

**PERINGKASAN DOKUMEN BAHASA INDONESIA  
PADA CERPEN MENGGUNAKAN METODE  
*MAXIMUM MARGINAL RELEVANCE***

*Diajukan Sebagai Syarat Untuk  
Menyelesaikan Pendidikan Program Strata-1  
Pada Jurusan Teknik Informatika Fakultas Ilmu Komputer Unsri*



**Oleh :**

**EVA SAPAYONA**

**09021381320042**

**JURUSAN TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA  
2019**

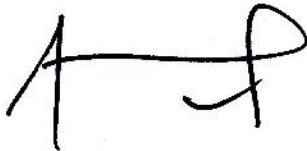
# LEMBAR PENGESAHAN TUGAS AKHIR

PERINGKASAN DOKUMEN BAHASA INDONESIA  
PADA CERPEN MENGGUNAKAN METODE  
*MAXIMUM MARGINAL RELEVANCE*

Oleh :

EVA SAPAYONA  
NIM. 09021381320042

Pembimbing I



M. Fachrurrozi, S.Si., M.T.  
NIP. 198005222008121002

Palembang, Agustus 2019

Pembimbing II



Novi Yuslimi, M.T.  
NIP. 198211082012122001

Menyetujui,

Ketua Jurusan Teknik Informatika



Rifkie Primartha, M.T.  
NIP. 197706012009121004

## TANDA LULUS UJIAN SIDANG TUGAS AKHIR

Pada hari selasa tanggal 23 Juli 2019 telah dilaksanakan Ujian Sidang Tugas Akhir oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

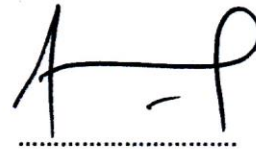
Nama : Eva Sapayona

NIM : 09021381320042

Judul : Peringkasan Dokumen Bahasa Indonesia Pada Cerpen  
Menggunakan Metode *Maximum Marginal Relevance*

### 1. Pembimbing I

M. Fachrurrozi, S.Si., MT  
NIP198005222008121002



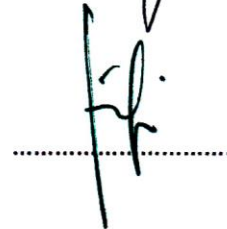
### 2. Pembimbing II

Novi Yusliani, M.T  
NIP 198211082012122001



### 3. Ketua Penguji

Rifkie Primartha, S.T. M.T  
NIP197706012009121004



### 4. Anggota Penguji

Muhammad Ali Buchari, M.T  
NIP198803302019031007



Mengetahui,  
Ketua Jurusan Teknik Informatika



Rifkie Primartha, M.T

## HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Eva Sapayona  
NIM : 09021381320042  
Fakultas : Ilmu Komputer  
Program Studi : Teknik Informatika Bilingual  
Judul Skripsi : Peringkasan Dokumen Bahasa Indonesia Pada  
Cerpen Menggunakan Metode *Maximum Marginal  
Relevance*  
Hasil Pengecekan Software  
*iThenticate/Turnitin* : 19 %

Menyatakan bahwa Laporan Penelitian saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan proyek ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.

Palembang, Agustus 2019



Eva Sapayona  
NIM. 09021381320042

## KATA PENGANTAR



Puji syukur kepada Allah swt. atas berkat dan rahmat-Nya yang telah diberikan kepada Penulis sehingga dapat menyelesaikan Tugas Akhir ini dengan baik. Tugas akhir ini disusun untuk memenuhi salah satu syarat guna menyelesaikan pendidikan program Strata-1 pada Fakultas Ilmu Komputer Program Studi Teknik Informatika di Universitas Sriwijaya.

Dalam menyelesaikan Tugas Akhir ini banyak pihak yang telah memberikan bantuan dan dukungan baik secara langsung maupun secara tidak langsung. Untuk itu Penulis ingin menyampaikan rasa terima kasih kepada:

1. Orang tuaku, Robinson dan Siti Ani, saudaraku, Evi Syahfitri, dan Adean Saputra dan juga seluruh keluarga besarku yang selalu mendoakan serta memberikan dukungan baik moril maupun materil.
2. Bapak Jaidan Jauhari, S.Pd., M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Bapak Rifkie Primartha, S.T., M.T. selaku Ketua Jurusan Teknik Informatika
4. Bapak Muhammad Fachrurrozi, S.Si, M.T. selaku dosen pembimbing I dan Ibu Novi Yusliani, S.Kom., M.T. selaku dosen pembimbing II yang telah memberikan arahan serta dukungan dalam proses pengerjaan Tugas Akhir
5. Bapak Rifkie Primartha, S.T., M.T., Ibu Desty Rodiah, S.Kom., M.T., Bapak Kanda Januar, M.T., dan Bapak Muhammad Ali Buchari, M.T selaku dosen penguji yang telah memberikan masukan dan dorongan dalam proses pengerjaan Tugas Akhir

6. Ibu Yunita. M.Cs selaku dosen pembimbing akademik
7. Seluruh dosen Program Studi Teknik Informatika dan staf tata usaha Fakultas Ilmu Komputer Universitas Sriwijaya yang telah membantu dalam kelancaran proses administrasi dan akademik selama masa perkuliahan
8. Rachman Hakim, Sepriansah sahabat-sahabatku yang selalu ada buatku dalam suka maupun duka dari semester 1 hingga penulis menyelesaikan studi.
9. Rizky Herlangga dan Fuad Wahid yang telah memotivasi dan berjuang bersama Penulis selama masa perkuliahan
10. Keluargaku, seluruh anggota IF Bilingual 2013
11. Semua pihak yang tidak bisa Penulis sebutkan satu persatu yang telah membantu dan berperan dalam Tugas Akhir ini.

Penulis menyadari dalam penyusunan Tugas Akhir ini masih terdapat banyak kekurangan disebabkan keterbatasan pengetahuan dan pengalaman, oleh karena itu kritik dan saran yang membangun sangat diharapkan untuk kemajuan penelitian selanjutnya.

Akhir kata dengan segala kerendahan hati, semoga Tugas Akhir ini dapat berguna dan bermanfaat bagi kita semua.

Palembang, Agustus 2019

Eva Sapayona

## DAFTAR ISI

Halaman

<b>HALAMAN JUDUL</b> .....	i
<b>HALAMAN PENGESAHAN</b> .....	ii
<b>TANDA LULUS UJIAN SIDANG TUGAS AKHIR</b> .....	iii
<b>HALAMAN PERNYATAAN</b> .....	iv
<b>MOTTO DAN PERSEMBAHAN</b> .....	v
<b>ABSTRACT</b> .....	vi
<b>ABSTRAK</b> .....	vii
<b>KATA PENGANTAR</b> .....	viii
<b>DAFTAR ISI</b> .....	x
<b>DAFTAR TABEL</b> .....	xv
<b>DAFTAR GAMBAR</b> .....	xvii

### **BAB I PENDAHULUAN**

1.1 Pendahuluan .....	I-1
1.2 Latar Belakang .....	I-1
1.3 Rumusan Masalah .....	I-3
1.4 Tujuan Penelitian .....	I-3
1.5 Manfaat Penelitian .....	I-4
1.6 Batasan Masalah .....	I-4
1.7 Sistematika Penulisan .....	I-5
1.8 Kesimpulan .....	I-6

### **BAB II LANDASAN TEORI**

2.1 Pendahuluan .....	II-1
2.2 Penelitian Terkait .....	II-1
2.3 Pemrosesan Bahasa Alami .....	II-3
2.4 Peringkasan Teks Dokumen .....	II-4
2.5 <i>Pre-processing</i> .....	II-5
2.6 Algoritma <i>Stemming</i> Nazief dan Andriani .....	II-7

2.7 Evaluasi Penelitian .....	II-11
2.8 <i>Maximum Marginal Relevance</i> (MMR).....	II-11
2.9 Algoritma TF-IDF .....	II-13
2.10 <i>F-Measure</i> .....	II-15
2.11 Kesimpulan .....	II-16

### **BAB III METODE PENELITIAN**

3.1 Pendahuluan .....	III-1
3.2 Unit Penelitian.....	III-1
3.3 Metode Pengumpulan Data .....	III-1
3.3.1 Jenis Data .....	III-1
3.3.2 Sumber Data .....	III-2
3.3.3 Teknik Pengumpulan Data .....	III-2
3.4 Tahapan Penelitian .....	III-2
3.4.1 Menentukan Ruang Lingkup dan Unit Penelitian .....	III-3
3.4.2 Menentukan Dasar Teori yang Berkaitan dengan Penelitian ..	III-3
3.4.3 Menetapkan Kriteria Pengujian.....	III-3
3.4.4 Menentukan Alat yang Digunakan dalam Pelaksanaan Penelitian .....	III-4
3.4.5 Melakukan Pengujian Penelitian .....	III-5
3.4.6 Melakukan Analisa Hasil Pengujian dan Membuat Kesimpulan Penelitian .....	III-5
3.5 Metode Pengembangan Perangkat Lunak .....	III-6
3.5.1 Fase Insepsi .....	III-6
3.5.2 Fase Elaborasi .....	III-6
3.5.3 Fase Konstruksi .....	III-7
3.5.4 Fase Transisi.....	III-8
3.6 Penjadwalan Penelitian .....	III-8
3.7 Kesimpulan.....	III-17



## BAB IV PENGEMBANGAN PERANGKAT LUNAK

4.1 Pendahuluan .....	IV-1
4.2 Fase Insepsi .....	IV-1
4.2.1 Pemodelan Bisnis .....	IV-1
4.2.2 Kebutuhan Sistem .....	IV-2
4.2.2.1 Fitur Pengolahan .....	IV-3
4.2.3 Analisis dan Desain .....	IV-4
4.2.3.1 Analisis Perangkat Lunak .....	IV-4
4.2.3.2 Analisis Kebutuhan Perangkat Lunak .....	IV-4
4.2.3.3 Analisis Data .....	IV-4
4.2.3.4 Analisis Pra-pengolahan .....	IV-5
4.2.3.4.1 Pemecahan Kalimat .....	IV-5
4.2.3.4.2 <i>Case Folding</i> .....	IV-6
4.2.3.4.3 <i>Tokenizing</i> .....	IV-7
4.2.3.4.4 <i>Stop-word Removal</i> .....	IV-9
4.2.3.4.5 <i>Stemming</i> .....	IV-10
4.2.3.5 Analisis TF-IDF .....	IV-11
4.2.3.6 <i>Maximum Marginal Relevance</i> .....	IV-18
4.2.3.7 <i>F-measure</i> .....	IV-54
4.2.3.8 Desain Perangkat Lunak .....	IV-56
4.3 Fase Elaborasi .....	IV-59
4.3.1 Pemodelan Bisnis .....	IV-59
4.3.1.1 Perancangan Data .....	IV-60
4.3.1.1.1 Perancangan Antar Muka .....	IV-60
4.3.1.1.2 Perancangan Antar Muka	
Meringkas Dokumen .....	IV-61
4.3.2 Kebutuhan Sistem .....	IV-62
4.3.3 Diagram <i>Sequence</i> Meringkas Dokumen .....	IV-62
4.4 Fase Kontruksi .....	IV-65
4.4.1 Kebutuhan Sistem .....	IV-65
4.4.2 Diagram Kelas .....	IV-65

4.4.3 Implementasi .....	IV-67
4.4.3.1 Implementasi Kelas.....	IV-67
4.4.3.2. Implementasi Antar Muka .....	IV-68
4.4.3.2.1 Antar Muka Meringkas Dokumen .....	IV-69
4.5 Fase Transisi.....	IV-70
4.5.1 Pemodelan Bisnis .....	IV-71
4.5.2 Kebutuhan Sistem .....	IV-71
4.5.3 Rencana Pengujian .....	IV-71
4.5.3.1 Rencana Pengujian <i>Use Case</i>	
Meringkas Dokumen.....	IV-72
4.5.4 Implementasi .....	IV-74
4.5.4.1 Kasus Uji <i>Use Case</i> .....	IV-74
4.5.4.1.1 Pengujian <i>Use Case</i>	
Meringkas Dokumen .....	IV-75
4.6 Kesimpulan.....	IV-77

## **BAB V HASIL DAN ANALISIS PENELITIAN**

5.1 Pendahuluan .....	V-1
5.2 Analisis Pengujian.....	V-1
5.3 Hasil Pengujian Data .....	V-3
5.4 Analisis Hasil Pengujian .....	V-5
5.5 Kesimpulan.....	V-9

## **BAB VI KESIMPULAN DAN SARAN**

6.1 Pendahuluan .....	VI-1
6.2 Kesimpulan.....	VI-1
6.3 Saran.....	VI-2

<b>DAFTAR PUSTAKA .....</b>	<b>xviii</b>
-----------------------------	--------------

<b>LAMPIRAN.....</b>	<b>L1-1</b>
<b>LAMPIRAN HASIL PENGUJIAN .....</b>	<b>L1-1</b>
<b>LAMPIRAN KODE PROGRAM.....</b>	<b>L2-1</b>

## DAFTAR TABEL

	Halaman
III-1. Tabel Rancangan Hasil MMR .....	III-5
III-2. Tabel Penjadwalan Penelitian dalam bentuk <i>Work Breakdown Structure (WBS)</i> .....	III-9
IV-1. Tabel Kebutuhan Fungsional Perangkat Lunak.....	IV-3
IV-2. Tabel Kebutuhan Non Fungsional Perangkat Lunak.....	IV-3
IV-3. Tabel Proses Pemecahan Kalimat.....	IV-6
IV-4. Tabel Proses <i>Case-Folding</i> .....	IV-7
IV-5. Tabel Proses <i>Tokenizing</i> .....	IV-8
IV-6. Tabel Proses <i>Stopword Removal</i> .....	IV-9
IV-7. Tabel Proses <i>Stemming</i> .....	IV-10
IV-8. Tabel Nilai TF dan IDF .....	IV-12
IV-9. Tabel Nilai TF-IDF.....	IV-13
IV-10. Tabel Nilai $W(Q)*W(D_i)$ .....	IV-14
IV-11. Tabel Nilai Panjang Vektor .....	IV-15
IV-12. Tabel Nilai <i>Cosine Similarity</i> .....	IV-18
IV-13. Tabel Nilai MMR iterasi-1 .....	IV-21
IV-14. Tabel Bobot <i>Similarity</i> untuk iterasi-2 .....	IV-22
IV-15. Tabel Nilai MMR iterasi-2 .....	IV-25
IV-16. Tabel Bobot <i>Similarity</i> untuk iterasi -3 .....	IV-28
IV-17. Tabel Nilai MMR iterasi-3 .....	IV-31
IV-18. Tabel Bobot <i>Similarity</i> untuk iterasi-4 .....	IV-35
IV-19. Tabel Nilai MMR iterasi-4 .....	IV-38
IV-20. Tabel Bobot <i>Similarity</i> untuk itersi-5 .....	IV-42
IV-21. Tabel Nilai MMR iterasi-5 .....	IV-45
IV-22. Tabel Bobot <i>Similarity</i> untuk iterasi-6 .....	IV-50
IV-23. Tabel Nilai MMR iterasi-6 .....	IV-53
IV-24. Tabel Hasil Perangkingan MMR.....	IV-54
IV-25. Tabel Definisi Aktor <i>Use Case</i> .....	IV-57
IV-26. Tabel Definisi <i>Use Case</i> .....	IV-57

IV-27. Tabel Skenario <i>Use Case</i> Meringkas Dokumen.....	IV-58
IV-28. Tabel Menunjukkan Implementasi kelas dalam bahasa java.....	IV-67
IV-29. Tabel Rencana Pengujian <i>Use Case</i> Meringkas Dokumen .....	IV-72
IV-30. Tabel Pengujian <i>Use Case</i> Meringkas Dokumen .....	IV-75
V-1. Tabel Perbandingan Ringkasan Manual dan Sistem.....	V-3
V-2. Tabel Hasil Pengujian Ringkasan “Hari Tanpa Sahabat” .....	V-4
V-3. Tabel Hasil Jumlah <i>Query</i> dan Akurasi .....	V-5
V-4. Tabel Perbandingan Hasil Ringkasan Sistem dan Manual .....	V-6
V-5. Hasil Akurasi Cerita Pendek.....	V-7

## DAFTAR GAMBAR

	Halaman
Gambar II-1. Awalan dan Akhiran yang Tidak Diizinkan.....	II-9
Gambar II-2. Aturan Pemenggalan <i>Stemmer Nazief</i> dan Andriani .....	II-10
Gambar II-3. Jenis Awalan beserta tipe awalannya .....	II-10
Gambar III-3. Penjadwalan Penelitian Bagian 1 .....	III-13
Gambar III-4. Penjadwalan Penelitian Bagian 2.....	III-14
Gambar III-5. Penjadwalan Penelitian Bagian 3.....	III-15
Gambar III-6. Penjadwalan Penelitian Bagian 4.....	III-16
Gambar IV-1. Dokumen txt. Yang akan dipakai proses masukan.....	IV-5
Gambar IV-2. Dokumen txt. Yang diringkas dengan sistem.....	IV-55
Gambar IV-3. Dokumen txt. Yang diringkas dengan manual .....	IV-55
Gambar IV-4. Diagram <i>Use Case</i> .....	IV-57
Gambar IV-5. Diagram Aktivitas Meringkas Dokumen.....	IV-59
Gambar IV-6. Rancangan <i>Interface</i> Antar Muka.....	IV-60
Gambar IV-7. Rancangan <i>Interface</i> Meringkas Dokumen .....	IV-61
Gambar IV-8. Diagram <i>Sequence</i> Meringkas Dokumen .....	IV-64
Gambar IV-9. Diagram Kelas .....	IV-66
Gambar IV-10. <i>Interface</i> awal perangkat lunak Form Utama .....	IV-69
Gambar IV-11. <i>Interface</i> Meringkas Dokumen .....	IV-70
Gambar V-1. Tampilan Form Peringkasan “Hari Tanpa Sahabat” .....	V-3
Gambar V-2. Grafik <i>precision</i> , <i>Recall</i> dan <i>F-measure</i> .....	V-8

INDONESIAN DOCUMENT SUMMARY ON CERPENS USING THE  
METHOD MAXIMUM MARGINAL RELEVANCE

By :

Eva Sapayona

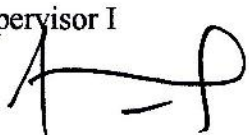
09021381320042

**ABSTRACT**

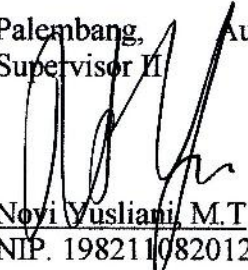
Reading activity is an activity that requires a relatively long time, a long time is very influential on a person's productivity. Therefore, summarizing text document automatically is needed to help get the short story core. The data used in this study are documents that have been formatted in the form of TXT. Data is obtained directly from the website [www.cerpenmu.com](http://www.cerpenmu.com). The amount of data collected is 20 short stories. In this study there are four stages, namely first, text preprocessing such as sentence solving folding case, tokenizing, stopword, and stemming. Second calculate the weight values in summary sentences using various features. The features contained are words that resemble titles, frequency of occurrence of words in sentences, and also the length of sentences. Third, determine the results of summary sentences obtained from sentence extraction using the maximum marginal relevance method. Fourth, sequencing the results of summary sentences so as to produce a summary that is easy to understand for the readers. The level of accuracy obtained by testing the summary results of the software compared to the summary results from expert experts so as to produce average precision 76.7%, recall 77.5%, f-measure 77%.

**Keywords:** Accuracy, Maximum Marginal Relevance Method, Automatic document text summarization, Textpreprocessing and TF-IDF.

Supervisor I

  
M. Fachrurrozi, S.Si., M.T.  
NIP. 198005222008121002

Palembang, August 2019  
Supervisor II

  
Noyi Yusliani, M.T  
NIP. 198211082012122001

Approved,  
Head of the Department of Informatics Engineering,



Rifkie Primartha, M.T.  
NIP. 197706012009121004

PERINGKASAN DOKUMEN BAHASA INDONESIA PADA CERPEN  
MENGUNAKAN METODE *MAXIMUM MARGINAL RELEVANCE*

Oleh :

Eva Sapayona

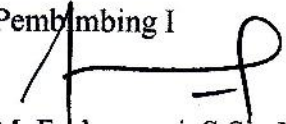
09021381320042

**ABSTRAK**

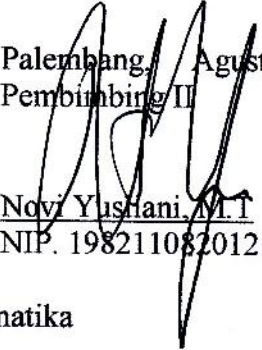
Kegiatan membaca merupakan kegiatan yang membutuhkan waktu relatif lama, waktu yang lama sangat berpengaruh pada produktifitas seseorang. Oleh karena itu, peringkasan teks dokumen secara otomatis sangat dibutuhkan untuk membantu mendapatkan inti cerpen. Data yang digunakan dalam penelitian ini berupa dokumen yang sudah diformat ke dalam bentuk TXT. Data diperoleh langsung dari website [www.cerpenmu.com](http://www.cerpenmu.com). Jumlah data yang telah dikumpulkan adalah 20 cerpen. Pada penelitian ini terdapat empat tahap yaitu pertama, text preprocessing seperti pemecahan kalimat, *case folding*, *tokenizing*, *stopword*, dan *stemming*. Kedua, menghitung nilai bobot dalam kalimat ringkasan dengan menggunakan berbagai fitur. Fitur-fitur yang terdapat yaitu kata yang menyerupai judul, frekuensi kemunculan kata dalam kalimat, dan juga panjangnya kalimat. Ketiga, menentukan hasil kalimat ringkasan yang didapat dari ekstraksi kalimat dengan menggunakan metode *maximum marginal relevance*. Keempat, mengurutkan hasil kalimat-kalimat ringkasan sehingga menghasilkan sebuah ringkasan yang mudah dipahami bagi para pembaca. Adapun tingkat akurasi yang diperoleh dengan cara melakukan pengujian terhadap hasil ringkasan perangkat lunak yang dibandingkan dengan hasil ringkasan dari ahli pakar sehingga menghasilkan rata-rata *precision* 76.7%, *recall* 77.5%, dan *f-measure* 77%.

**Kata Kunci:** Akurasi, Metode *Maximum Marginal Relevance*, Peringkasan teks dokumen otomatis, *Textpreprocessing* dan *TF-IDF*.

Pembimbing I

  
M. Fachrurrozi, S.Si., M.T.  
NIP. 198005222008121002

Palembang, Agustus 2019  
Pembimbing II

  
Novi Yustiani, M.T.  
NIP. 198211082012122001

Menyetujui,  
Ketua Jurusan Teknik Informatika



Rifkie Primartha, M.T.  
NIP. 197706012009121004



# **BAB I**

## **PENDAHULUAN**

### **1.1 Pendahuluan**

Bab 1 memberikan penjelasan umum mengenai keseluruhan penelitian. Antara lain membahas latar belakang masalah penelitian, perumusan masalah penelitian, tujuan penelitian, manfaat penelitian dan batasan masalah. Pendahuluan dimulai dengan penjelasan mengenai permasalahan waktu untuk membaca sebuah cerpen yang memerlukan waktu relatif lama dan penggunaan algoritma *Maximum Marginal Relevance* (MMR) sebagai metode untuk peringkasan teks sebuah cerpen pada penelitian ini.

### **1.2 Latar Belakang**

Pada era-globalisasi sekarang ini budaya membaca perlahan mulai menghilang, dikarenakan generasi anak muda zaman sekarang merupakan generasi milenial. Dimana generasi ini teknologi berkembang dengan pesatnya yang membuat anak muda lebih memilih *handphone* dibandingkan buku. Selain itu juga tebalnya buku dan panjangnya cerita merupakan alasan lain yang membuat generasi sekarang malas untuk membaca, sekalipun itu cerita pendek atau cerpen. Kegiatan membaca merupakan kegiatan yang membutuhkan waktu relatif lama, waktu yang lama sangat berpengaruh pada efisiensi waktu dan produktifitas seseorang.

Peringkasan teks menjadi alternatif atau solusi dalam penyelesaian permasalahan ini, munculah sebuah ide untuk meringkas teks pada cerpen dengan menggunakan sebuah metode yaitu metode *Maximum Marginal Relevance* (MMR). Peringkasan teks (*Automatic Text Summarization*) adalah penyortiran beberapa paragraf menjadi bentuk yang lebih singkat menggunakan aplikasi yang dioperasikan dalam komputer (Ajmal dan Haroon, 2016). Algoritma *Maximum Marginal Relevance* (MMR) merupakan salah satu metode peringkasan yang digunakan untuk meringkas dokumen tunggal. MMR meringkas dokumen dengan menghitung kesamaan (*Similarity*) antara bagian teks (Indriani A. 2014).

Metode MMR dipilih agar bisa menyelesaikan permasalahan panjangnya sebuah cerita, karena untuk mendapatkan inti pokok dari informasi tersebut membutuhkan waktu yang cukup lama. Setidaknya pembaca harus membaca halaman demi halaman dari sumber referensi. Sebuah teks yang dimasukkan dalam komputer diolah dan diringkas dari teks asli. Teknik ini dipelopori oleh Luhan sejak tahun 1958 yang telah dikembangkan hingga kini menggunakan berbagai metode (Mustaqhfiri et al., 2012). Beberapa metode yang sering digunakan diantaranya adalah metode Genetika, *Lexrank*, *Maximum Marginal Relevance*, *Graph Based*. Metode ekstraksi dokumen yang digunakan dalam penelitian ini adalah metode *Maximum Marginal Relevance*.

Metode MMR pernah digunakan pada peringkasan sebuah dokumen diantaranya yaitu peringkasan teks otomatis sinopsis buku berbahasa Indonesia, yang dilakukan oleh (Indriani A, 2014). Selain itu metode MMR pernah digunakan untuk peringkasan dokumen berupa dokumen buku yang dilakukan

oleh (Pardedel J dan Sinatria J, 2015). Ditinjau dari latar belakang metode MMR diharapkan mampu untuk menyelesaikan permasalahan mengenai peringkasan dokumen teks pada cerpen

### **1.3 Perumusan Masalah**

Rumusan masalah pada penelitian ini adalah dengan menerapkan metode *Maximum Marginal Relevance* (MMR) dan pembobotan kata menggunakan *Term Frequency-Invers Dokument Frequency* (TF-IDF) ke dalam peringkasan dokumen bahasa Indonesia untuk meringkas cerpen. untuk menjawab rumusan masalah tersebut, diuraikan beberapa *research question* sebagai berikut :

1. Bagaimana mengimplementasi metode *Maximum Marginal Relevance* (MMR) untuk sistem peringkasan dokumen bahasa Indonesia pada perangkat lunak.
2. Bagaimana cara untuk mengetahui Seberapa tinggi tingkat akurasi sistem peringkasan dokumen bahasa indonesia yang menggunakan metode *Maximum Marginal Relevance* (MMR).

### **1.4 Tujuan Penelitian**

Tujuan dilakukannya penelitian adalah sebagai berikut :

1. Mengembangkan sistem peringkasan dokumen bahasa Indonesia pada cerpen menggunakan metode *Maximum Marginal Relevance* (MMR).

2. Mengetahui tingkat akurasi peringkasan dokumen bahasa Indonesia pada cerpen menggunakan metode *Maximum Marginal Relevance* (MMR).

### **1.5 Manfaat Penelitian**

Manfaat yang dapat diperoleh dari penelitian ini adalah sebagai berikut :

1. Menghemat waktu bagi para pembaca dalam memahami isi cerita melalui ringkasan cerpen.
2. Menambah ilmu pengetahuan, menjadi bahan referensi dan perbandingan untuk penelitian yang terkait dengan penggunaan metode *Maximum Marginal Relevance* (MMR).

### **1.6 Batasan Masalah**

Batasan masalah pada penelitian ini adalah sebagai berikut :

1. Dokumen yang digunakan pada penelitian ini berupa dokumen dalam bahasa Indonesia.
2. Dokumen yang digunakan berupa teks dengan format \*.txt.
3. Tidak menangani kesalahan penulisan pada teks dokumen yang akan di ringkas.
4. Sistem peringkasan teks hanya memproses kata dan kalimat, tidak untuk rumus, simbol dan table.

## **1.7 Sistematika Penulisan**

Sistematika penulisan skripsi ini adalah sebagai berikut :

### **BAB I. PENDAHULUAN**

Pada bab ini diuraikan mengenai latar belakang, perumusan masalah, tujuan, manfaat penelitian dan batasan masalah/ruang lingkup, metodologi penelitian, dan sistematika penulisan.

### **BAB II. KAJIAN LITERATUR**

Pada bab ini akan dibahas dasar-dasar teori yang digunakan dalam penelitian, seperti definisi-definisi *TF-IDF*, *Cosine Similarity*, *Maximum Marginal Relevance*, *Precision* dan *Recall* serta *f-measure*.

### **BAB III. METODOLOGI PENELITIAN**

Pada bab ini akan dibahas mengenai tahapan yang akan dilaksanakan pada penelitian ini. Masing-masing rencana tahapan penelitian dideskripsikan dengan rinci mengacu pada suatu kerangka kerja. Di akhir bab ini berisi perancangan manajemen proyek pada pelaksanaan penelitian.

### **BAB IV. PENGEMBANGAN PERANGKAT LUNAK**

Pada bab ini akan dibahas mengenai perancangan dan lingkungan Implementasi Metode *Maximum Marginal Relevance* terhadap peringkasan dokumen cerita pendek, hasil peringkasan, dan hasil pengujian.

## **BAB V. HASIL DAN ANALISIS PENELITIAN**

Pada bab ini, hasil pengujian berdasarkan langkah-langkah yang telah direncanakan disajikan. Analisis diberikan sebagai basis dari kesimpulan yang diambil dalam penelitian ini.

## **BAB VI. KESIMPULAN DAN SARAN**

Pada bab ini berisi kesimpulan dari semua uraian-uraian pada bab-bab sebelumnya dan juga berisi saran-saran yang diharapkan berguna dalam peringkasan cerpen.

### **1.8 Kesimpulan**

Penelitian ini akan berfokus untuk pembobotan hubungan suatu kata pada dokumen atau kalimat dalam mengimplementasikan *Term Frequency-Invers Dokumen Frequency* (TF-IDF) terhadap peringkasan cerpen. Selain itu juga menggunakan *Cosine Similarity* untuk menghitung pendekatan relevansi *Query* terhadap dokumen dan metode *Maximum Marginal Relevance* yang dipergunakan untuk meringkas dokumen tunggal guna mengukur hasil ringkasan yang *Relevance* serta *Precision*, *Recall* dan *F-measure* digunakan untuk mengetahui tingkat akurasi.

## DAFTAR PUSTAKA

- Agusta, L. (2009). Perbandingan algoritma stemming Porter dengan algoritma Nazief & Adriani untuk stemming dokumen teks bahasa indonesia. *Konferensi Nasional Sistem dan Informatika*, 196-201.
- Ajmal, E. B., & Haroon, R. P. (2016). Maximal Marginal Relevance Based Malayalam Text Summarization with Successive Thresholds. *International Journal on Cybernetics & Informatics*, 5, 349-356.
- Budi Susanto, "Text Summarization", in *Sebuah Pengantar TI UKDW*, Februari 4, 2013.
- C.D. Manning, P.Raghavan dan H.Schütze,"An Introduction toInformation Retrieval", in Cambridge University Press, pp.26, April 1, 2009.
- Carbonell, J., & Goldstein, J. (1998, August). The use of MMR, diversity-based reranking for reordering documents and producing summaries. In *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 335-336). ACM.
- Flores, J. G., Gillard, L., Ferret, O., & de Chalendar, G. (2008, November). Bag of Senses Versus Bag of Words: Comparing Semantic and Lexical Approaches on Sentence Extraction. In *TAC*.
- Fuji Mulia, <http://www.trigonalworld.com/2013/05/pengertian-dan-tujuan-ringkasan.html>.
- Hovy, E. H., Hermjakob, U., & Lin, C. Y. (2001, November). The Use of External Knowledge of Factoid QA. In *TREC* (Vol. 2001, pp. 644-52).

- Indriani, A. (2014). MAXIMUM MARGINAL RELEVANCE UNTUK PERINGKASAN TEKS OTOMATIS SINOPSIS BUKU BERBAHASA INDONESIA. *SEMNASTEKNOMEDIA ONLINE*, 2(1), 3-05.
- Jones, K. S., & Galliers, J. (1996). Evaluating natural language processing systems: An analysis and review. *Nat. Lang. Eng.*, 4(2), 175-190.
- Mani, I., & Maybury, M. T. (Eds.). (1999). *Advances in automatic text summarization* (Vol. 293). Cambridge, MA: MIT press.
- Mohammad, S., Dorr, B. J., Egan, M., Madnani, N., Zajic, D. M., & Lin, J. J. (2008). Multiple Alternative Sentence Compressions and Word-Pair Antonymy for Automatic Text Summarization and Recognizing Textual Entailment. In *TAC*.
- Mustaqhfiri, M., Abidin, Z., & Kusumawati, R. (2012). Peringkasan Teks Otomatis Berita Berbahasa Indonesia Menggunakan Metode Maximum Marginal Relevance. *MATICS*.
- Suciadi, J. (2001). Studi analisis metode-metode parsing dan interpretasi semantik pada natural language processing. *Jurnal Informatika*, 2(1), pp-13.
- Xie, S., & Liu, Y. (2008, March). Using corpus and knowledge-based similarity measure in maximum marginal relevance for meeting summarization. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on* (pp. 4985-4988). IEEE.