

PERINGKASAN TEKS DENGAN METODE K-MEANS

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika



Oleh :

Ari Firdaus
NIM : 09021381722129

Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2021

PERINGKASAN TEKS DENGAN METODE K-MEANS

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika



Oleh :

Ari Firdaus
NIM : 09021381722129

Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2021

LEMBAR PENGESAHAN SKRIPSI

PERINGKASAN TEKS DENGAN METODE K-MEANS

Oleh :

Ari Firdaus
NIM : 09021381722129

Palembang, 04 Agustus 2021

Pembimbing I



Novi Yusufiani, S.Kom., M.T
NIP. 198211082012122001

Pembimbing II



Desty Rodiah, S.Kom., MT.
NIP. 198912212020122011

Mengetahui,
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom.
NIP 197812222006042003

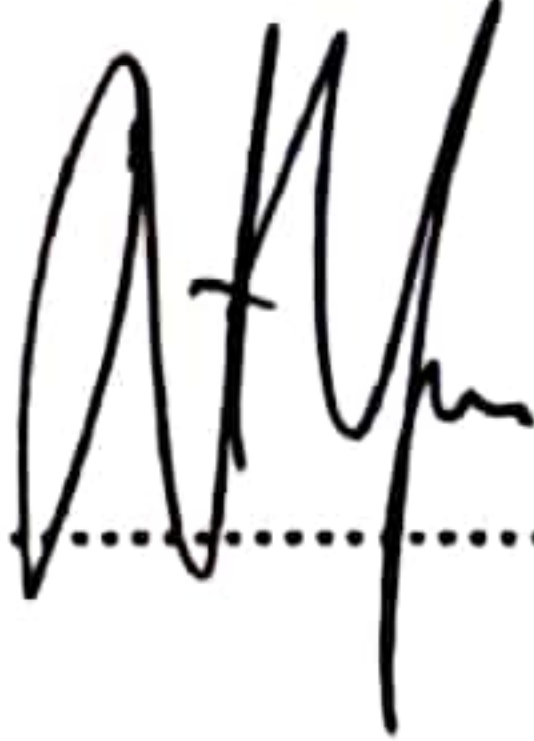
TANDA LULUS UJIAN SIDANG SKRIPSI

Pada hari Kamis tanggal 5 Agustus 2021 telah dilaksanakan ujian sidang skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

Nama : Ari Firdaus
NIM : 09021381722129
Judul : Peringkasan Teks Dengan Metode K-Means

1. Pembimbing I

Novi Yusliani, M.T.
NIP. 198211082012122001



.....

2. Pembimbing II


Desty Rodiah, S.Kom., MT.
NIP. 198912212020122011



.....

3. Penguji I


Dian Palupi Rini, M.Kom., Ph.D.
NIP. 198410012009121005



.....

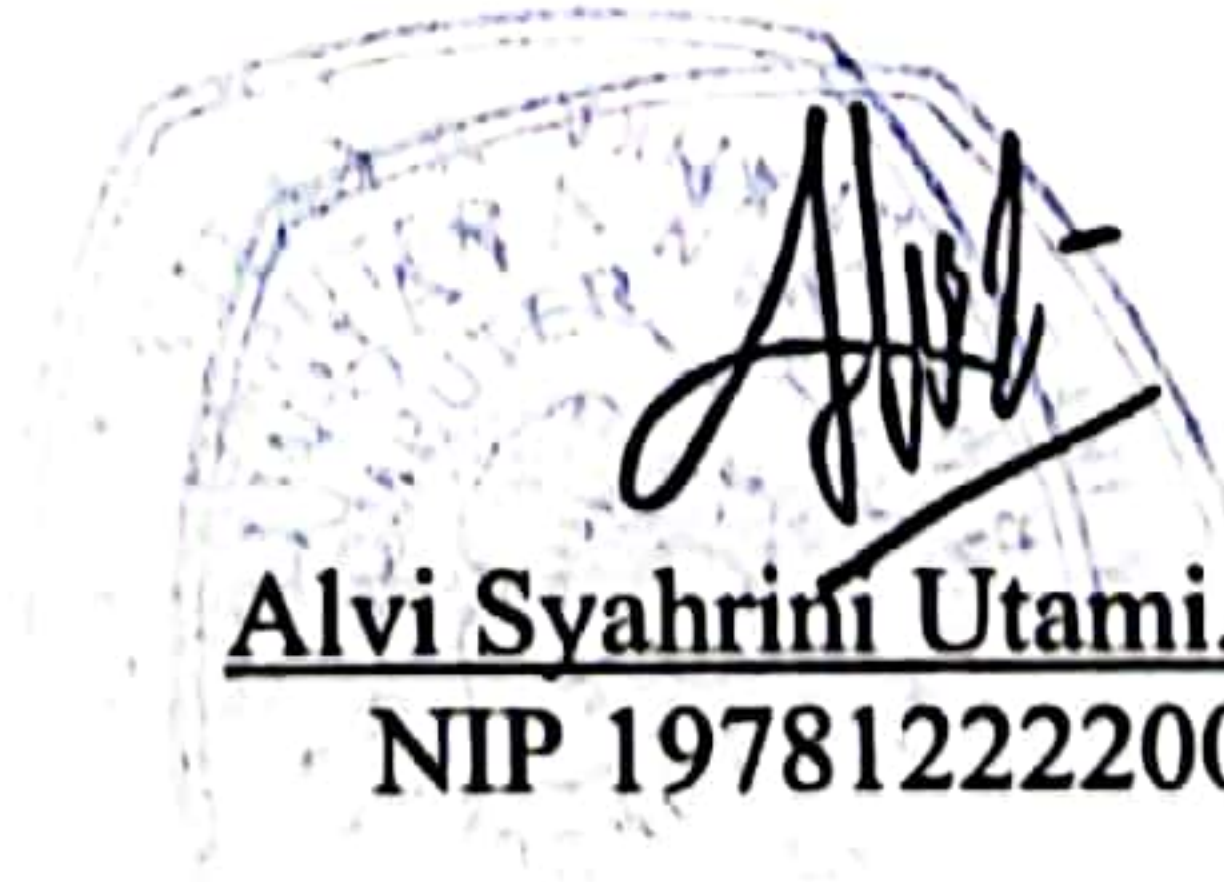
4. Penguji II

Muhammad Naufal Rachmatullah, M.T.
NIP.



.....

Mengetahui,
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom.
NIP 197812222006042003

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Ari Firdaus
NIM : 09021381722129
Program Studi : Teknik Informatika Bilingual
Judul Skripsi : Peringkasan Teks dengan Metode K-Means
Hasi Pengecekan Software
iThenticate/Turnitin : 11%

Menyatakan bahwa laporan proyek saya merupakan hasil karya saya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan proyek ini, makasaya 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, 25 Agustus 2021



(Ari Firdaus)

NIM. 09021381722129

MOTTO DAN PERSEMBAHAN

MOTTO:

“Always remember that you are absolutely unique. Just like everyone else.”

- Margaret Mead

“Whether you think you can or you think you can't, you're right.”

- Henry Ford

“Just do it!”

- Penulis

Kupersembahkan karya tulis ini kepada :

- *Allah SWT & Nabi Muhammad SAW*
- *Ibu & ayahku tercinta*
- *Keluarga besarku*
- *Dosen pembimbing & dosen penguji*
- *Sahabat & teman seperjuanganku*
- *Almamater*

ABSTRACT

Text Summarization is a tool used to generate a short form of text that contains important information that is needed by the user automatically. In this study, Text Summarization was conducted on Indonesian news using K-Means method. The news is taken from CNN Indonesia with a free topic. K-Means is used to classify sentences that already have weight in the news with 2 clusters, namely text summaries and not text summaries. The initial centroid is selected based on the sentence with the largest value and the sentence with the smallest value. The test conducted on Indonesian news with a total 50 news and tested for feasibility using a questionnaire. K-Means was successfully summarizing the news with an average 27.3 % of original news length and gain 87% good summarize based on respondents from questionnaire.

Keywords: Text Summarization, K-Means, Natural Language Processing (NLP)

ABSTRAK

Peringkasan teks adalah alat yang digunakan untuk menghasilkan bentuk ringkas suatu teks yang berisi informasi penting didalamnya yang diperlukan oleh pengguna secara otomatis. Pada penelitian ini, peringkasan teks dilakukan pada teks berita berbahasa Indonesia menggunakan metode K-Means. Teks berita diambil dari CNN Indonesia dengan topik bebas. K-Means digunakan untuk mengelompokkan kalimat yang telah memiliki bobot pada berita tersebut dengan 2 kelompok yaitu ringkasan teks dan bukan ringkasan teks. Centroid awal yang digunakan dipilih berdasarkan kalimat dengan nilai terbesar dan kalimat dengan nilai terkeci. Pengujian dilakukan pada berita berbahasa Indonesia sebanyak 50 berita dan diuji kelayakannya dengan menggunakan kuesioner. K-Means berhasil meringkas berita dengan rata-rata 27.3% dari panjang berita asli dan mendapatkan 87% ringkasan yang layak berdasarkan responden dari kuesioner.

Kata Kunci: Peringkasan Teks, K-Means, Pemrosesan Bahasa Alami

KATA PENGANTAR

Assalamualaikum warahmatullahi wabarakatuh, puji syukur kepada Allah atas berkat dan rahmat-Nya yang diberikan kepada penulis sehingga dapat menyelesaikan tugas akhir yang berjudul “PERINGKASAN TEKS DENGAN METODE K-MEANS” dengan baik. Tugas akhir ini disusun untuk memenuhi salah satu syarat untuk menyelesaikan pendidikan program Strata-1 pada Fakultas Ilmu Komputer Program Studi Teknik Informatika di Universitas Sriwijaya.

Untuk selanjutnya penyusun mengucapkan banyak terima kasih kepada pihak-pihak yang telah membantu dalam penyelesaian skripsi ini, yaitu:

1. Bapak Jaidan Jauhari, S. Pd. M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
2. Ibu Alvi Syahrini Utami, M.Kom. selaku Kepala Jurusan Teknik Informatika Universitas Sriwijaya.
3. Ibu Desty Rodiah, S.Kom., MT. selaku dosen pembimbing akademik dan dosen pembimbing 2 tugas akhir yang telah membantu penulis dalam mengikuti, membimbing, dan menyelesaikan studi di Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Ibu Novi Yusliani, S.Kom., M.T selaku dosen pembimbing 1 tugas akhir yang telah membantu penulis dalam membimbing dan menyelesaikan tugas akhir.
5. Seluruh staff pengajar Fakultas Ilmu Komputer yang telah memberikan ilmu pengetahuan yang tak ternilai selama penulis menempuh pendidikan di Fakultas Ilmu Komputer Universitas Sriwijaya.
6. Kedua orang tua penulis, Abdul Rahman dan Siti Aisyah, yang selalu memberikan kasih sayang, doa, nasehat, serta atas kesabarannya yang luar biasa dalam setiap langkah penulis, yang merupakan anugrah terbesar.

7. Saudara-saudaraku, Ria, Ira, dan Raihan, terimakasih atas doa dan segala dukungan.
8. Teman-teman kuliah, terimakasih atas kesenangan, canda tawa yang membahagiakan dan menjadi keluarga baru bagi penulis.
9. Seluruh responden yang telah meluangkan waktunya untuk mengisi kuesioner.
10. Semua pihak yang tidak dapat disebutkan satu pe rsatu, yang telah memberikan doa dengan ikhlas serta dukungan sehingga dapat menyelesaikan skripsi ini.

Penulis menyadari dalam penyusunan tugas akhir ini masih banyak kekurangan yang disebabkan kurangnya pengetahuan dan pengalaman. Oleh karena itu kritik dan saran sangat diterima untuk kemajuan penelitian berikutnya. Akhir kata semoga tugas akhir ini dapat berguna dan bermanfaat bagi kita semua. Wassalam.

Palembang, 04 Agustus 2021



Ari Firdaus

DAFTAR ISI

	Halaman
LEMBAR PENGESAHAN SKRIPSI	ii
TANDA LULUS UJIAN SIDANG SKRIPSI	iii
MOTTO DAN PERSEMBAHAN	iv
ABSTRACT	v
ABSTRAK	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xv
BAB I PENDAHULUAN	I-1
1.1 Pendahuluan	I-1
1.2 Latar Belakang	I-1
1.3 Rumusan Masalah	I-3
1.4 Tujuan Penelitian	I-4
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 KAJIAN LITERATUR	II-1
2.1 Pendahuluan	II-1
2.2 Landasan Teori	II-1
2.2.1 Peringkasan Teks Otomatis	II-1
2.2.3 Praproses Teks	II-2
2.2.4 Pembobotan	II-6
2.2.5 <i>K-Means</i>	II-7
2.2.6 <i>Extreme Programming</i>	II-9

2.3	Penelitian Terkait.....	II-10
2.4	Kesimpulan.....	II-11
BAB III METODOLOGI PENELITIAN		III-1
3.1	Pendahuluan	III-1
3.2	Pengumpulan Data	III-1
3.2.1	Jenis Data	III-1
3.2.2	Sumber Data.....	III-1
3.2.3	Metode Pengumpulan Data.....	III-1
3.3	Tahapan Penelitian	III-2
3.3.1	Kerangka Kerja	III-2
3.3.2	Kriteria Pengujian	III-12
3.3.3	Format Data Pengujian.....	III-13
3.3.4	Alat yang Digunakan dalam Penelitian.....	III-15
3.3.5	Pengujian Penelitian.....	III-15
3.3.6	Analisis Hasil Pengujian dan Membuat Kesimpulan.....	III-16
3.4	Metode Pengembangan Perangkat Lunak	III-17
3.5	Manajemen Proyek Penelitian.....	III-18
BAB IV PENGEMBANGAN PERANGKAT LUNAK		IV-1
4.1	Pendahuluan	IV-1
4.2	Perencanaan Awal	IV-1
4.3	Iterasi Pertama	IV-4
4.3.1	Planning	IV-4
4.3.2	<i>Design</i>	IV-16
4.3.3	<i>Coding</i>	IV-28
4.3.4	<i>Testing</i>	IV-28
4.4	Iterasi Kedua.....	IV-32
4.4.1	<i>Planning</i>	IV-32
4.4.2	<i>Design</i>	IV-32
4.4.3	<i>Coding</i>	IV-37
4.4.4	<i>Testing</i>	IV-37

BAB V HASIL DAN ANALISIS PENELITIAN.....	V-1
5.1 Pendahuluan	V-1
5.2 Data Hasil Percobaan/Penelitian	V-1
5.2.1. Konfigurasi Percobaan	V-1
5.2.2 Data Hasil Konfigurasi I	V-1
5.2.3 Data Hasil Konfigurasi II	V-4
5.3 Analisis Hasil Penelitian	V-7
BAB VI KESIMPULAN DAN SARAN.....	VI-1
6.1 Pendahuluan	VI-1
6.2 Kesimpulan.....	VI-1
6.3 Saran	VI-2
Daftar Pustaka	iv
Lampiran	vi

DAFTAR TABEL

	Halaman
Tabel II-1 Contoh Segmentasi Kalimat.....	II-3
Tabel II-2 Contoh Case Folding.....	II-4
Tabel II-3 Contoh Tokenisasi.....	II-5
Tabel II-4. Contoh Stemming	II-6
Tabel II-5. Metode <i>Extreme Programming</i>	II-9
Tabel III-1. Penyajian Dokumen Asli dan Dokumen Ringkasan.....	III-13
Tabel III-2 Tabel Persentase Jumlah Kalimat Ringkasan	III-14
Tabel III-3 Tabel Kelayakan Hasil Sistem.....	III-14
Tabel III-4. Hasil Pengujian Rata-Rata Panjang Ringkasan	III-16
Tabel III-5. Hasil Pengujian Kealayaan Ringkasan Sistem	III-17
Tabel IV-1. Kebutuhan Fungsional.....	IV-2
Tabel IV-2 Kebutuhan Non-Fungsional	IV-3
Tabel IV-3. Iterasi dan Fokusnya.....	IV-3
Tabel IV-4. Contoh masukan dari pengguna dan statusnya.....	IV-5
Tabel IV-5 Contoh Segmentasi Kalimat	IV-6
Tabel IV-6. Contoh Penyeragaman Karakter Dalam Kalimat	IV-7
Tabel IV-7. Contoh Tokenisasi	IV-8
Tabel IV-8. Contoh <i>Stemming</i>	IV-9
Tabel IV-9. Contoh Pembobotan Kata Menggunakan TF-IDF	IV-10
Tabel IV-10 Contoh Pembobotan Kalimat	IV-11

Tabel IV-11 Centroid Awal	IV-13
Tabel IV-12 Pengelompokan Kalimat perulangan pertama.....	IV-13
Tabel IV-13 Penentuan Nilai <i>Centroid</i> perulangan pertama.....	IV-14
Tabel IV-14 Pengelompokan Kalimat perulangan Kedua	IV-14
Tabel IV-15 Penentuan Nilai centroid perulangan kedua	IV-15
Tabel IV-16 Kelompok dan Anggota Hasil <i>K-Means</i>	IV-15
Tabel IV-17 Rancangan Skenario Memasukan Tautan	IV-16
Tabel IV-18 Rancangan Skenario Melihat Hasil Ringkasan	IV-16
Tabel IV-19 Rancangan Skenario Melihat Rincian Ringkasan	IV-16
Tabel IV-20 Definisi Aktor.....	IV-17
Tabel IV-21. Definisi <i>Use Case</i>	IV-18
Tabel IV-22. Skenario <i>Use Case</i> Memasukan Tautan Berita	IV-19
Tabel IV-23 Skenario <i>Use Case</i> Melihat Hasil Ringkasan.....	IV-20
Tabel IV-24 Skenario <i>Use Case</i> Melihat Rincian Ringkasan.....	IV-21
Tabel IV-25 Tabel Kelas.....	IV-27
Tabel IV-26. Pengujian Skenario Memasukan Tautan Berita	IV-29
Tabel IV-27 Pengujian Skenario Melihat Hasil Ringkasan	IV-30
Tabel IV-28 Pengujian Skenario Melihat Rincian Ringkasan	IV-31
Tabel IV-29 Pengujian Skenario Memasukan Tautan Berita	IV-38
Tabel IV-30 Pengujian Skenario Melihat Hasil Ringkasan	IV-39
Tabel IV-31 Pengujian Skenario Melihat Rincian Ringkasan	IV-40
Tabel V-1. Persentase Jumlah Kalimat Ringkasan	V-2
Tabel V-2. Hasil Kuesioner.....	V-5

Tabel V-3 Persentase Rata-Rata Panjang Ringkasan.....	V-7
Tabel V-4 Hasil Kuesioner.....	V-8

DAFTAR GAMBAR

	Halaman
Gambar II-1. Peringkasan Teks Otomatis.....	II-2
Gambar III-1. Tahapan Penelitian dalam Bentuk Card Based.....	III-2
Gambar III-2. Identifikasi Masalah.....	III-3
Gambar III-3. Analisis Masalah.....	III-3
Gambar III-4. Studi Literatur	III-4
Gambar III-5. Dokumentasi	III-4
Gambar III-6. Proposal Skripsi	III-5
Gambar III-7. Skripsi	III-6
Gambar III-8. Pengembangan Perangkat Lunak	III-6
Gambar III-9. <i>Front-End</i>	III-7
Gambar III-10. <i>Back-End</i>	III-8
Gambar III-11. <i>Flowchart</i> Sistem Peringkasan Teks.....	III-8
Gambar III-12. Pengujian Sistem.....	III-10
Gambar III-13. Pengujian Penelitian	III-11
Gambar III-14. Analisis Hasil Sistem	III-11
Gambar III-15. Kuisisioner Hasil	III-12
Gambar III-16. Kalender Perencanaan Proyek Penelitian	III-19
Gambar IV-1. <i>Use Case Diagram</i>	IV-17
Gambar IV-2 <i>Activity Diagram</i> Memasukan Tautan Berita	IV-22

Gambar IV-3 <i>Activity Diagram</i> Melihat Hasil Ringkasan.....	IV-23
Gambar IV-4 <i>Activity Diagram</i> Melihat Rincian Ringkasan.....	IV-24
Gambar IV-5. <i>Sequence Diagram</i> Memasukan Tautan Berita	IV-25
Gambar IV-6. <i>Sequence Diagram</i> Melihat Hasil Ringkasan	IV-26
Gambar IV-7. <i>Sequence Diagram</i> Melihat Rincian Ringkasan	IV-26
Gambar IV-8 <i>Class Diagram</i> Perangkat Lunak.....	IV-28
Gambar IV-9 Tampilan Utama Perangkat Lunak Untuk Platform PC	IV-33
Gambar IV-10 Tampilan Utama Perangkat Lunak Untuk <i>Platform Mobile</i>	IV-33
Gambar IV-11 Tampilan Hasil Ringkasan Perangkat Lunak Untuk <i>Platform</i> PC	IV-34
Gambar IV-12 Tampilan Hasil Ringkasan Perangkat Lunak Untuk <i>Platform Mobile</i>	IV-35
Gambar IV-13 Tampilan Rincian Ringkasan Perangkat Lunak Untuk <i>Platform</i> PC	IV-36
Gambar IV-14 Tampilan Rincian Ringkasan Perangkat Lunak Untuk <i>Platform</i> <i>Mobile</i>	IV-36

BAB I

PENDAHULUAN

1.1 Pendahuluan

Bab ini akan menjelaskan secara umum keseluruhan penelitian, diantaranya adalah latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, sistematika penulisan, dan kesimpulan dari bab ini.

1.2 Latar Belakang

Berita sangat bermanfaat untuk menginformasikan kepada masyarakat tentang peristiwa-peristiwa yang ada di sekitar mereka. Namun, berkembangnya internet dengan pesat berdampak pada banyaknya berita dan jumlah informasi yang mengakibatkan sulitnya mendapatkan maksud dari informasi secara efisien dan efektif (Lewis, 2006). Oleh sebab itu diperlukan alat untuk meringkas suatu informasi sehingga memudahkan pembaca mendapatkan informasi. Alat ini disebut dengan peringkasan teks otomatis. Peringkasan teks otomatis adalah alat yang digunakan untuk menghasilkan bentuk ringkas suatu teks yang berisi informasi penting didalamnya yang diperlukan oleh pengguna secara otomatis (Agrawal & Gupta, 2014).

Berdasarkan jumlah dokumen yang diberikan sebagai masukan, peringkasan teks terbagi menjadi dua yaitu, *Single Document Summarization* dan *Multi Document Summarization*. *Single Document Summarization* menggunakan satu dokumen sebagai masukannya, sedangkan *Multi Document Summarization* menggunakan dokumen lebih dari satu sebagai masukannya (Pandya, 2019).

Secara umum ada dua metode dalam melakukan peringkasan teks otomatis, yaitu ekstraktif dan abstraktif (Akter, et al., 2017). Metode ekstraktif melakukan peringkasan dengan cara memilih kalimat yang ada pada teks, sedangkan metode abstraktif melakukan peringkasan dengan membuat kalimat baru (Agrawal & Gupta, 2014). Menurut Akter, et al. (2017) terdapat beberapa teknik yang dapat digunakan untuk pendekatan metode ekstraktif, antara lain: metode TF-IDF, *cluster based*, *graph theoretic approach*, *machine learning approach*, dan *k-means clustering*. Pendekatan metode ekstraktif sering memberikan hasil yang lebih baik dibandingkan dengan pendekatan metode abstraktif. Hal ini disebabkan karena masalah dalam pendekatan metode abstraktif seperti *semantic representation*, *inference* dan *natural language generation* relatif lebih sulit dibandingkan pendekatan berbasis data seperti ekstraksi kalimat. Karena itu, metode ekstraktif yang dibantu dengan metode *unsupervised learning* lebih banyak dipilih dalam melakukan pengelompokan pada peringkasan teks (Agrawal & Gupta, 2014).

Metode *K-Means* merupakan metode paling sederhana dan paling banyak digunakan dalam metode pengelompokan. Pada peringkasan teks, *K-Means* digunakan untuk mengelompokkan kalimat yang merupakan ringkasan dan bukan ringkasan. *K-Means* memiliki kemampuan untuk mengelompokkan data dalam jumlah yang besar dengan cepat dan efisien (Khotimah, Irhamni, & Sundarwati, 2016). Namun, *K-Means* memiliki kelemahan yaitu bergantung pada pengelompokan awal. Jika pemilihan *Centroid* awal dalam pengelompokannya tidak tepat maka hasil pengelompokan hanya akan menjadi lokal optimum,

sehingga penentuan *Centroid* awal yang baik sangat diperlukan (Celebi, A., & Vela, 2012).

Penelitian yang dilakukan (Agrawal & Gupta, 2014) menggunakan *K-Means Clustering* untuk melakukan peringkasan teks. Peringkasan dilakukan dengan masukan satu dokumen yang menggunakan bahasa Inggris. Untuk mengelompokkan ringkasan digunakan *Centroid* awal yang banyaknya ditentukan melalui aturan dan nilainya dipilih secara acak dari kalimat yang telah diberi nilai. Penelitian ini memperoleh peringkasan teks berupa 35-50% kalimat dari dokumen yang ada.

Penelitian yang dilakukan (Akter, et al., 2017) menggunakan *K-Means Clustering Algorithm* untuk mendapatkan bentuk ringkas suatu teks. Bahasa yang digunakan adalah Bangali dan jumlah masukan berupa satu dokumen dan lebih dari satu dokumen. Kalimat dengan nilai tertinggi dan kalimat dengan nilai terendah dipilih sebagai *Centroid* awal. Penelitian ini memperoleh teks berupa 30% kalimat dari dokumen yang ada untuk dipilih sebagai ringkasan.

Oleh karena itu, penelitian ini akan membuat sistem peringkasan teks dengan pendekatan ekstraktif menggunakan metode *K-Means* dengan *Centroid* awal yang ditentukan berdasarkan kalimat dengan nilai terbesar dan kalimat dengan nilai terendah. Bahasa yang digunakan adalah bahasa Indonesia dengan masukan satu dokumen berupa teks berita.

1.3 Rumusan Masalah

Berdasarkan permasalahan yang telah dijelaskan pada latar belakang maka rumusan masalah pada penelitian ini adalah :

1. Bagaimana mekanisme *K-Means Clustering* dalam melakukan peringkasan teks?
2. Bagaimana kinerja *K-Means Clustering* dalam melakukan peringkasan teks berdasarkan jumlah kalimat yang dihasilkan?
3. Apakah hasil ringkasan teks menggunakan *K-Means Clustering* dapat diterima oleh pembaca?

1.4 Tujuan Penelitian

Sesuai dengan rumusan masalah, maka dapat ditentukan tujuan penelitian sebagai berikut :

1. Menghasilkan aplikasi peringkasan teks menggunakan metode *K-Means Clustering*.
2. Mengetahui rata-rata jumlah kalimat ringkasan yang dihasilkan oleh sistem.
3. Mengetahui kelayakan ringkasan teks yang dihasilkan sistem.

1.5 Manfaat Penelitian

Manfaat yang diperoleh dari penelitian ini adalah :

1. Dapat membantu pembaca dokumen/artikel untuk lebih cepat dalam mendapatkan informasi penting pada dokumen/artikel tersebut.
2. Membantu penelitian dibidang peringkasan teks khususnya bahasa Indonesia.

1.6 Batasan Masalah

Batasan masalah di penelitian ini adalah :

1. Bahasa yang digunakan adalah bahasa Indonesia.
2. Masukan berupa tautan berita bebas topik yang berasal dari CNN Indonesia.

1.7 Sistematika Penulisan

Penyusunan skripsi ini disusun dengan sistematika penulisan sebagai berikut:

BAB I. PENDAHULUAN

Pada bab ini diuraikan mengenai pokok-pokok pikiran yang melandasi pembuatan penelitian, seperti latar belakang, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah serta sistematika penulisan.

BAB II. KAJIAN LITERATUR

Pada bab ini membahas dasar-dasar teori yang digunakan dalam penelitian, seperti definisi peringkasan teks, algoritma *k-means*, serta beberapa kajian literatur mengenai penelitian lain yang relevan dengan penelitian ini.

BAB III. METODOLOGI PENELITIAN

Pada bab ini membahas mengenai tahapan yang akan dilaksanakan pada penelitian. Seperti pengumpulan data, analisis data, serta perancangan sistem yang akan dibangun. Masing-masing rencana tahapan penelitian dideskripsikan dengan rinci mengacu pada suatu kerangka kerja.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Pada bab ini membahas mengenai analisis dan perancangan perangkat lunak yang akan digunakan sebagai alat penelitian. Dimulai dari pengumpulan dan analisa kebutuhan, rancangan dan konstruksi perangkat

lunak serta pengujian untuk memastikan semua kebutuhan pengembangan perangkat lunak sesuai dengan dengan kebutuhan.

BAB V. HASIL DAN ANALISA PENELITIAN

Pada bab ini diuraikan hasil pengujian berdasarkan langkah-langkah yang telah direncanakan. Tabel hasil pengujian serta analisisnya disajikan sebagai basis dari kesimpulan yang akan diambil dalam penelitian ini.

BAB VI. KESIMPULAN DAN SARAN

Pada bab ini berisi kesimpulan dari semua uraian-uraian pada bab-bab sebelumnya dan juga saran yang diberikan berdasarkan hasil dari penelitian.

1.8 Kesimpulan

Pada penelitian ini akan dijelaskan bagaimana cara melakukan peringkasan teks dengan metode *K-Means*. Menjelaskan mekanisme sistem dalam melakukan peringkasan teks serta mengetahui banyak keluaran yang dihasilkan oleh sistem.

Daftar Pustaka

- Agrawal, A., & Gupta, U. (2014). Extraction based approach for text summarization using k-means clustering. *International Journal of Scientific and Research Publications*, 1-4.
- Akter, S., Asa, A. S., Uddin, M. P., Hossain, M. D., Roy, S. K., & Afjal, M. I. (2017). An Extractive Text Summarization Technique for Bengali Document(s) using K-means Clustering Algorithm. *IEEE*.
- Carolina, I., & Supriyatna, A. (2019). Penerapan Metode Extreme Programming Dalam Perancangan Aplikasi Perhitungan Kuota SKS Mengajar Dosen. *Jurnal IKRA-ITH Informatika Vol 3 No 1*, 106-113.
- Celebi, M. E., A., H. K., & Vela, P. A. (2012). A Comparative Study of Efficient Initialization Methods for the K-Means Clustering Algorithm. *Expert Systems with Applications*, 200-210.
- Forgy, E. W. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics* 21, 768-769.
- Javadi, S., Mohammadi, K., Shahdany, M. H., & Neshat, A. (2017). Classification of aquifer vulnerability using K-means cluster analysis. *Journal of Hydrology*, 27-37.
- Jones, K. S. (1972). A statistical interpretation of term specificity and its. *Journal of Documentation*.
- Khotimah, B. K., Irhamni, F., & Sundarwati, T. (2016). A GENETIC ALGORITHM FOR OPTIMIZED INITIAL CENTERS K-MEANS

CLUSTERING IN SMEs. *Journal of Theoretical and Applied Information Technology*, 23-30.

Librian, A. (2021, Januari 18). *damzaky/sastrawijs: Indonesian language stemmer. Javascript port of PHP Sastrawi project*. Retrieved from Github: <https://github.com/damzaky/sastrawijs>

Librian, A. (2021, Januari 18). *sastrawijs*. Retrieved from npm: <https://www.npmjs.com/package/sastrawijs>

Mallick, C., Das, A. K., Dutta, M., Das, A. K., & Sarkar, A. (2019). Graph-Based Text Summarization Using Modified TextRank. *Soft Computing in Data Analytics Advances in Intelligent Systems and Computing*, 137-146.

Mujilahwati, S. (2016). Pre-Processing Text Mining pada Data Twitter. *Seminar Nasional Teknologi Informasi dan Komunikasi 2016*, 49-56.

Nazief, B., Adriani, M., Asian, J., Tahaghoghi, S. M., & Williams, H. E. (2005). Stemming Indonesian. *Conferences in Research and Practice in Information Technology Series*, 307-314.

Pandya, V. (2019). AUTOMATIC TEXT SUMMARIZATION OF LEGAL CASES: A HYBRID APPROACH. *arXiv preprint*.

Prathima, M. R., & Divakar, H. R. (2018). Automatic Extractive Text Summarization Using K-Means Clustering. *International Journal of Computer Sciences and Engineering* , 782-787.