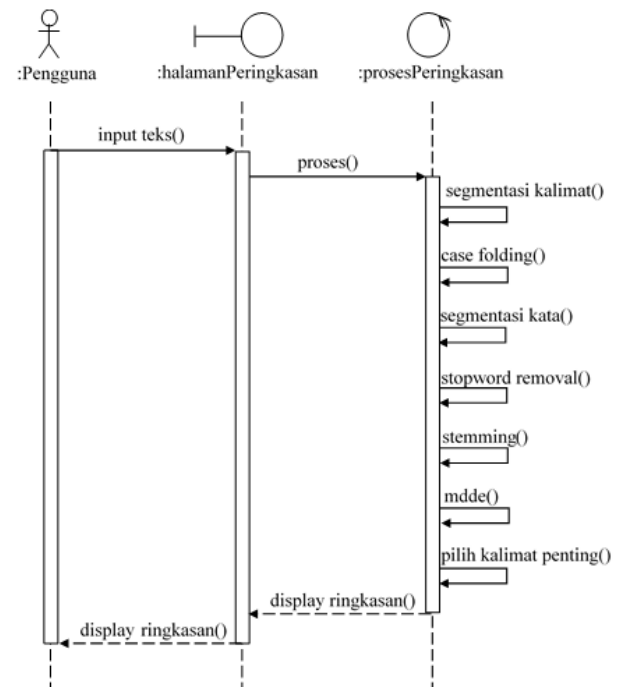
### Perancangan Sistem

Rancangan proses sistem peringkasan teks dapat ditunjukkan dengan diagram *use case* pada Gambar 1.



Gambar 1. Diagram *Use Case* Sistem Peringkasan Teks

Urutan proses sistem peringkasan teks dapat ditunjukkan dengan diagram *sequence* pada Gambar 2.



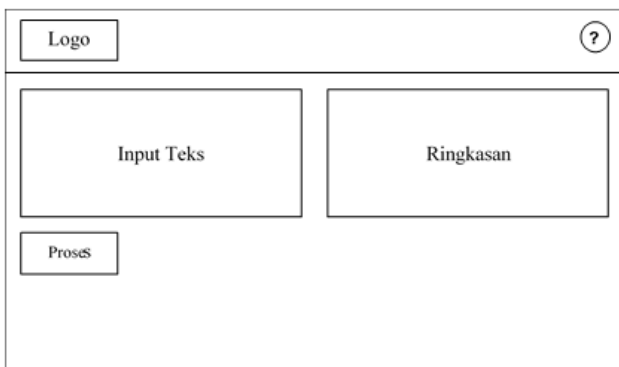
Gambar 2. Diagram *Sequence* Sistem Peringkasan Teks

Berikut penjelasan masing-masing proses:

1. Segmentasi kalimat: proses untuk memisahkan kalimat dari teks yang akan diringkas berdasarkan tanda baca.
2. *Case folding*: proses mengubah setiap huruf pada kalimat menjadi huruf kecil.
3. Segmentasi kata: proses memisahkan kata pada setiap kalimat.

4. *Stopword removal*: proses menghilangkan kata-kata yang kurang mempunyai makna dan sering muncul pada kalimat.
5. *Stemming*: proses mengembalikan kata turunan ke bentuk akar kata (kata dasar) menggunakan aturan tertentu.
6. MDDE: proses mengelompokkan kalimat berdasarkan kemiripan antar kalimat.
7. Pilih kalimat penting: proses pemilihan kalimat terpenting pada setiap kluster.

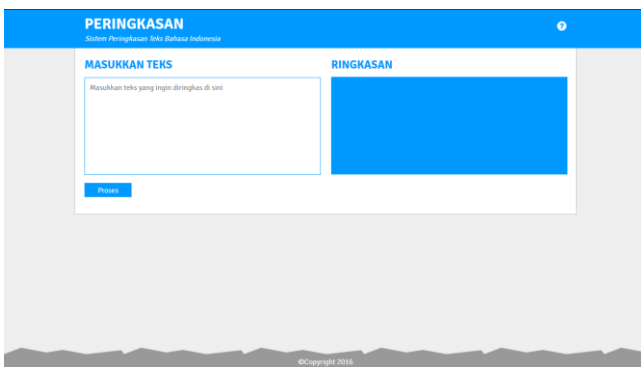
Antarmuka halaman peringkasan dibuat sederhana agar pengguna lebih fokus kepada hasil ringkasan. Rancangan antarmuka halaman peringkasan ditunjukkan pada Gambar 3.



Gambar 2. Rancangan Antarmuka Sistem Peringkasan Teks

### Implementasi Sistem

Sistem dibuat menggunakan bahasa pemrograman PHP dan menggunakan desain *responsive* sehingga tampilan akan menyesuaikan ukuran layar yang digunakan. Berdasarkan rancangan antarmuka pada Gambar 2, implementasi antarmuka sistem ditunjukkan pada Gambar 3.



Gambar 3. Implementasi Antarmuka Sistem Peringkasan Teks

### Pengujian Sistem

Ada dua jenis pengujian yang dilakukan pada sistem peringkasan teks, yakni evaluasi ringkasan dan pengujian *Betha*. Evaluasi ringkasan bertujuan untuk menguji tingkat akurasi sistem dari segi kebahasaan, sedangkan pengujian *betha* bertujuan untuk menguji kelayakan sistem dari segi pengguna. Pada evaluasi ringkasan, sistem diuji dengan membandingkan ringkasan sistem dengan ringkasan pakar sedangkan pada pengujian *betha*, sistem diujikan kepada pengguna dan ahli komputer.

Evaluasi ringkasan dilakukan dengan menghitung nilai *F-Score* dari 10 teks dengan jenis narasi, deskripsi, persuasi, eksposisi, dan argumentasi. Selain itu, ketiga *fitness* yang dapat digunakan pada MDDE juga diuji untuk menentukan *fitness* terbaik.

Menurut Alguliev & Aliguliyev (2009:131), tiga *fitness* yang dapat digunakan pada MDDE adalah:

$$F = (1 + \text{sigm}(F_1))^{F_2}$$

$$F_1 = \sum_{p=1}^k |C_p| \sum_{S_i, S_j \in C_p} \text{sim}_{NGD}(S_i, S_j)$$

$$F_2 = \sum_{p=1}^{k-1} \frac{1}{|C_p|} \sum_{q=p+1}^k \frac{1}{|C_q|} \sum_{S_i \in C_p} \sum_{S_j \in C_q} \text{sim}_{NGD}(S_i, S_j)$$

Dengan

$$\text{sigm}(z) = \frac{1}{1 + \exp(-z)}$$

Perhitungan nilai *F-Score* menggunakan persamaan yang dikemukakan oleh Steinberger & Ježek (2009) di bawah ini:

$$F\text{-Score} = \frac{2PR}{P+R}$$

Dengan

$$P = \frac{\text{correct}}{\text{correct} + \text{wrong}}$$

$$R = \frac{\text{correct}}{\text{correct} + \text{missed}}$$

*Correct* adalah kalimat yang ada pada ringkasan pakar dan ringkasan sistem. *Missed* adalah kalimat hasil ringkasan pakar yang tidak ada dalam ringkasan sistem. *Wrong* adalah kalimat hasil ringkasan sistem yang tidak termasuk dalam ringkasan pakar.

Hasil *F-Score* masing-masing teks berdasarkan jenis *fitness* ditunjukkan pada Tabel 1.

Tabel 1. Nilai *F-Score* Setiap Teks Sesuai Jenis *Fitness*

Teks	<i>Fitness</i>	<i>fitness1</i>	<i>fitness2</i>
1	0	0	0.3333
2	0.3	0.3	0.2
3	0.75	0.5	0.5
4	0.4286	0.4286	0.2857
5	0.25	0.25	0.25
6	0.5714	0.1429	0.4286
7	1	0.5	0.5
8	0.3333	0.5	0.5
9	1	1	0.5
10	0.5	0.333	0.6666
<b>Akurasi</b>	<b>0.5133</b>	<b>0.3955</b>	<b>0.4164</b>

Akurasi sistem peringkasan teks berdasarkan jenis teks ditampilkan pada Tabel 2.

Tabel 2. Akurasi Sistem Peringkasan Teks Berdasarkan Jenis Teks

Jenis Teks	Akurasi
Narasi	0.1889
Deskripsi	0.4822
Persuasi	0.3155
Eksposisi	0.5556
Argumentasi	0.6666

Pengujian *betha* dibagi menjadi dua, yakni pengujian kepada pengguna dan ahli komputer. Pengujian kepada pengguna ditujukan kepada 20 mahasiswa yang telah menggunakan sistem. Pengujian kepada ahli komputer ditujukan kepada dosen untuk *me-review* sistem secara keseluruhan.

Pengujian sistem menggunakan kuesioner yang terdiri dari 7 butir pertanyaan untuk pengguna dan 11 butir pertanyaan untuk ahli komputer. Kuesioner menggunakan skala jawaban Sangat Tidak Setuju (STS), Tidak Setuju (TS), Setuju (S), dan Sangat Setuju (SS). Dari hasil pengujian kuesioner tersebut, dilakukan perhitungan untuk menentukan kesimpulan terhadap penilaian aplikasi yang dibangun. Hasil

pengujian sistem kepada pengguna ditunjukkan pada Tabel 3.

Tabel 3. Hasil Pengujian *Betha*

Faktor	Hasil	Keterangan
<i>Correctness</i>	3.025	Baik
<i>Reliability</i>	2.85	Cukup Baik
<i>Integrity</i>	3.25	Baik
<i>Usability</i>	3.2167	Baik

Berdasarkan Tabel 3, sistem peringkasan membantu memahami isi teks, menghasilkan ringkasan yang relevan, aman, dan mudah digunakan.

## SIMPULAN DAN SARAN

### Simpulan

Sistem peringkasan dibangun menggunakan bahasa pemrograman PHP dan dirancang menggunakan model *Waterfall*. Ada empat tahap perancangan, yaitu analisis sistem, perancangan sistem, implementasi sistem dan pengujian sistem. Berikut penjelasan masing-masing tahap:

- Analisis sistem : Ada satu prosedur pada sistem peringkasan, yakni prosedur peringkasan. Sistem bersifat umum, sehingga hanya ada satu tipe pengguna.
- Perancangan sistem : Antarmuka sistem dirancang sederhana agar pengguna lebih fokus kepada hasil peringkasan.
- Implementasi sistem : Sistem dibuat dengan menggunakan bahasa pemrograman PHP dan didesain secara *responsive*.
- Pengujian sistem : Berdasarkan hasil pengujian *betha*, *correctness* dalam kriteria baik, *reliability* dalam kriteria cukup baik, *integrity* dalam kriteria baik, dan *usability* dalam kriteria baik. Artinya, sistem peringkasan membantu memahami isi teks, menghasilkan ringkasan yang relevan, aman, dan mudah digunakan.

Berdasarkan pengujian evaluasi ringkasan, sistem peringkasan teks mempunyai akurasi 51.33%. Akurasi sistem berdasarkan jenis teks, adalah sebagai berikut: narasi = 18.89%,

deskripsi = 48.22%, persuasi = 31.55%, eksposisi = 55.56%, dan argumentasi = 66.66%. Artinya, sistem paling baik digunakan pada teks berjenis argumentasi dan tidak baik pada teks berjenis narasi.

### **Saran**

Beberapa saran untuk penelitian selanjutnya antara lain: pengujian dilakukan pada tipe teks lainnya, seperti jurnal dan skripsi, aturan segmentasi kalimat diperbaiki agar pembagian kalimat lebih akurat, penambahan tipe pengguna admin yang dapat mengganti daftar kamus, daftar *stopword*, dan daftar singkatan menggunakan sistem, dan sistem diperbaiki sehingga dapat melakukan peringkasan teks dari suatu file.

### **DAFTAR PUSTAKA**

Jogiyanto H., (2005). *Analisis dan Desain Sistem Informasi*. Yogyakarta: Andi.

Josef Steinberger & Karel Jeřek. (2009). EVALUATION MEASURES FOR TEXT SUMMARIZATION. *Computing and Informatics* (Vol. 28). Hlm. 1001-1026.

Niladri Chatterjee & Shiwali Mohan. (2007). Extraction-Based Single-Document Summarization Using Random Indexing. *IEEE International Conference on Tools with Artificial Intelligence* (Vol.19, 2007). Hlm 448-455.

Pressman, R. S. (2001). *Software Engineering A Practitioner Approach*. 5th. Ed. New York: McGraw Hill.

Rasim Alguliev & Ramiz Aliguliyev. (2009). Evolutionary Algorithm for Extractive Text Summarization. *Intelligent Information Management* (Nomor 1 tahun 2009). Hlm. 128-138.