

**RANCANG BANGUN APLIKASI BACA BERITA ONLINE DENGAN
PERINGKAS TEKS OTOMATIS MENGGUNAKAN ALGORITMA
TEXTRANK**

LAPORAN TUGAS AKHIR

Laporan ini disusun untuk memenuhi salah satu syarat memperoleh gelar Sarjana
Strata 1 (S1) Program Studi Teknik Informatika Fakultas Teknologi Industri
Universitas Islam Sultan Agung Semarang



DISUSUN OLEH : RYAN FAJAR KHAQIQI

NIM 32601400839

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ISLAM SULTAN AGUNG
SEMARANG NOVEMBER 2021**

**DESIGN AND BUILD ONLINE NEWS READING APPLICATION WITH
AUTOMATIC TEXT SUMMARY USING TEXTRANK ALGORITHM**

*Proposed to complete the requirement to obtain a bachelor's degree (S-1) at
Informatics Engineering Departement of Industrial Technology faculty Sultan
Agung Islamic University*



RYAN FAJAR KHAQIQI

NIM 32601400839

MAJORING OF INFORMATICS ENGINEERING

INDUSTRIAL TECHNOLOGY FACULTY

SULTAN AGUNG ISLAMIC UNIVERSITY

SEMARANG JULY 2021

LEMBAR PENGESAHAN PEMBIMBING

Laporan Tugas Akhir dengan judul “**RANCANG BANGUN APLIKASI BACA BERITA ONLINE DENGAN PERINGKAS TEKS OTOMATIS MENGGUNAKAN ALGORITMA TEXTRANK**” ini disusun oleh:

Nama : Ryan Fajar Khaqiqi

NIM : 32601400839

Program Studi : Teknik Informatika

Telah disahkan oleh dosen pembimbing pada:


Hari : Selasa

Tanggal : 30 November 2021

Mengesahkan,

Pembimbing I

Pembimbing II


M. Qomaruddin, ST.MSc, Ph.D


Bagus Satrio WP, S.Kom M.Cs.

NIDN.0631057101

NIDN. 1027118801

Ketua Progam Studi Teknik Informatika

Fakultas Teknologi Industri

Universitas Islam Sultan Agung


Andi Kianiyah, ST, M. Kom.

NIDN. 0609108802

LEMBAR PENGESAHAN PENGUJI

Laporan tugas akhir dengan judul **“RANCANG BANGUN APLIKASI BACA BERITA ONLINE DENGAN PERINGKAS TEKS OTOMATIS MENGGUNAKAN ALGORITMA TEXTRANK”** ini telah dipertahankan di depan dosen penguji Tugas Akhir pada :

Hari : Selasa

Tanggal : 30 November 2021

TIM PENGUJI

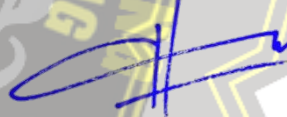
Anggota I

 20/12/2021

Badie'ah, ST, M. Kom

NIDN. 0619018701

Anggota II



Dedy Kurniadi, ST, M. Kom.

NIDN. 0622058802

Ketua Penguji



Andi Riansyah, ST, M. Kom.

NIDN. 0609108802

SURAT PERNYATAAN KEASLIAN TUGAS AKHIR

SURAT PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan dibawah ini :

Nama : Ryan Fajar Khaqiqi

NIM : 32601400839

Judul Tugas Akhir : RANCANG BANGUN APLIKASI BACA BERITA
ONLINE DENGAN PERINGKAS TEKS OTOMATIS
MENGUNAKAN ALGORITMA TEXTRANK

Dengan bahwa ini saya menyatakan bahwa judul dan isi Tugas Akhir yang saya buat dalam rangka menyelesaikan Pendidikan Strata Satu (S1) Teknik Informatika tersebut adalah asli dan belum pernah diangkat, ditulis ataupun dipublikasikan oleh siapapun baik keseluruhan maupun sebagian, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka, dan apabila di kemudian hari ternyata terbukti bahwa judul Tugas Akhir tersebut pernah diangkat, ditulis ataupun dipublikasikan, maka saya bersedia dikenakan sanksi akademis. Demikian surat pernyataan ini saya buat dengan sadar dan penuh tanggung jawab.

Semarang, November 2021

Yang Menyatakan,



Ryan Fajar Khaqiqi

PERNYATAAN PUBLIKASI KARYA ILMIAH

PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Saya yang bertanda tangan dibawah ini :

Nama : Ryan Fajar Khaqiqi
NIM : 32601400839
Program Studi : Teknik Informatika
Fakultas : Teknologi industri
Alamat Asal : Ds. Kebonsari, Kec. Rowosari Kab. Kendal

Dengan ini menyatakan Karya Ilmiah berupa Tugas akhir dengan Judul :
RANCANG BANGUN APLIKASI BACA BERITA ONLINE DENGAN
PERINGKAS TEKS OTOMATIS MENGGUNAKAN ALGORITMA TEXTRANK

Menyetujui menjadi hak milik Universitas Islam Sultan Agung serta memberikan Hak bebas Royalti Non-Eksklusif untuk disimpan, dialihmediakan, dikelola dan pangkalan data dan dipublikasikan di internet dan media lain untuk kepentingan akademis selama tetap menyahtumkan nama penulis sebagai pemilik hak cipta. Pernyataan ini saya buat dengan sungguh-sungguh. Apabila dikemudian hari terbukti ada pelanggaran Hak Cipta/Plagiarisme dalam karya ilmiah ini, maka segala bentuk tuntutan hukum yang timbul akan saya tanggung secara pribadi tanpa melibatkan Universitas Islam Sultan agung.

Semarang, November 2021



Ryan Fajar Khaqiqi

KATA PENGANTAR

Puji syukur senantiasa penulis panjatkan atas kehadiran Allah SWT yang telah memberikan rahmat dan inayah-Nya, sehingga dapat menyelesaikan Tugas Akhir dengan judul **“RANCANG BANGUN APLIKASI BACA BERITA ONLINE DENGAN PERINGKAS TEKS OTOMATIS MENGGUNAKAN ALGORITMA TEXTRANK”** ini untuk memenuhi salah satu syarat menyelesaikan studi serta dalam rangka memperoleh gelar sarjana (S-1).

Tugas Akhir ini disusun dan dibuat dengan adanya bantuan dari berbagai pihak, materi maupun teknis, oleh karena itu saya selaku penulis mengucapkan terima kasih kepada :

1. Kedua orang tua yang selalu memberikan motivasi, semangat, doa serta dukungan kepada penulis.
2. Bapak M.Qomaruddin, ST.MSc, Ph.D selaku Dosen Pembimbing I dan Bapak Bagus Satrio WP, S.Kom, M.Cs selaku Dosen Pembimbing II yang telah meluangkan waktu untuk membimbing serta memberikan ilmu kepada penulis, sehingga penulis dapat menyelesaikan tugas akhir dengan baik.
3. Para Dosen FTI Unissula yang telah memberikan banyak ilmu yang bermanfaat.
4. Saudara dan teman serta kawan-kawan teknik informatika angkatan 2014 yang selalu memberikan semangat kepada penulis.

Dengan segala kerendahan hati, penulis menyadari masih banyak terdapat banyak kekurangan – kekurangan dari segi kualitas atau kuantitas maupun dari ilmu pengetahuan dalam penyusunan laporan, sehingga penulis mengharapkan adanya saran dan kritikan yang bersifat membangun laporan ini di masa mendatang.

Semarang, November 2021

Penulis

DAFTAR ISI

RANCANG BANGUN APLIKASI BACA BERITA ONLINE DENGAN PERINGKAS TEKS OTOMATIS MENGGUNAKAN ALGORITMA TEXTRANK	1
DESIGN AND BUILD ONLINE NEWS READING APPLICATION WITH AUTOMATIC TEXT SUMMARY USING TEXTRANK ALGORITHM	2
LEMBAR PENGESAHAN PEMBIMBING	i
LEMBAR PENGESAHAN PENGUJI	ii
Surat Pernyataan Keaslian Tugas Akhir	iii
Pernyataan Publikasi Karya Ilmiah	iv
Kata Pengantar	v
Daftar Isi	vi
Daftar Tabel	ix
Daftar Gambar	x
Abstrak	xii
Bab 1 Pendahuluan	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Batasan Masalah	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
Bab 2 Tinjauan Pustaka dan Dasar Teori	5
2.1 Tinjauan Pustaka	5

2.2	Dasar Teori	6
2.2.1	TEXTRANK	6
2.2.2	Peringkas Teks Otomatis.....	8
2.2.3	Berita.....	10
2.2.4	Information Retrieval	12
2.2.5	Natural Language Processing.....	14
2.2.6	Text Mining.....	15
2.2.7	Text Preprocessing	15
2.2.8	Kata	16
2.2.9	Kalimat.....	16
2.2.10	Unified Modeling Language (UML).....	17
2.2.11	Model Pengembangan Sistem	20
2.2.12	Waterfall.....	20
2.2.13	Paragraf	22
2.2.14	PHP	22
2.2.15	MySQL.....	24
2.2.16	Hypertext Markup Language (HTML)	25
2.2.17	Cascading Style Sheet (CSS)	25
2.2.18	Javascript.....	26
Bab 3	Metode Penelitian	27
3.1	Metode Pengumpulan Data	27
3.1.1	Observasi.....	27
3.1.2	Studi Pustaka.....	27
3.2	Metode Pengembangan Sistem	27
3.3	Model yang Diusulkan	30

3.3.1	Data Teks Berita.....	30
3.3.2	Algoritma TEXRANK	31
3.4	Analisa Aplikasi	37
3.4.1	Deskripsi Umum Aplikasi.....	37
3.4.2	Analisa Kebutuhan	38
3.5	Perancangan Aplikasi	38
3.5.1	Use Case Diagram.....	38
3.5.2	Activity Diagram.....	43
3.5.3	Sequence Diagram	44
3.6	Perancangan Antarmuka.....	45
Bab 4	Hasil dan Analisis Penelitian.....	48
4.1.1	Pengujian Aplikasi	48
4.1.2	Pengujian Fungsional.....	53
4.1.3	Pengujian Ringkasan.....	54
Bab 5	Kesimpulan dan Saran	56
5.1	Kesimpulan.....	56
5.2	Saran.....	56
Daftar Pustaka	57

DAFTAR TABEL

Tabel 2.1 Simbol <i>Use Case</i>	18
Tabel 2.2 Simbol <i>Activity Diagram</i>	18
Tabel 3.1: Teks Splitting	31
Tabel 3.2: Teks Cleansing	32
Tabel 3.3: Hasil Tokenizing	33
Tabel 3.4: Stopword Removal	33
Tabel 3.5: Hasil Teks Preprocessing	34
Tabel 3.6: Text Similiarity	35
Tabel 3.7: Matrix Text Similiarity	36
Tabel 3.8: Perhitungan Bobot Kalimat	36
Tabel 3.9: Hasil Urutan Pembobotan Kalimat	37
Tabel 4.6: Pengujian Fungsional Admin	53
Tabel 4.7: Pengujian Fungsional User	53
Tabel 4.8: Pengujian Ringkasan	54

DAFTAR GAMBAR

Gambar 2.1 Model Waterfall (Tujni & Hutrianto, 2020).....	21
Gambar 2.2: alur PHP	23
Gambar 3.1 Model Waterfall	28
Gambar 3.2: Model yang Diusulkan	30
Gambar 3.3: Use Case Diagram Aplikasi	39
Tabel 3.1: Skenario Membuka Aplikasi.....	39
Tabel 3.2: Skenario Baca Berita Ringkasan.....	40
Tabel 3.3: Skenario Baca Berita Asli.....	40
Tabel 3.4: Skenario Melihat Halaman About	40
Tabel 3.5: Skenario Login.....	41
Tabel 3.6: Skenario Mengelola Kategori	41
Tabel 3.7: Skenario Mengelola Berita	41
Tabel 3.8: Skenario Ringkasan Berita TEXTRANK.....	41
Gambar 3.4: Activity Diagram Aplikasi	43
Gambar 3.5: Sequence Diagram Ringkasan Berita.....	44
Gambar 3.6: Sequence Diagram Berita Asli	45
Gambar 3.7: Sequence Diagram Berita Asli	45
Gambar 3.8: Perancangan Antarmuka	46
Gambar 4.1: Login Admin	48
Gambar 4.2: Halaman Kategori	49
Gambar 4.3: Halaman Data Berita.....	49
Gambar 4.4: Halaman Input Berita	50

Gambar 4.5: Halaman Hasil Ringkasan	51
Gambar 4.6: Halaman Beranda User	51
Gambar 4.7: Halaman Berita Ringkasan.....	52
Gambar 4.8: Halaman Detail Berita Asli	52



ABSTRAK

Dalam berita media online terdapat teks highlight yang menunjukkan titik berat berita sebelum dipilih untuk dibaca namun banyak media berita yang hanya menampilkan highlight hanya pada awal dari berita. Pengambilan teks highlight pada awal berita tidak dapat menampilkan informasi berita ringkas yang mewakili isi berita. Maka dibutuhkan ringkasan yang merupakan isi dari keseluruhan artikel secara ringkas. Tujuan dari penelitian ini adalah Membuat aplikasi peringkasan teks otomatis pada data teks berita seputar Semarang menggunakan algoritma TEXTRANK. Metode yang diusulkan yaitu TEXTRANK telah berhasil diimplementasikan untuk peringkasan teks otomatis pada aplikasi berita online sekitar kota Semarang. Penerapan metode TEXTRANK mempunyai performa ringkasan yang mampu meringkas teks hingga 27 persen. Aplikasi yang dibuat telah diuji menggunakan pengujian blackbox yang menguji keseluruhan fungsi, hasil dari pengujian menunjukkan fungsi-fungsi yang terdapat pada aplikasi berjalan dengan baik dan memiliki output yang sesuai.

Kata kunci: peringkasan teks otomatis, textrank, berita

ABSTRACT

In online media news, there is a highlight text that shows the focus of the news before it is chosen to be read, but many news media that only show highlights are only at the beginning of the news. Retrieving text highlights at the beginning of the news cannot display concise news information that represents the contents of the news. Then a summary is needed which is the content of the whole article in a concise manner. The purpose of this study is to make an automatic text summarization application on news text data around Semarang using the TEXTRANK algorithm. The proposed method, namely TEXTRANK has been successfully implemented for automatic text summarizing on online news applications around the city of Semarang. The application of the TEXTRANK method has a summary performance that is able to summarize text up to 27 percent. Applications that have been made have been tested using blackbox testing which tests the entire function, the results of the tests show that the functions contained in the application run well and have the appropriate output.

Keywords: automatic text summarization, textrank, news

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Seiring berkembangnya teknologi informasi mudah di akses melalui portal-portal berita online. Dengan teknologi internet kemudahan mendapatkan informasi melalui portal berita online. Media online sebagai portal berita seperti detik, kompas, okezone, liputan6 dan lain-lain sangatlah diminati oleh masyarakat karena kemudahannya mendapatkan informasi secara lokal maupun global sampai luar negeri dengan cepat, selain itu juga tidak memerlukan biaya yang mahal. Banyaknya minat dan pembaca berita secara online menjadikan muncul-muncul portal berita online yang jumlahnya sangat banyak (Kurniawan, 2012).

Untuk mencari informasi pada suatu berita masyarakat sebagai pembaca memerlukan waktu yang cukup lama karena membaca pada keseluruhan konten yang berupa teks (Widiantoro, 2014). Bagi yang memiliki waktu singkat namun ingin memperoleh informasi hal tersebut tentunya akan mengganggu dan perolehan informasi tidak maksimal. Permasalahan tersebut dapat diatasi jika terdapat ringkasan atau rangkuman dari berita yang berupa teks.

Ringkasan atau rangkuman dibutuhkan untuk mengetahui informasi yang ada pada suatu dokumen secara cepat dan ringkas. Ringkasan didefinisikan sebagai penyajian bacaan dalam bentuk singkat dengan mempertahankan urutan isi dan sudut pandang dengan mempersingkat bacaan dengan mengambil inti bacaan tanpa mengubah struktur wacana. Ringkasan suatu berita sangatlah penting karena menggunakan ringkasan pembaca akan mendapatkan kemudahan untuk mengetahui inti dari informasi yang ada pada berita karena tidak memakan banyak waktu karena pengguna cenderung ingin mengetahui informasi dengan cepat (Mustaqfiri, 2012).

Peringkasan teks otomatis (*automatic text summarization* atau ATS) adalah teknik pembuatan ringkasan dari sebuah teks secara otomatis dengan

memanfaatkan metode tertentu untuk menghasilkan informasi yang paling penting dari dokumen aslinya. Peringkat diberi masukan berupa teks, kemudian melakukan peringkasan, dan menghasilkan keluaran berupa teks yang lebih singkat dari teks asli. Hasil peringkasan mengandung poin-poin penting atau informasi utama dari teks sumber. Salah satu algoritma yang dapat dijadikan peringkasan dokumen otomatis adalah TEXTRANK (Abdurrohman, 2018). Algoritma TEXTRANK menghasilkan ekstraksi kata sampai kalimat dengan hasil akhir bobot setiap kalimat. Kalimat yang memiliki bobot paling tinggi digunakan sebagai inti dari keseluruhan paragraf.

Dari penjelasan diatas maka diusulkan penelitian yang menerapkan algoritma TEXTRANK untuk peringkasan teks otomatis pada berita online yang diterapkan pada aplikasi mobile web. Penelitian disusun dengan judul “Rancang Bangun Aplikasi Baca Berita Online Dengan Peringkasan Teks Otomatis Menggunakan Algoritma Textrank”.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah di atas, maka dapat ditemukan masalah yang akan diselesaikan dalam penelitian ini adalah Bagaimana mengimplementasikan teknik peringkasan teks otomatis (automatic text summarization atau ATS) dengan algoritma TEXTRANK untuk membuat peringkasan teks otomatis pada aplikasi baca berita online berbasis mobile web..

1.3 Tujuan Penelitian

Tujuan penelitian ini adalah

1. Mengimplementasikan algoritma TEXTRANK untuk peringkasan dokumen otomatis dengan data berita online
2. Merancang dan membangun aplikasi portal berita berbasis website dengan menerapkan peringkasan teks otomatis menggunakan algoritma TEXTRANK.

1.4 Batasan Masalah

Agar tidak terjadi perluasan dalam pembahasan masalah, maka diperlukan adanya pembatasan masalah untuk dapat memberikan gambaran yang lebih terarah dan fokus. Adapun batasan masalah yang diambil yaitu:

1. Data yang digunakan mengambil secara langsung dari portal berita online detik.com
2. Hasil akhir dari penelitian merupakan aplikasi berbasis mobile web
3. Aplikasi dibuat menggunakan html, css dan javascript. Bahasa pemrograman yang digunakan adalah PHP (Hypertext Preprocessor)

1.5 Manfaat Penelitian

1. Manfaat Teoritis

Penggunaan Metode Automatic Text Summarization dengan algoritma TEXTRANK bermanfaat dalam meringkas teks secara otomatis. Dengan demikian penelitian ini bermanfaat sebagai referensi terhadap penelitian selanjutnya dalam pemanfaatan teknik peringkas teks otomatis.

2. Manfaat Praktis

Berkontribusi dalam ilmu *text mining* Bahasa Indonesia yang dapat diterapkan untuk membuat peringkas teks secara otomatis. Dengan adanya peringkas teks otomatis dapat menjadikan highlight berita menjadi lebih baik dan proses mendapatkan informasi dalam membaca berita menjadi lebih cepat.

1.6 Sistematika Penulisan

BAB 1: PENDAHULUAN Bab ini menjelaskan mengenai latar belakang pemilihan judul tugas akhir “Rancang Bangun Aplikasi Baca Berita Online Dengan Peringkas Teks Otomatis Menggunakan Algoritma Textrank”, rumusan masalah, batasan masalah, tujuan penelitian, metodologi penelitian, dan sistematika penulisan.

BAB 2: TINJAUAN PUSTAKA DAN DASAR TEORI. Bab ini memuat

dasar teori yang berfungsi sebagai sumber atau alat dalam memahami permasalahan yang berkaitan dengan sistem, Algoritma TEXTRANK dan Peringkat Teks Otomatis serta teori yang berhubungan dan diperlukan dalam pembuatan aplikasi ini.

BAB 3: METODE PENELITIAN Bab ini

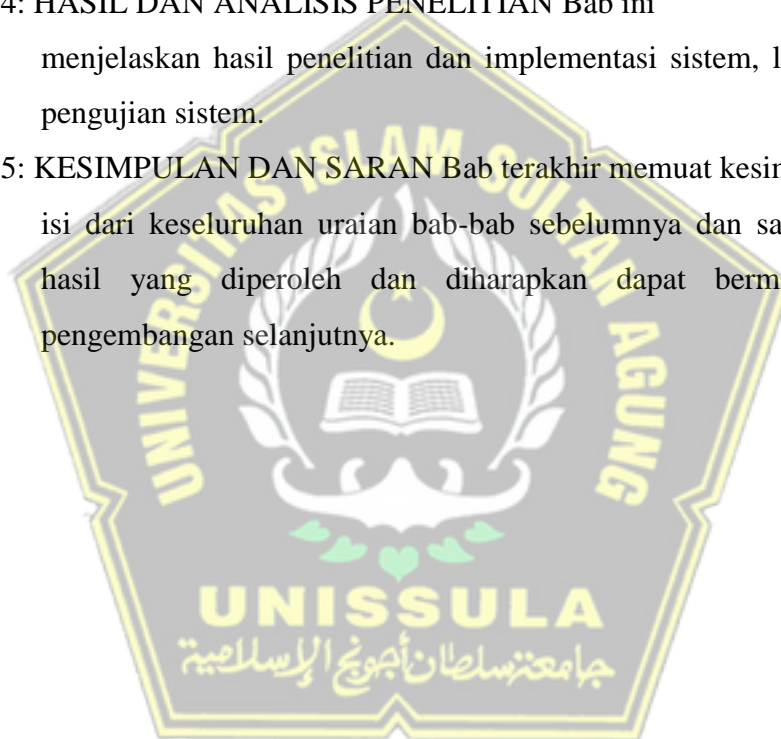
menjabarkan tentang tujuan dari perancangan sistem, peringkat teks otomatis menggunakan Algoritma TEXTRANK

BAB 4: HASIL DAN ANALISIS PENELITIAN Bab ini

menjelaskan hasil penelitian dan implementasi sistem, lalu dilakukan pengujian sistem.

BAB 5: KESIMPULAN DAN SARAN Bab terakhir memuat kesimpulan

isi dari keseluruhan uraian bab-bab sebelumnya dan saran-saran dari hasil yang diperoleh dan diharapkan dapat bermanfaat dalam pengembangan selanjutnya.



BAB 2

TINJUAN PUSTAKA DAN DASAR TEORI

2.1 Tinjauan Pustaka

Terdapat beberapa penelitian yang dijadikan acuan pada penelitian ini. Penelitian sebagai acuan merupakan penelitian yang memiliki kesamaan baik tema, metode maupun data yang digunakan. Penelitian terkait diantaranya:

Sebelumnya terdapat penelitian yang dijadikan acuan dan referensi pada penelitian ini. Penelitian yang dijadikan tinjauan pustaka merupakan penelitian yang terkait tentang objek penelitian atau metode yang digunakan. Penelitian dengan judul Rancang Bangun Sistem Peringkasan Teks Otomatis Artikel Berbahasa Indonesia Menggunakan Metode Tf-Idf (*Term Frequency-Inversed Document Frequency*) dilakukan oleh (Budianto, 2016) Paper ini membahas tentang peringkasan teks otomatis pada dokumen berbahasa Indonesia. Data yang digunakan merupakan artikel yang bersumber dari internet. Dalam media internet artikel merupakan suatu kebutuhan dan pengetahuan. Pada umumnya pembaca ingin mendapatkan rangkuman suatu artikel dengan cepat tanpa membaca isi dari keseluruhan dokumen teks sumber. Agar lebih mempermudah dalam mencari rangkuman diperlukan sistem peringkasan teks otomatis (*auto text summarization*) yang dapat menghasilkan rangkuman (*summary*). *Summary* merupakan hasil dari sebuah kata-kata, kalimat-kalimat dan paragraph yang panjangnya tidak lebih dari setengah teks sumber asli. Salah satu metode yang dapat digunakan dalam tujuan untuk peringkasan teks otomatis, dimulai dari pengelolaan dokumen teks dengan menggunakan metode text mining dan perangkuman dokumen artikel menggunakan algoritma *Term Frequency - Inversed Document Frequency* (TF-IDF). Tingkat akurasi dari algoritma TF-IDF dalam perangkuman dokumen artikel yaitu dengan melakukan hasil pengujian dibandingkan dengan ringkasan manual kepada 20 responden yang menghasilkan nilai rata-rata akurasi 70%.

Penelitian selanjutnya yang dijadikan tinjauan pustaka dengan judul Penerapan Algoritma TEXTRANK Untuk Automatic Summarization Pada

Dokumen Berbahasa Indonesia oleh (Christianti, et al., 2017). Pada paper ini penelitian membahas tentang peringkas teks otomatis pada dokumen berbahasa Indonesia. Data yang diambil bersumber dari media online. dalam penelitian ini Automatic Summarization dibangun menggunakan algoritma TextRank. TextRank adalah algoritma peringkat berbasis grafik untuk memproses teks. TextRank menghasilkan ekstraksi kalimat sebagai ringkasan. Perumusan dilakukan pada tahap berikut: preprocessing, hitung nilai kesamaan konten yang tumpang tindih, hitung nilai TextRank pada setiap kalimat, dan buat grafik. Hasilnya adalah teks ringkasan yang informatif. Automatic Summarization diuji dengan Q & A Evaluation yang diberikan kepada beberapa responden. Pengujian menunjukkan bahwa, algoritma ini mampu memberikan ringkasan dengan konten informatif hingga 82,48% untuk teks ringkasan 50% dan konten informatif 93,76% untuk teks ringkasan yang dirangkum 75%.

2.2 Dasar Teori

2.2.1 TEXTRANK

TextRank merupakan metode yang termasuk kedalam pendekatan unsupervised dan menggunakan pemodelan berbasis graf. Metode ini dikembangkan berdasarkan dari metode PageRank (Pragantha & Eris, 2017). Dasar dari model berbasis perankingan graf yaitu dengan mengimplementasikan tahapan “voting” pada setiap kata (vertex) dalam graf. Suatu vertex akan dianggap penting jika vertex tersebut di „vote’ lebih banyak dibandingkan dengan vertex lainnya. Nilai skor pada tiap vertex didalam graf ditentukan dari Persamaan berikut:

$$S(V_i) = (1 - d) + d \times \sum_{j \in In(V_i)} \frac{1}{|Out(V_j)|} S(V_j) \quad (1)$$

dimana nilai $S(V_i)$ sebagai nilai score vertex V_i , dengan nilai d sebagai damping factor yang biasanya di set dengan nilai 0.85.

Persamaan diatas dapat juga diterapkan untuk melakukan perankingan pada graf yang berbobot dengan menggunakan nilai bobot edge antar vertexnya sehingga dibuatlah perumusan baru seperti pada Persamaan dibawah.

$$WS(V_i) = (1 - d) + d \times \sum_{V_j \in In(V_i)} \frac{1}{\sum_{V_k \in Out(V_j)} w_{jk}} WS(V_j) \quad (2)$$

dimana dalam rumus diatas terdapat w_{ji} yaitu sebagai bobot edge dari vertex V_j yang berelasi dengan vertex V_i . w_{ji} juga disebut sebagai bobot frekuensi kemunculan vertex V_j dan V_i yang saling co-occurrence di dalam ukuran window dengan maksimum L_{kata} . TextRank menggunakan hubungan kemunculan bersama satu kata dengan kata yang lainnya (co-occurrence relation) dalam window maksimum L kata tersebut dimana ukuran L bisa di set antara 2 – 10 kata.

Pada metode textrank ini kandidat kata kunci didapat dengan filter kata menggunakan part-of-speech tagging. Setiap kata yang akan dijadikan vertex dalam graf akan difilter berdasarkan filter kelas katanya seperti kata benda, gabungan kata benda dan kata kerja, atau filter semua jenis kata. Dari hasil penelitian sebelumnya hasil terbaik didapatkan dari filter kata berupa kata benda dan kata sifat saja. Untuk metode TextRank ini frase atau kandidat kata kunci berbentuk multiword di proses pada tahapan post-processing untuk menghindari besarnya hubungan relasi didalam graf. Berdasarkan hal itu cara ini tidak bisa menjamin semua kata kunci yang dihasilkan merupakan kata-kata yang sudah benar dalam satuan leksikal bahasanya dan tidak semua kata-kata yang ada didalam kombinasi kata kunci bisa didapat dari ranking teratas. Algoritma textrank didalam konsep jaringan frase (phrase network) yaitu dengan meranking langsung kandidat kata kunci berbentuk frase ataupun multiword. Biasanya vertex dalam graph pada algoritma textrank sebelumnya hanya berupa satu kata saja, tetapi dengan jaringan frase ini maka akan diperluas hubungan co-occurrence kata menjadi berupa kata dan frase. Kandidat kata kunci atau vertex yang digunakan pada penelitian tersebut didapat dengan memanfaatkan metode Document Frequency Accessor Variety (DF-AV) dikarenakan pada bahasa Cina pemfilteran menggunakan POS tagging tidak menunjukkan hasil yang cukup baik (Nugraha, 2016).

Umumnya algoritma perankingan berbasis graf memiliki beberapa tahapan seperti berikut:

- Identifikasi satuan unit teks yang akan digunakan dan tambahkan setiap katanya sebagai vertex dalam graf
- Identifikasi hubungan yang terbentuk untuk setiap unit teks untuk menggambarkan edge dalam graf. Edge dapat berupa berarah atau tidak berarah (directed or undirected) ,berbobot atau tidak berbobot (weighted or unweighted).
- Iterasi algoritma perankingan berbasis graf tersebut hingga konvergen (Convergence). Konvergen didapatkan ketika rata-rata nilai error untuk setiap vertex di dalam graf berada dibawah nilai threshold yang diberikan.

Urutkan vertex yang terbentuk berdasarkan skor bobot akhirnya. Gunakan nilai skor bobot tiap kata tersebut sebagai urutan ranking.

2.2.2 Peringkasan Teks Otomatis

Sering kali kita membutuhkan ringkasan dari sebuah bacaan untuk mendapatkan secara ringkas dan cepat isi dari bacaan. Konsep sederhana dari ringkasan adalah mengambil bagian penting yang menggambarkan keseluruhan isi dari dokumen asal. Menurut Mani dan Maybury ringkasan adalah mengambil isi yang paling penting dari sumber informasi yang kemudian menyajikan kembali dalam bentuk yang lebih ringkas bagi penggunaanya. Dalam Hovy, summary atau ringkasan didefinisikan sebagai sebuah teks yang dihasilkan dari satu atau lebih teks, mengandung informasi dari teks asli dan panjangnya tidak lebih dari setengah teks asli (Moreno & Manuel, 2014).

Peringkasan teks otomatis (*automatic text summarization*) adalah pembuatan versi yang lebih singkat dari sebuah teks dengan memanfaatkan aplikasi yang dijalankan pada komputer. Hasil peringkasan ini mengandung poin-poin penting dari teks asli.

Berdasar teknik pembuatan, suatu ringkasan diambil dari bagian terpenting dari teks aslinya, terdapat 2 tipe yaitu:

1. Abstraktif

Tipe peringkasan abstraktif menghasilkan sebuah interpretasi terhadap teks aslinya. Dimana sebuah kalimat akan ditransformasikan menjadi kalimat yang lebih singkat dan kalimat baru yang tidak terdapat dalam dokumen yang asli atau dengan kalimat yang berbeda.

2. Ekstraktif

Tipe peringkasan ekstraktif menghasilkan suatu ringkasan dengan memilih sebagian dari kalimat yang ada dalam dokumen asli. Metode ini menggunakan metode statistical, linguistical dan heuristic atau kombinasi dari semuanya dalam menetapkan ringkasan suatu teks.

Berdasarkan teori, hasil ringkasan ekstraktif lebih baik dibandingkan dengan ringkasan abstraktif. Hal ini dikarenakan peringkasan abstraktif, seperti representasi semantik, inferens dan pembangun natural language relatif lebih sulit, dibandingkan pendekatan data driven, seperti ekstraksi kalimat. Sehingga kebanyakan penelitian dilakukan menggunakan metode ekstraktif .

Sedangkan model peringkasan teks otomatis ada dua yaitu ringkasan yang umum (*generic summary*) merupakan perwakilan dari teks asli yang mencoba untuk mempresentasikan semua fitur penting dari sebuah teks asal. Mengikuti pendekatan bottom up (*information retrieval*) dan yang kedua ringkasan berpusat pada pemakai (*query driven*) yaitu peringkasan bersandar pada spesifikasi kebutuhan informasi pemakai, seperti topik atau query dan mengikuti pendekatan top down (*information extraction*).

Tujuan dari peringkasan teks (*teks summarization*) dapat dikategorikan berdasarkan maksud, fokus dan cakupannya, sebagai berikut:

1. Informatif

Informatif, ringkasan ini menyatakan informasi - informasi penting yang terdapat pada dokumen asal.

2. Indikatif

Indikatif, tujuan dari ringkasan ini adalah untuk dijadikan sebuah referensi, yang membantu pembaca untuk mengetahui isi dari teks daripada membaca keseluruhan teks yang ada. Ringkasan ini meliputi topik kunci dari teks asal.

3. Evaluatif

Evaluatif, atau ringkasan yang melibatkan pembuatan sebuah pertimbangan pada teks asal, seperti suatu tinjauan ulang atau opini.

4. *User-focused (query-relevant)*

User-focused, ringkasan yang dibuat berdasarkan topik yang dipilih oleh user, sering merupakan jawaban dari query yang dimiliki oleh user.

5. *Generic*

Generic, disebut juga *author-focused*, sifatnya lebih umum dan berdasarkan pada teks aslinya.

6. Dokumen tunggal (*single document*)

Dokumen tunggal, ringkasan merupakan ringkasan dari satu dokumen.

7. Banyak dokumen (*multi document*)

Banyak dokumen, ringkasan merupakan hasil ringkasan dari banyak dokumen

2.2.3 Berita

Kata "berita" berasal dari bahasa sansekerta yaitu dari kata "vrit" yang sebenarnya berarti "terjadi" atau "ada". Berita (news) adalah laporan mengenai suatu peristiwa atau kejadian yang terbaru (aktual); laporan mengenai fakta-fakta yang aktual, menarik perhatian, dinilai penting, atau luar biasa (Barus, 2016).

Berita adalah informasi baru tentang kejadian yang baru, penting, dan bermakna, yang berpengaruh pada para pendengarnya serta relevan dan layak dinikmati.

2.2.3.1 Nilai Berita

Dalam menulis berita, ada beberapa hal yang perlu diperhatikan terkait nilai berita itu sendiri. Ada beberapa nilai berita yang dapat dikelompokkan sebagai acuan dalam sebuah penulisan. Beberapa nilai berita tersebut adalah sebagai berikut :

1. *Magnitude* (pengaruh) artinya seberapa luas pengaruh suatu berita terhadap khalayak
2. *Significant* (Arti) artinya seberapa penting arti dari suatu kejadian atau peristiwa.
3. *Actuality* (Aktualitas) artinya seberapa besar tingkat aktualitas suatu kejadian atau peristiwa.
4. *Proximity* (Kedekatan) artinya bertia lokal lebih pas diberitakan di daerah bersangkutan.
5. *Prominence* (Keakraban) artinya akrabnya suatu peristiwa terhadap khalayak.
6. *Surprise* (Kejutan).
7. *Clarity* (Kejelasan) kejadian atau peristiwa.
8. Dampak (*Impact*) artinya berdampak apakah berita tersebut terhadap khalayak.
9. Konflik.
10. *Human Interest* artinya kemampuan suatu peristiwa menyentuh perasaan khalayak

2.2.3.2 Unsur Berita

Dalam penulisan berita kita harus memahami unsur dari suatu berita supaya memberi kemudahan kita dalam mendeskripsikan berita tersebut dan berita yang kita buat mudah untuk dipahami oleh khalayak ramai. Unsur-unsur berita tersebut adalah:

1. What (apa) artinya apa yang tengah terjadi. Peristiwa apa yang tengah terjadi.
2. Who (siapa) artinya siapa saja yang terlibat dalam peristiwa itu.
3. Where (dimana) artinya dimana lokasi terjadinya peristiwa itu.
4. When (kapan) artinya kapan peristiwa itu berlangsung.
5. Why (mengapa) artinya mengapa kejadian itu bisa terjadi.
6. How (bagaimana) artinya bagaimana kejadian itu bisa berlangsung.

2.2.3.3 Anatomi Berita

Seperti tubuh manusia, berita juga mempunyai bagian-bagian, diantaranya adalah sebagai berikut:

1. Judul atau Kepala Berita (*Headline*) Headline mewakili isi berita yang ingin disampaikan dan memiliki daya tarik yang kuat.
2. Baris Tanggal (*Dateline*) Dateline terdiri atas nama media massa, tempat kejadian dan tanggal kejadian. Tujuannya adalah untuk menunjukkan tempat kejadian dan inisial media.
3. Teras Berita (*Lead atau Intro*) Lead biasanya ditulis pada paragraf pertama sebuah berita. Lead merupakan unsur yang paling penting dari sebuah berita, yang menentukan apakah isi berita akan dibaca atau tidak.
4. Tubuh Berita (*Body*) *Body* isinya menceritakan peristiwa yang dilaporkan dengan bahasa yang singkat, padat, dan jelas baik yang sudah dikemukakan dalam teras maupun yang belum diungkapkan.

2.2.3.4 Geografi Berita

Berita ditinjau dari sudut geografi yang terbagi menjadi 3 yaitu:

1. Berita lokal
Berita tentang peristiwa yang terjadi pada area tertentu pada umumnya mencakup area administrasi desa atau kelurahan, kecamatan dan kabupaten atau kota.
2. Berita nasional
Berita tentang peristiwa yang terjadi pada Negara tertentu atau mencakup berita nasional.
3. Berita internasional
Berita tentang peristiwa antarnegara atau mencakup hubungan internasional.

2.2.4 Information Retrieval

Information Retrieval System atau Sistem Temu Balik Informasi merupakan bagian dari computer science tentang pengambilan informasi dari dokumen-dokumen yang

didasarkan pada isi dan konteks dari dokumen-dokumen itu sendiri. Menurut Gerald J. Kowalski di dalam bukunya “*Information Storage and Retrieval Systems Theory and Implementation*”, sistem temu balik informasi adalah suatu sistem yang mampu melakukan penyimpanan, pencarian, dan pemeliharaan informasi. Informasi dalam konteks ini dapat terdiri dari teks (termasuk data numerik dan tanggal), gambar, audio, video, dan objek multimedia lainnya (Zhai & Lafferty, 2017).

Tujuan dari sistem IR adalah memenuhi kebutuhan informasi pengguna dengan me-retrieve semua dokumen yang mungkin relevan, pada waktu yang sama me-retrieve sesedikit mungkin dokumen yang tidak relevan. Sistem IR yang baik memungkinkan pengguna menentukan secara cepat dan akurat apakah isi dari dokumen yang diterima memenuhi kebutuhannya. Agar representasi dokumen lebih baik, dokumen-dokumen dengan topik atau isi yang mirip dikelompokkan bersama-sama.

Model Information Retrieval adalah model yang digunakan untuk melakukan pencocokan antara term-term dari query dengan term-term dalam document collection, Model yang terdapat dalam Information retrieval terbagi dalam 3 model besar, yaitu

1. *Set-theoretic models*, model merepresentasikan dokumen sebagai himpunan kata atau frase. Contoh model ini ialah standard Boolean model dan *extended Boolean model*.
2. *Algebraic model*, model merepresentasikan dokumen dan query sebagai vektor atau matriks similarity antara vektor dokumen dan vektor query yang direpresentasikan sebagai sebuah nilai skalar. Contoh model ini ialah vektor space model (model ruang vektor) dan latent semantic indexing (LSI).
3. *Probabilistic model*, model memperlakukan proses pengambilan dokumen sebagai sebuah probabilistic inference. Contoh model ini ialah penerapan teorema bayes dalam model probabilistik.

2.2.5 Natural Language Processing

Natural Language Processing (NLP) adalah penerapan ilmu komputer, khususnya linguistik komputasional (*computational linguistics*), untuk mengkaji interaksi antara komputer dengan bahasa (alami) manusia. NLP berupaya memecahkan masalah untuk memahami bahasa alami manusia, dengan segala aturan gramatika dan semantiknya, dan mengubah bahasa tersebut menjadi representasi formal yang dapat diproses oleh computer (Collobert, 2011).

Berdasarkan dalam penerapannya, tujuan NLP untuk memahami bahasa manusia ini memiliki banyak tantangan, yang antara lain adalah sebagai berikut:

1. Penandaan kelas kata (*part-of-speech tagging*). Sulit untuk menandai kelas kata (kata benda, kata kerja, kata sifat, dsb.) suatu kata dalam teks karena pengelasan kata sangat bergantung kepada konteks penggunaannya.
2. Segmentasi teks (*text segmentation*). Penentuan segmentasi sulit dilakukan pada bahasa tulis yang tidak memiliki pembatas kata spesifik (misal: bahasa Mandarin, Jepang, dan Thailand) serta pada bahasa lisan yang kadang membaurkan bunyi antarkata.
3. Disambiguasi makna kata (*word sense disambiguation*). Banyak kata memiliki lebih dari satu makna, baik dalam bentuk homonim (makna berbeda dan tidak terkait, contohnya “bisa” dalam makna “dapat” dan “racun”) maupun polisemi (makna berbeda, namun terkait, mis. “ragu” dalam makna “bimbang” dan “sangsai”). Pembedaan makna hanya dapat dilakukan dengan melihat konteks penggunaan.
4. Ambiguitas sintaksis (*syntactic ambiguity*). Suatu bahasa memiliki berbagai kemungkinan struktur kalimat. Pemilihan struktur yang paling tepat biasanya membutuhkan gabungan informasi semantik dan kontekstual.
5. Masukan yang tak sempurna atau tak teratur (*imperfect or irregular input*). Aksentuasi dalam bahasa lisan serta kesalahan ejaan dan gramatikal dalam bahasa tulis menyulitkan pemrosesan bahasa alami.

6. Pertuturan (*speech act*). Struktur kalimat saja kadang tidak dapat dengan tepat menggambarkan maksud penutur atau penulis. Kadang gaya bahasa dan konteks menentukan maksud yang diinginkan.

2.2.6 Text Mining

Text mining adalah cara pengekstraksian informasi yang tidak diketahui sebelumnya dari suatu sumber data yang berbeda ke dalam bentuk baru dimana bentuk baru dan sumber data tersebut mempunyai keterkaitan. *Text Mining* merupakan salah satu bentuk eksplorasi dan analisis data teks yang bertujuan untuk mendapatkan pengetahuan baru baik itu melalui cara otomatis maupun semi otomatis (Weiss, 2010). Berdasarkan kutipan-kutipan diatas, penulis berpendapat bahwa Text mining merupakan proses penggalian informasi-informasi baru dan berguna dari data lama menjadi data baru dimana kedua data tersebut saling berhubungan.

2.2.7 Text Preprocessing

Dalam proses *text mining* tahap yang perlu dilakukan adalah *text preprocessing*. Tujuannya adalah mempersiapkan dan mengolah teks menjadi data yang dapat diproses pada tahapan berikutnya. Teks yang tidak terstruktur dan mempunyai banyak komponen yang tidak diperlukan diolah agar lebih terstruktur. Dalam text preprocessing biasanya terdapat tahap-tahap diantaranya adalah tokenization, stemming dan stopword removal (Laksana & Purwarianti, 2014).

Tokenization adalah tugas pemotongan urutan karakter dan sebuah set dokumen yang diberikan menjadi potongan-potongan kata atau karakter yang sesuai dengan kebutuhan sistem. Potongan-potongan tersebut dikenal dengan istilah token.

Stemming merupakan salah satu proses dari mengubah token yang berimbuhan menjadi kata dasar, dengan menghilangkan semua imbuhan yang ada pada token tersebut. Pentingnya stemming dalam proses pembuatan sistem adalah untuk menghilangkan imbuhan pada awalan dan akhiran. Berdasarkan hasil proses tersebut, akan didapatkan sebuah informasi mengenai banyaknya fitur yang muncul dalam sebuah dokumen.

Stopwords dapat diartikan sebagai menghilangkan karakter, tanda baca, serta kata-kata umum yang tidak memiliki makna atau informasi yang dibutuhkan. *Stopwords* umumnya digunakan dalam pengambilan informasi salah satu contohnya adalah mesin pencari Google. Pengurangan ukuran indeks dalam teks dengan penghilangan beberapa kata kerja, kata sifat, dan kata keterangan lainnya dapat dimasukkan ke dalam daftar stopwords.

2.2.8 Kata

Kata adalah kesatuan terkecil yang diperoleh sesudah kalimat dibagi atas bagianbagiannya dan mengandung suatu ide (Sutami, 2020). Kategori kata berdasarkan sintaksisnya terdiri dari lima kata, yaitu:

1. Kata Benda (Nomina)

Kata benda adalah kata yang mengacu pada manusia, binatang, benda dan konsep atau pengertian.

2. Kata Kerja (Verba)

Kata kerja adalah kata yang menyatakan tindakan.

3. Kata Sifat (Adjektiva)

Kata sifat adalah kata yang memberi keterangan yang lebih khusus tentang sesuatu yang dinyatakan oleh nomina dalam kalimat.

4. Kata Keterangan (adverbia)

Kata keterangan adalah kategori yang dapat mendampingi adjektiva, numeralia atau preposisi dalam konstruksi sintaksis.

5. Kata Tugas

Kata tugas adalah kata yang hanya memiliki arti gramatikal dan tidak memiliki arti leksikal.

2.2.9 Kalimat

Kalimat adalah satuan bahasa terkecil dalam wujud lisan atau tulisan, yang mengungkapkan pikiran yang utuh (Sutami, 2020). Kalimat terdiri atas deret kata yang dimulai dengan huruf kapital dan diakhiri dengan tanda titik (.), tanda tanya (?), atau tanda seru (!).

Unsur-unsur kalimat terdiri dari kata, kelompok kata dan lagu kalimat. Di dalam kalimat terdapat pengaturan hubungan kedudukan antara bagian-bagiannya. Ada bagian didalam kalimat yang menunjukkan sebagai “pelaku”, ada bagian yang menunjukkan sebagai “perbuatan”, ada bagian yang menunjukkan “bagaimana perbuatan itu dilakukan”.

Berdasarkan jabatannya kalimat terdiri dari :

1. Subyek, yaitu bagian yang menjadi pangkal atau pokok pembicaraan.
2. Predikat, yaitu bagian yang menerangkan subyek, biasanya berdiri sesudah subyek.
3. Obyek, yaitu bagian yang menjadi tujuan.
4. Keterangan, yaitu bagian yang menunjukkan waktu (keterangan waktu), tempat (keterangan tempat), alat (keterangan alat) dan sebagainya.

Sedangkan kalimat berdasarkan fungsinya, dapat dikategorikan sebagai berikut:

1. Kalimat pernyataan
2. Kalimat pertanyaan
3. Kalimat perintah

Kalimat seruan



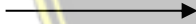
2.2.10 Unified Modeling Language (UML)

Unified Modeling Language (UML) merupakan salah satu bentuk *language* atau bahasa, menurut pencetusnya UML di definisikan sebagai bahasa visual untuk menjelaskan, memberikan spesifikasi, merancang, membuat model, dan mendokumentasikan aspek-aspek dari sebuah sistem. Sedangkan menurut Munawar UML adalah salah satu alat bantu yang sangat handal di dunia pengembangan sistem yang berorientasi objek. Dengan menggunakan UML dapat dibuat model untuk semua jenis aplikasi piranti lunak, di mana aplikasi tersebut dapat berjalan pada piranti keras, sistem operasi dan jaringan apapun, serta ditulis dalam bahasa pemrograman apapun. Tetapi karena UML juga menggunakan *class* dan *operation* dalam konsep dasarnya, maka UML lebih cocok untuk penulisan piranti lunak dalam bahasa berorientasi objek (Pressman, 2005).

2.2.10.1 Use Case Diagram

Use case diagram bersifat statis, diagram ini memperlihatkan himpunan *use case* dan aktor-aktor suatu jenis khusus dari kelas. Terutama sangat penting untuk mengorganisasikan dan memodelkan perilaku dari suatu sistem yang dibutuhkan serta diharapkan pengguna (Mubarak, 2019).

Tabel 2.1 Simbol Use Case




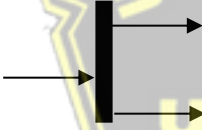
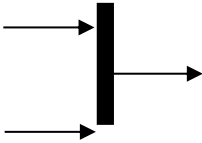
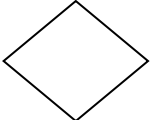
No.	Simbol	Keterangan
1	Aktor 	Merupakan kesatuan <i>eksternal</i> yang berinteraksi dengan sistem.
2.	Use Case 	Rangkaian/uraian sekelompok yang saling terkait dan membentuk sistem.
3.	Generelation 	Menggambarkan hubungan khusus atau interaksi dalam objek.

2.2.10.2 Activity Diagram

Activity diagram bersifat dinamis, diagram state ini memperlihatkan state-state pada sistem, diagram ini terutama penting untuk memperlihatkan sifat dinamis dari antarmuka, kalaborasi dan pemodelan sistem-sistem yang reaktif (Mubarak, 2019).

Tabel 2.2 Simbol Activity Diagram

No.	Simbol	Keterangan

1	Start State 	<i>Start state</i> adalah sebuah kondisi awal sebuah <i>object</i> sebelum ada perubahan keadaan. Start state digambarkan dengan sebuah lingkaran solid.
2.	End State 	<i>End state</i> adalah menggambarkan ketika objek berhenti memberi respon terhadap sebuah event. <i>End state</i> digambarkan dengan lingkaran solid di dalam sebuah lingkaran kosong.
3.	State/Activities 	<i>State</i> atau <i>activities</i> menggambarkan kondisi sebuah entitas, dan digambarkan dengan segiempat yang pinggirnya.
4.	Fork (Percabangan) 	<i>Fork</i> atau percabangan merupakan pemisalah beberapa aliran konkuren dari suatu aliran tunggal.
5.	Join (Penggabungan) 	<i>Join</i> atau penggabungan merupakan penggabungan beberapa aliran konkuren dalam aliran tunggal.
6.	Decision 	<i>Decision</i> merupakan suatu logika aliran konkuren yang mempunyai dua cabang aliran konkuren.

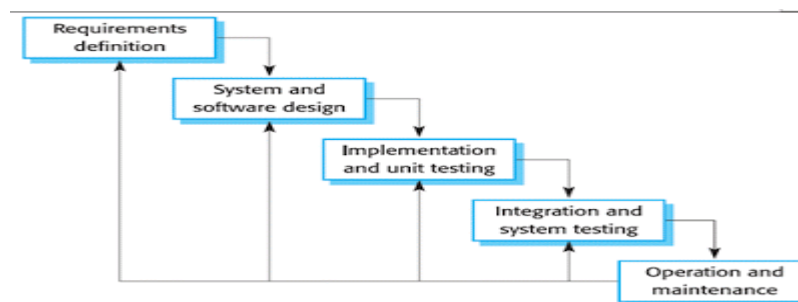
--	--	--

2.2.11 Model Pengembangan Sistem

Suatu sistem yang baik dirancang dengan menggunakan perencanaan dan mengikuti salah satu model proses perangkat lunak yang sesuai dengan sistem yang akan dibangun. Suatu model proses adalah juga dikenal sebagai sebagai suatu paradigma teknik pengembangan perangkat lunak. Model proses tersebut dipilih didasarkan dari kebutuhan sebuah sistem dan sifat alami proyek atau aplikasi, metoda dan *tools* untuk digunakan, kendali serta *deliverables* yang diperlukan. Masing-masing model menunjuk tahap yang umum yang diperlukan, antar lain tahap definisi, tahap pengembangan dan tahap pemeliharaan. Model yang biasanya digunakan adalah *Linear Sequential model, prototyping, the incremental model, the spiral model and the component assembly model*. (Tujni & Hutrianto, 2020)

2.2.12 Waterfall

Metodologi Waterfall adalah metode yang menggambarkan proses pembuatan sistem apabila suatu proses dalam perkembangan project belum terselesaikan, maka proses lain dalam project ini tidak dapat dimulai. Jika anda telah memasuki tahap selanjutnya dalam project ini, maka anda tidak bisa kembali ke tahap sebelumnya. Biasanya waterfall model digunakan dalam project bersekala kecil, ketika requirement ditentukan pada awal project. Biasanya model ini menganggap bahwa requirement dari suatu user sudah tetap dan tidak akan berubah lagi.



Gambar 2.1 Model Waterfall (Tujni & Hutrianto, 2020)

Berikut adalah penjelasan detail dari masing-masing tahap dalam model pengembangan *waterfall*.

a. Requirements definition (analisis kebutuhan)

Analisis kebutuhan merupakan tahap pertama yang menjadi dasar dalam proses pembuatan sistem pakar. Kelancaran proses pembuatan aplikasi secara keseluruhan dan kelengkapan fitur *software* yang di hasilkan sangat tergantung pada analisis kebutuhan. Untuk memperoleh informasi tentang proses dan kebutuhan data, penulis melakukan survey dan mencari data penelitian sebelumnya.

b. Sistem and software design (desain sistem)

Merupakan gambaran yang di berikan kepada user tentang kegiatan yang akan di lakukan dan merupakan persiapan dari desain sistem setelah di lakukan analisis sistem.

c. Impelentation and Unit Testing (penulisan koding)

Merupakan penerjemahan design dalam bahasa yang bisa dikenali oleh komputer dalam bentuk perintah-perintah.. Tahapan inilah yang merupakan tahapan secara real dalam mengerjakan suatu sistem. Ini. Setelah pengkodean selesai maka akan dilakukan testing terhadap sistem yang telah dibuat tadi. Tujuan testing yaitu menemukan kesalahan-kesalahan terhadap sistem tersebut dan kemudian bisa diperbaiki. Berbeda dengan tahap 2 sebelumnya, pada tahap ini penulisan kode program ini keterlibatan narasumber tidaklah besar.

d. Intergration and Sistem Testing (penerapan dan pengujian sistem)

Pengujian perangkat lunak yang di lakukan untuk memastikan bahwa aplikasi yang di buat telah sesuai dengan desainnya dan semua fungsi dapat di pergunakan dengan baik tanpa adanya kendala. Pengujian yang dilakukan pada aplikasi ini menggunakan *blackbox*.

Pengujian *blackbox* berfokus pada persyaratan fungsional perangkat lunak. Pengujian ini memungkinkan analis sistem memperoleh kumpulan kondisi input yang akan mengerjakan seluruh keperluan fungsional program. Tujuan metode ini di gunakan adalah untuk mencari kesalahan pada fungsi yang salah atau hilang. Kesalahan pada *interface*, kesalahan pada struktur data atau akses ke database, kesalahan informasi dan tujuan akhir.

e. Maintenance (perawatan)

Setelah sebuah sistem dapat terselesaikan maka sistem membutuhkan supaya sistem dapat berjalan dengan baik dan data dapat tersimpan dengan benar dan aman.

2.2.13 Paragraf

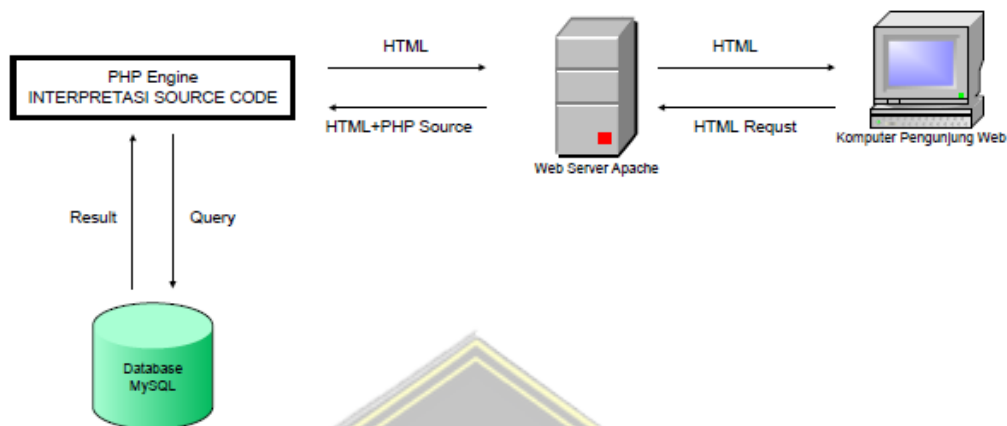
Paragraf disebut juga alinea. Kata paragraf merupakan kata serapan dari bahasa Inggris *paragraph*, sedangkan kata alinea dari bahasa Belanda dengan ejaan yang sama. Paragraf adalah seperangkat kalimat yang membicarakan suatu gagasan atau topic (Sutami, 2020). Terdapat dua syarat dalam membentuk paragraf :

1. Menulis pernyataan (kalimat) tentang pokok bahasan dengan baik.
2. Mengangkat pola susunan rincian dengan patut.

2.2.14 PHP

PHP adalah bahasa skrip dalam *server (server-side embedded scripting language)*. Artinya, PHP bekerja di dalam HTML dengan tugas membuat isi dokumen sesuai

permintaan (Mubarak, 2019). Alur kerja PHP itu sendiri dapat kita lihat pada gambar 2.3



Gambar 2.2: alur PHP

Pada gambar di atas tampak alur kerja *engine* PHP. Pada awalnya, sebuah komputer pengunjung web melakukan HTTP *request* terhadap halaman tertentu. Jika halaman yang di-*request* tersebut adalah halaman PHP, maka *Web Server Apache* akan meneruskan halaman PHP tersebut ke *PHP engine*. *PHP engine* akan melakukan interpretasi terhadap *source* PHP dalam halaman tersebut, dengan bantuan dari *database* jika perlu. Setelah interpretasi selesai, maka PHP akan mengembalikan hasilnya yang berupa HTML murni (tanpa adanya *source* PHP sama sekali) kepada *Web Server Apache*. Setelah itu *Web Server Apache* akan meneruskannya kepada komputer pengunjung web tadi dalam bentuk HTML yang dapat ditampilkan oleh internet browser.

PHP tidak hanya didesain untuk *Web Server Apache* dan *database MySQL*. Bahkan PHP tidak hanya bisa dipakai untuk halaman web. PHP dapat berjalan pada hampir semua *web server* yang mendukung konsep CGI-binary atau ISAPI module, di antaranya *Personal Web Server (PWS)* 3 ke atas, *Internet Information Service (IIS)* 3 ke atas, *Apache*, *OmniHTTPd* 2 ke atas, *Oreilly Website Pro*, *Xitami*, dan *Netscape Enterprise Server*.

Koneksi *database* pada PHP amat mudah, cukup dengan menggunakan 2 atau 3 baris perintah saja. Fungsi-fungsi untuk *database* cukup lengkap, dan *database*

yang didukung amat banyak, di antaranya MySQL, PostgreSQL, dBase, DB++, FrontBase, ODBC, Sybase-CT, Oracle, Informix, dan Interbase.

PHP memiliki empat kelebihan utama yang menarik minat banyak pengguna. Kelebihan utama PHP tersebut diringkas dalam 4P berikut:

1. *Practicality*. PHP dibuat dengan menitikberatkan pada kepraktisan. Hasilnya, PHP adalah bahasa pemrograman minimalis, dilihat dari segi kebutuhan pengguna dan kebutuhan sintaks.
2. *Power*. PHP memiliki banyak kemampuan, mulai dari kemampuan untuk terhubung dengan basis data, membuat halaman *web* dinamis, membuat dan memanipulasi berkas gambar, *Flash* dan *PDF*, berkomunikasi dengan bermacam protokol seperti *IMAP* dan *POP3*, dan masih banyak lagi.
3. *Possibility*. PHP dapat menyediakan lebih dari satu solusi untuk suatu masalah
4. *Price*. PHP selalu dirilis kepada publik tanpa ada batasan penggunaan, modifikasi, atau redistribusi.

2.2.15 MySQL

MySQL merupakan salah satu perangkat lunak perangkat system manajemen basis data atau RDBMS (*Relational Database Management System*) seperti halnya oracle, PostgreSQL, dan Microsoft SQL. MySQL digunakan sebagai penyimpanan data.

MySQL menggunakan bahasa Structured Query Language (SQL). MySQL dalam operasi client-server melibatkan server daemon MySQL disisi server dan berbagai macam program serta library yang berjalan disisi client. MySQL mampu menangani data yang cukup besar. Perusahaan yang mengembangkan MySQL yaitu TcX, mengaku menyimpan data lebih dari 40 database, 10.000 tabel dan sekitar 7 juta baris, totalnya kurang lebih 100 Gigabyte data.

SQL adalah bahasa standar yang digunakan untuk mengakses database server. Bahasa ini pada awalnya dikembangkan oleh IBM, namun telah diadopsi dan digunakan sebagai standar industri. Dengan menggunakan SQL, proses akses database menjadi lebih user-friendly dibandingkan dengan menggunakan dBASE atau Clipper yang masih menggunakan perintah-perintah pemrograman.

MySQL merupakan software database yang paling populer di lingkungan Linux, kepopuleran ini karena ditunjang performa query dari databasenya yang saat ini bisa dikatakan paling cepat dan jarang bermasalah. MySQL ini juga sudah dapat berjalan pada lingkungan Windows. (A.S Rossa, 2011)

2.2.16 Hypertext Markup Language (HTML)

HTML adalah bahasa markup yang umum digunakan. Kepopuleran HTML disebabkan karena HTML ini mudah digunakan. Pembuatan dokumen WEB dengan HTML dapat dilakukan dengan mudah dan cepat. Dokumen web dapat tersaji dengan cepat ke banyak pembaca di seluruh dunia sekaligus. HTML mudah melakukan kontrol terhadap tampilan halaman web baik berupa teks, gambar, suara, animasi maupun video.

HTML merupakan singkatan dari Hypertext Markup Language. HTML digunakan untuk membangun halaman web. Sekalipun banyak orang menyebutnya sebagai bahasa pemrograman, HTML sebenarnya sama sekali bukan bahasa pemrograman, hal ini terlihat dari namanya, HTML adalah suatu bahasa Markup. HTML digunakan untuk melakukan penandaan terhadap sebuah dokumen teks. Tanda tersebut digunakan untuk menentukan format atau style dari teks yang ditandai. (Sovia, 2017)

2.2.17 Cascading Style Sheet (CSS)

Cascading Style Sheets (CSS) adalah suatu bahasa stylesheet yang digunakan untuk mengatur tampilan suatu dokumen yang ditulis dalam bahasa markup. CSS bekerja

sebagai pelengkap pada HTML. Dan tugas utama CSS adalah menetapkan aturan tampilan/style yang akan digunakan pada sebuah website. CSS dapat mengubah besar kecilnya text, mengganti warna background pada sebuah halaman, atau dapat pula mengubah warna border pada tabel, dan masih banyak lagi hal yang dapat dilakukan oleh CSS. Singkatnya, CSS digunakan untuk mengatur susunan tampilan pada halaman HTML. (Sovia, 2017)

2.2.18 Javascript

Javascript merupakan scripting language yang terintegrasi dengan web browser untuk memberikan fleksibilitas tambahan bagi programmer untuk mengontrol elemen elemen dalam halaman web. Aplikasi javascript sebenarnya cukup luas namun pada aplikasi yang dirancang, digunakan javascript yang dibatasi untuk DHTML (Dinamic HyperText Markup Language), yaitu javascript yang digunakan untuk mengakses property, method, dan event handler yang disediakan oleh DOM dan CSS.

Property didefinisikan sebagai setting nilai suatu objek tertentu. Contohnya adalah warna suatu teks, action untuk suatu form, nama file untuk suatu gambar, dll. Method adalah fungsi-fungsi yang dapat diterapkan dalam suatu objek. Misalnya maximize untuk window. Event handler menspesifikasikan bagaimana suatu objek merespon terhadap suatu kejadian, misalnya suatu button di klik, window di resize, dan lain-lain. (Sovia, 2017)

BAB 3

METODE PENELITIAN

3.1 Metode Pengumpulan Data

3.1.1 Observasi

Metode yang digunakan untuk mengumpulkan data dalam penelitian dengan cara pengambilan data teks berita secara langsung pada portal-portal berita online.

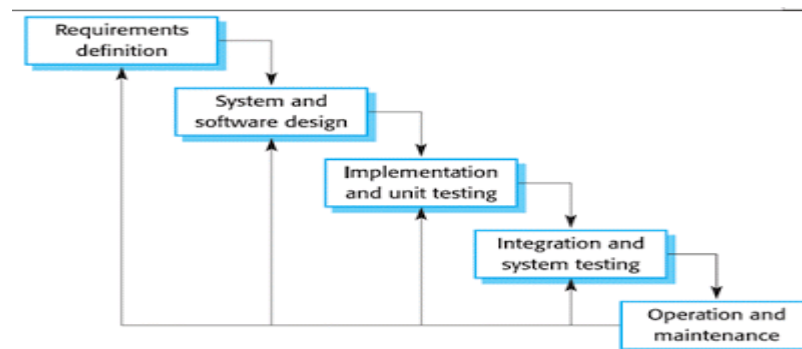
3.1.2 Studi Pustaka

Selain menggunakan metode pengumpulan data dokumen observasi, penulis juga menggunakan metode pengumpulan data studi pustaka yaitu dengan cara mencari sumber - sumber dari jurnal, buku, media internet, dan beberapa media yang berhubungan dengan penelitian. Beberapa studi pustaka yang didapatkan adalah:

- a. Buku – buku yang membahas tentang Algoritma TEXTRANK yang digunakan untuk membuat peringkasan teks otomatis.
- b. Jurnal tentang Peringkasan Dokumen Otomatis yang berkaitan dengan Algoritma TEXTRANK.

3.2 Metode Pengembangan Sistem

Perancangan sistem penelitian menggunakan model waterfall. Langkah-langkah model waterfall yaitu



Gambar 3.1 Model Waterfall

Berikut adalah penjelasan detail dari masing-masing tahap dalam model pengembangan *waterfall*.

a. Requirements definition (analisis kebutuhan)

Analisis kebutuhan merupakan tahap pertama yang menjadi dasar dalam proses pembuatan sistem pakar. Kelancaran proses pembuatan aplikasi secara keseluruhan dan kelengkapan fitur *software* yang di hasilkan sangat tergantung pada analisis kebutuhan. Untuk memperoleh informasi tentang proses dan kebutuhan data, penulis melakukan survey dan mencari data penelitian sebelumnya.

b. Sistem and software design (desain sistem)

Merupakan gambaran yang di berikan kepada user tentang kegiatan yang akan di lakukan dan merupakan persiapan dari desain sistem setelah di lakukan analisis sistem.

c. Impelentation and Unit Testing (penulisan koding)

Merupakan penerjemahan design dalam bahasa yang bisa dikenali oleh komputer dalam bentuk perintah-perintah.. Tahapan inilah yang merupakan tahapan secara real dalam mengerjakan suatu sistem. Ini. Setelah pengkodean selesai maka akan dilakukan testing terhadap sistem yang telah dibuat tadi. Tujuan testing yaitu menemukan kesalahan-kesalahan terhadap sistem tersebut dan kemudian bisa diperbaiki. Berbeda dengan tahap 2 sebelumnya, pada tahap ini penulisan kode program ini keterlibatan narasumber tidaklah besar.

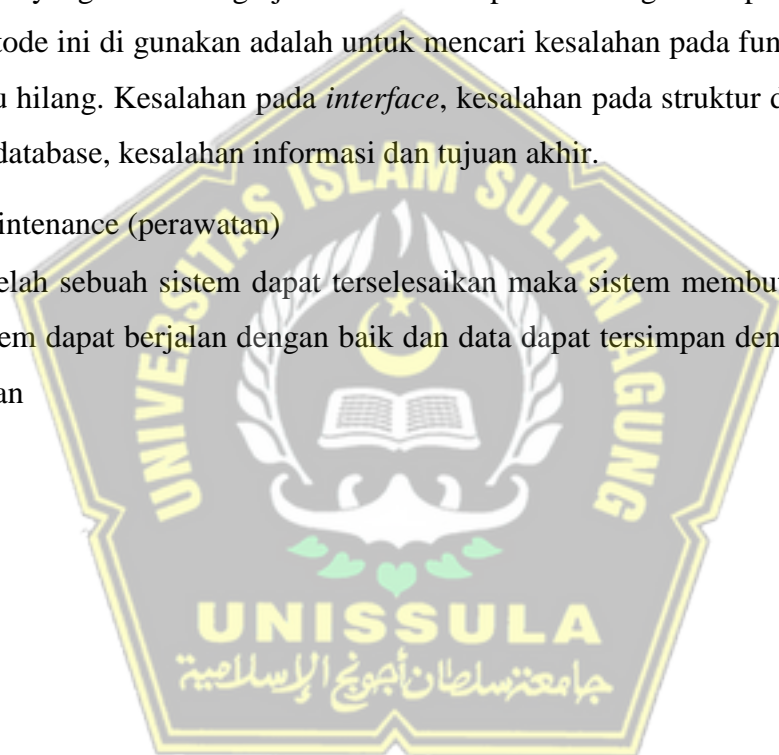
d. Intergration and Sistem Testing (penerapan dan pengujian sistem)

Pengujian perangkat lunak yang di lakukan untuk memastikan bahwa aplikasi yang di buat telah sesuai dengan desainnya dan semua fungsi dapat di pergunakan dengan baik tanpa adanya kendala. Pengujian yang dilakukan pada aplikasi ini menggunakan *blackbox*.

Pengujian *blackbox* berfokus pada persyaratan fungsional perangkat lunak. Pengujian ini memungkinkan analis sistem memperoleh kumpulan kondisi input yang akan mengerjakan seluruh keperluan fungsional program. Tujuan metode ini di gunakan adalah untuk mencari kesalahan pada fungsi yang salah atau hilang. Kesalahan pada *interface*, kesalahan pada struktur data atau akses ke database, kesalahan informasi dan tujuan akhir.

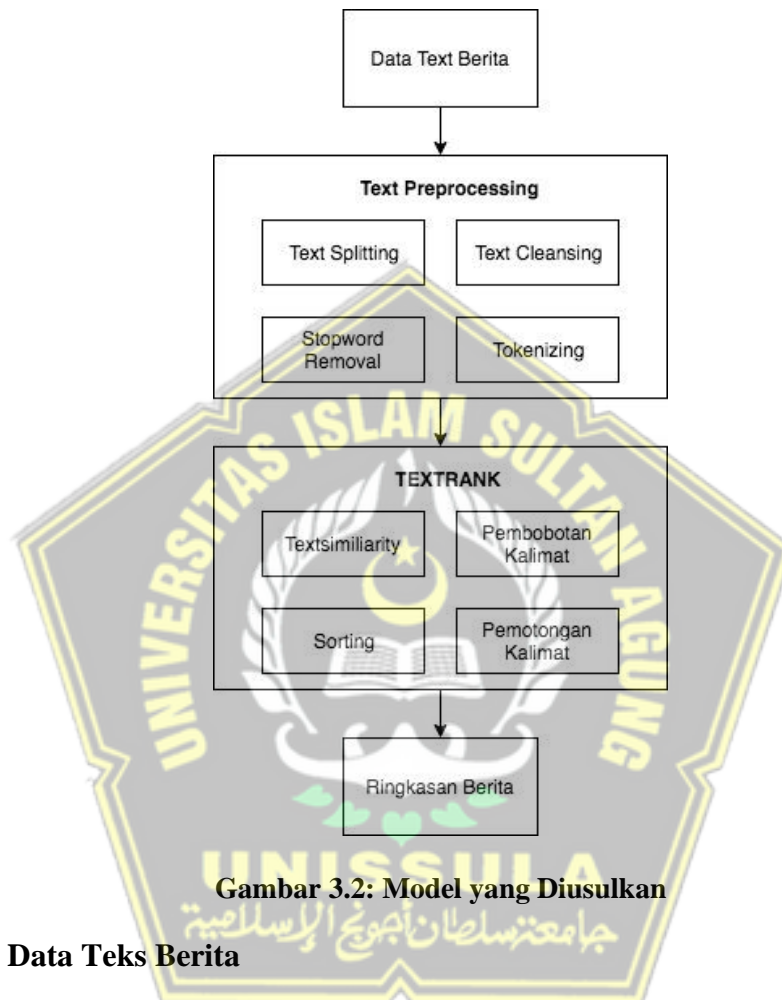
e. Maintenance (perawatan)

Setelah sebuah sistem dapat terselesaikan maka sistem membutuhkan supaya sistem dapat berjalan dengan baik dan data dapat tersimpan dengan benar dan aman



3.3 Model yang Diusulkan

Sistem peringkasan teks otomatis ini dibangun dengan memiliki 2 tahapan utama teks preprocessing dan implementasi textrank yang dijelaskan sebagai berikut:



Gambar 3.2: Model yang Diusulkan

3.3.1 Data Teks Berita

Data teks berita pada penelitian ini diambil dari media berita online terkait dengan kota semarang dan sekitarnya. Sebagai sampel diambil potongan paragraf berikut:

Sumber: <http://www.tribunnews.com/regional/2019/03/11/merasa-diselingkuhi-seorang-pria-pasang-foto-wanita-di-baliho-bertulis-aku-mau-putus>

Kategori: Hiburan

Isi Berita:

TRIBUNNEWS.COM - Biasanya [baliho](#) jumbo di pinggir jalan berisi foto model iklan atau calon legislatif (caleg).

Namun, ada yang unik dari jalanan Kota [Semarang](#) ini, foto yang terpampang di [baliho](#) itu adalah foto seorang perempuan yang bukan artis maupun caleg, tetapi wanita yang diajak putus oleh sang pacar.

Foto [baliho](#) tersebut diunggah oleh akun Instagram [@makassar_iinfo](#) pada Minggu (10/3/2019).

Dalam keterangan unggahan itu, disebutkan oleh pengunggah lokasi [baliho](#) tersebut berada di pertigaan Jalan Sultan Agung Kota [Semarang](#).

Tak hanya foto [baliho](#) yang diunggah, akun tersebut juga mengunggah video pria dan wanita yang diduga orang yang sama dalam [baliho](#) itu sedang bercecekok di seberang jalan tempat [baliho](#) itu dipasang.

Artikel ini telah tayang di [Tribunnews.com](#) dengan judul Merasa Diselingkuhi, Seorang Pria Pasang Foto Wanita di Baliho Bertulis 'Aku Mau Putus', <http://www.tribunnews.com/regional/2019/03/11/merasa-diselingkuhi-seorang-pria-pasang-foto-wanita-di-baliho-bertulis-aku-mau-putus>.

Editor: Wahid Nurdin

3.3.2 Algoritma TEXRANK

Tahap pertama proses praproses yang merupakan tahap dimana dokumen yang akan digunakan dibersihkan dari noise atau dari hal yang tidak mempunyai pengaruh dalam ringkasan. Praproses yang dilakukan dalam tahap ini antara lain: Splitting, tokenizing, filtering, Case Folding dan stopword removal.

1. Splitting

Splitting, teks berita yang utuh akan dipisah menjadi beberapa kalimat-kalimat.

Tabel 3.1: Teks Splitting

No	Kalimat
S1	Biasanya baliho jumbo di pinggir jalan berisi foto model iklan atau calon legislatif (caleg).
S2	Namun, ada yang unik dari jalanan Kota Semarang ini, foto yang terpampang di baliho itu adalah foto seorang perempuan yang bukan artis maupun caleg, tetapi wanita yang diajak putus oleh sang pacar.
S3	Foto baliho tersebut diunggah oleh akun Instagram @makassar_iinfo pada Minggu (10/3/2019).

S4	Dalam keterangan unggahan itu, disebutkan oleh pengunggah lokasi baliho tersebut berada di pertigaan Jalan Sultan Agung Kota Semarang.
S5	Tak hanya foto baliho yang diunggah, akun tersebut juga mengunggah video pria dan wanita yang diduga orang yang sama dalam baliho itu sedang bercecek di seberang jalan tempat baliho itu dipasang.

2. Text Cleansing

Text cleansing merupakan tahap dimana teks pada kalimat yang telah dipisah dibersihkan dari tanda baca, angka dan membuat teks menjadi lowercase atau huruf kecil.

Tabel 3.2: Teks Cleansing

No	Kalimat
S1	biasanya baliho jumbo di pinggir jalan berisi foto model iklan atau calon legislatif caleg
S2	namun ada yang unik dari jalanan kota semarang ini foto yang terpampang di baliho itu adalah foto seorang perempuan yang bukan artis maupun caleg tetapi wanita yang diajak putus oleh sang pacar
S3	foto baliho tersebut diunggah oleh akun instagram Makassar iinfo pada minggu
S4	dalam keterangan unggahan itu disebutkan oleh pengunggah lokasi baliho tersebut berada di pertigaan jalan sultan agung kota semarang
S5	tak hanya foto baliho yang diunggah akun tersebut juga mengunggah video pria dan wanita yang diduga orang yang sama dalam baliho itu sedang bercecek di seberang jalan tempat baliho itu dipasang

3. Tokenizing

Tokenizing kata, pemotongan kalimat menjadi kata-kata tunggal dengan (“ ”) spasi sebagai delimeter

Tabel 3.3: Hasil Tokenizing

No	Kalimat
S1	biasanya baliho jumbo di pinggir jalan berisi foto model iklan atau calon legislatif caleg
S2	namun ada yang unik dari jalanan kota semarang ini foto yang terpampang di baliho itu adalah foto seorang perempuan yang bukan artis maupun caleg tetapi wanita yang diajak putus oleh sang pacar
S3	foto baliho tersebut diunggah oleh akun instagram Makassar iinfo pada minggu
S4	dalam keterangan unggahan itu disebutkan oleh pengunggah lokasi baliho tersebut berada di pertigaan jalan sultan agung kota semarang
S5	tak hanya foto baliho yang diunggah akun tersebut juga mengunggah video pria dan wanita yang diduga orang yang sama dalam baliho itu sedang bercecok di seberang jalan tempat baliho itu dipasang

4. Stopword removal

Stopword Removal, menghilangkan kata-kata umum yang sering muncul.

Kata-kata ini dihilangkan karena tidak memiliki arti

Tabel 3.4: Stopword Removal

No	Kalimat Asli	Hasil Removal
S1	biasanya baliho jumbo di pinggir jalan berisi foto model iklan atau calon legislatif caleg	biasanya baliho jumbo pinggir jalan berisi foto model iklan calon legislatif caleg
S2	namun ada yang unik dari jalanan kota semarang ini foto yang terpampang di baliho itu adalah foto seorang perempuan yang bukan artis maupun caleg tetapi wanita yang diajak putus oleh sang pacar	namun ada yang unik jalanan kota semarang foto terpampang baliho foto seorang perempuan bukan artis maupun caleg wanita diajak putus sang pacar

S3	foto baliho tersebut diunggah oleh akun instagram Makassar iinfo pada minggu	foto baliho diunggah akun instagram Makassar iinfo pada minggu
S4	dalam keterangan unggahan itu disebutkan oleh pengunggah lokasi baliho tersebut berada di pertigaan jalan sultan agung kota semarang	keterangan unggahan disebutkan pengunggah lokasi baliho tersebut berada pertigaan jalan sultan agung kota semarang
S5	tak hanya foto baliho yang diunggah akun tersebut juga mengunggah video pria dan wanita yang diduga orang yang sama dalam baliho itu sedang bercecek di seberang jalan tempat baliho itu dipasang	hanya foto baliho diunggah akun tersebut mengunggah video pria wanita diduga orang sama baliho sedang bercecek seberang jalan tempat baliho dipasang

4. Teks hasil preprocessing

Berdasarkan langkah-langkah teks preprocessing menghasilkan teks yang telah siap untuk masuk pada algoritma TEXTRANK. Hasil dari proses text preprocessing ditampilkan pada tabel 3.5 dibawah.

Tabel 3.5: Hasil Teks Preprocessing

No	Kalimat	Panjang Kalimat (jumlah kata)
S1	biasanya baliho jumbo pinggir jalan berisi foto model iklan calon legislatif caleg	12
S2	namun ada yang unik jalanan kota semarang foto terpampang baliho foto seorang perempuan bukan artis maupun caleg wanita diajak putus sang pacar	22
S3	foto baliho diunggah akun instagram Makassar iinfo pada minggu	9

S4	keterangan unggahan disebutkan pengunggah lokasi baliho tersebut berada pertigaan jalan sultan agung kota semarang	14
S5	hanya foto baliho diunggah akun tersebut mengunggah video pria wanita diduga orang sama baliho sedang bercecek seberang jalan tempat baliho dipasang	21

5. Perhitungan TEXTRANK

Pada tahap awal dari algoritma textrank ini adalah menghitung kesamaan/similarity antar segmen atau antar kalimat dengan content overlap, dari pengambilan kalimat yang sama terhadap setiap segmen kalimat. Jika terdapat 2 kata yang sama dalam 1 kalimat, maka tetap dihitung 1. Berikut adalah rumus dari content overlap similarity:

$$Similarity(S_i, S_j) = \frac{|{\{W_k | W_k \in S_i \& W_k \in S_j\}}|}{\log(|S_i|) + \log(|S_j|)}$$

Setelah itu, terapkan rumus similarity content overlap pada semua kalimat. Dibawah ini adalah contoh perhitungan content overlap similarity untuk kalimat 1 dan 2

Kata yang sama antara kalimat (S1) dan kalimat 2 (S2) =

Tabel 3.6: Text Similiarity

S1	biasanya baliho jumbo pinggir jalan berisi foto model iklan calon legislatif caleg
S2	namun ada yang unik jalanan kota semarang foto terpampang baliho foto seorang perempuan bukan artis maupun caleg wanita diajak putus sang pacar
Similiarity Word	Baliho, foto, caleg

Jumlah	3
--------	---

Maka berdasarkan formula similarity :

$$\text{Similarity}(S1,S2) = \frac{3}{\log(12) + \log(22)} = 1.2388$$

Tabel 3.7: Matrix Text Similarity

	S1	S2	S3	S4	S5
S1	-	1.2388	0.9835	0.8987	1.2491
S2	1.2388	-	0.8708	1.2055	1.1258
S3	0.9835	0.8708	-	0.4761	0.8785
S4	0.8987	1.2055	0.4761	-	0.2153
S5	1.2491	1.1258	0.8785	0.2153	-

Setelah dilakukan pre-processing dan similarity, pada tahap ini dilakukan perhitungan bobot antar kalimat yang nantinya akan diranking untuk mendapatkan hasil ringkasan. Untuk menghitung bobot kalimat dengan textrank diperukan formula:

$$WS(S_i) = (1 - d) + d * \sum_{v_j \in Adj(v_i)} \frac{W_{ji}}{\sum_{v_k \in Adj(v_j)} W_{jk}} WS(S_j)$$

Dimana d merupakan dumping factor yang diisi dengan 0.85, W_{ji} merupakan similarity dari kalimat yang ingin dihitung, W_{jk} merupakan kalimat yang terhubung dengan W_j, S_j diisi dengan 1

Tabel 3.8: Perhitungan Bobot Kalimat

	S1	S2	S3	S4	S5	$\frac{W_{ji}}{\sum_{v_k \in Adj(v_j)} W_{jk}}$				WS (si)
S1	0.0000	1.2388	0.9835	0.8987	1.2491	0.2835	0.2113	0.2478	0.4321	1.1486

S2	1.2388	0.0000	0.8708	1.2055	1.1258	0.2790	0.1961	0.3463	0.3062	1.1084
S3	0.9835	0.8708	0.0000	0.4761	0.8785	0.3065	0.3439	0.2375	0.3918	1.2377
S4	0.8987	1.2055	0.4761	0.0000	0.2153	0.3215	0.5434	0.3059	0.1553	1.2772
S5	1.2491	1.1258	0.8785	0.2153	0.0000	0.3601	0.4364	0.4647	0.1458	1.3460

Rumus ini diterapkan pada semua kalimat untuk mendapatkan bobot setiap kalimat yang hasilnya akan di ranking.

Tabel 3.9: Hasil Urutan Pembobotan Kalimat

Kalimat	Bobot
S5	1.34600375
S4	1.27721408
S3	1.23768692
S1	1.14858187
S2	1.10837849

Jika ekstraksi kalimat dilakukan 25% dari teks asli, maka terdapat 2 kalimat yang masuk kedalam hasil ringkasan, yaitu kalimat 5 dan kalimat 4. Setelah mendapatkan hasil ringkasan pada dokumen tunggal, dilakukan proses ini sebanyak dokumen yang ada, lalu semua hasil ringkasan digabungkan. Maka hasil ringkasan dapat dilihat pada paragraf berikut:

“Dalam keterangan unggahan itu, disebutkan oleh pengunggah lokasi baliho tersebut berada di pertigaan Jalan Sultan Agung Kota Semarang. Tak hanya foto baliho yang diunggah, akun tersebut juga mengunggah video pria dan wanita yang diduga orang yang sama dalam baliho itu sedang bercecok di seberang jalan tempat baliho itu dipasang.”

3.4 Analisa Aplikasi

3.4.1 Deskripsi Umum Aplikasi

Tujuan dari penelitian ini adalah membuat aplikasi berbasis web meringkas berita secara otomatis. Berita didapatkan dari portal berita online yang terkait dengan berita local kota semarang dan sekitarnya. Kemudian dilakukan peringkasan dari konten isi berita menggunakan metode TEXTRANK.

3.4.2 Analisa Kebutuhan

Dengan analisa aplikasi dapat menguraikan secara utuh proses menjadi komponen-komponen dasar dengan tujuan mengidentifikasi serta mengevaluasi permasalahan dan kebutuhan. Analisa kebutuhan aplikasi meliputi kebutuhan fungsional.

Kebutuhan fungsional berisi bagian dan proses-proses yang harus ada pada sistem yang akan dibangun. Kebutuhan fungsional diuraikan berdasarkan pengguna dari system sebagai berikut

Admin

1. Admin dapat melakan login untuk masuk kedalam sistem pengelolaan data
2. Admin dapat mengelola data berita. Pengelolaan meliputi tambah, edit dan hapus.
3. Admin dapat mengelola data kategori. Pengelolaan meliputi tambah, edit dan hapus.
4. Sistem dapat meringkas berita secara otomatis menggunakan metode TEXTRANK

User

1. User dapat mengakses aplikasi
2. User dapat melihat daftar berita yang tersimpan dalam database.
3. User dapat melihat ringkasan berita
4. User dapat melihat berita berdasarkan kategori

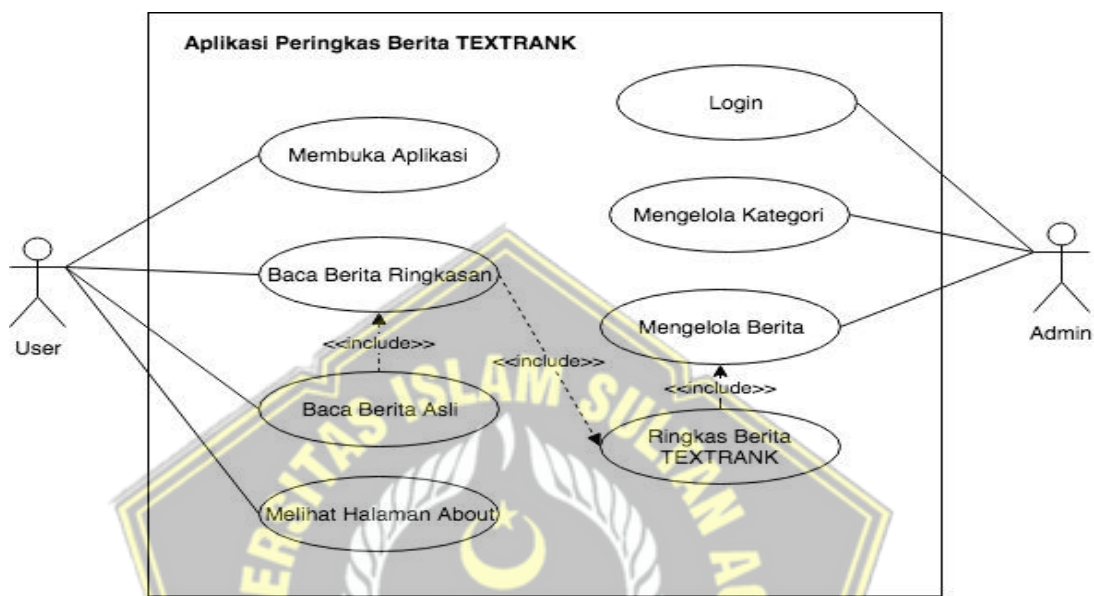
3.5 Perancangan Aplikasi

Pada bab ini dijelaskan tahapan-tahapan perancangan aplikasi berdsasarkan hasil analisa kebutuhan. Perancangan aplikasi dimulai dari perancangan usecase, activity diagram dan sequence diagram.

3.5.1 Use Case Diagram

Dalam use case diagram memperlihatkan hubungan antara actor dan use case. Actor merepresentasikan seorang user yang berinteraksi dengan sistem. Fungsionalitas

sistem didefinisikan kedalam use case dari sudut eksternal yang berguna untuk uji kelayakan sistem. Setiap usecase akan disertai penjelasan yang diuraikan dalam usecase scenario, yang menguraikan nama usecase, usecase yang terkait, aksi actor dalam system.



Gambar 3.3: Use Case Diagram Aplikasi

Pada gambar 3.3 ditampilkan use case diagram dari aplikasi peringkat berita otomatis. Terdapat 2 aktor yaitu user dan admin. User memiliki 4 usecase yang menggambarkan interaksi antara user dengan aplikasi. Sedangkan admin memiliki 3 usecase yang menggambarkan interaksi antara admin dengan aplikasi

Berdasarkan use case diagram tersebut dijelaskan interaksi pada scenario dibawah ini:

Tabel 3.1: Skenario Membuka Aplikasi

Nama Usecase	Membuka aplikasi
Aktor	User
Tujuan	Melihat dan mengoperasikan fitur-fitur pada aplikasi
Kondisi Sebelum	User membuka browser dan menginput alamat URL dari aplikasi

Kondisi Sesudah	Aplikasi menampilkan halaman awal
------------------------	-----------------------------------

Tabel 3.2: Skenario Baca Berita Ringkasan

Nama Usecase	Baca berita ringkasan
Aktor	User
Tujuan	User membaca hasil ringkasan berita
Kondisi Sebelum	Aplikasi telah diakses dan proses ringkasan selesai
Kondisi Sesudah	Aplikasi menampilkan daftar berita-berita dari sumber online yang telah diringkas

Tabel 3.3: Skenario Baca Berita Asli

Nama Usecase	Baca berita asli
Aktor	User
Tujuan	User membaca berita asli dari sumber online
Kondisi Sebelum	Aplikasi telah diakses, proses ringkasan selesai dan User memilih tombol asli
Kondisi Sesudah	Aplikasi menampilkan berita asli dari sumber online

Tabel 3.4: Skenario Melihat Halaman About

Nama Usecase	Membuka Melihat Halaman About
Aktor	User
Tujuan	Melihat halaman about dari aplikasi
Kondisi Sebelum	User memilih halaman about
Kondisi Sesudah	Aplikasi menampilkan halaman about

Tabel 3.5: Skenario Login

Nama Usecase	Login
Aktor	Admin
Tujuan	Masuk kedalam sistem administrator
Kondisi Sebelum	Admin mengakses halaman user dan menginput data login
Kondisi Sesudah	Sistem mengarahkan ke halaman administrator

Tabel 3.6: Skenario Mengelola Kategori

Nama Usecase	Mengelola Kategori
Aktor	Admin
Tujuan	Mengelola data kategori
Kondisi Sebelum	Memilih menu dan tombol fitur kategori
Kondisi Sesudah	Menyimpan, mengedit dan menghapus data kategori

Tabel 3.7: Skenario Mengelola Berita

Nama Usecase	Mengelola Berita
Aktor	Admin
Tujuan	Mngelola data berita
Kondisi Sebelum	User memilih menu dan tombol fitur berita
Kondisi Sesudah	Menyimpan, mengedit dan menghapus data berita

Tabel 3.8: Skenario Ringkasan Berita TEXTRANK

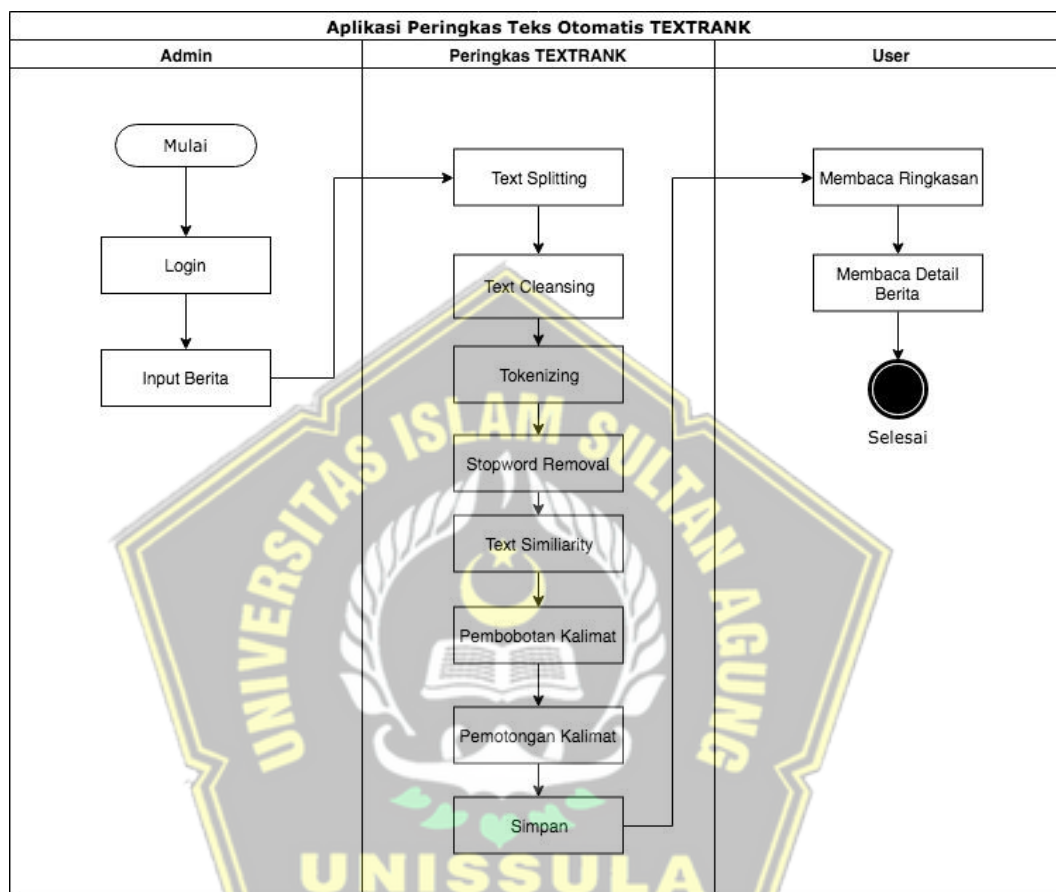
Nama Usecase	Ringkasan Berita TEXTRANK
Aktor	Admin

Tujuan	Memproses teks menjadi ringkasan secara otomatis
Kondisi Sebelum	Admin menginput berita
Kondisi Sesudah	Sistem menyimpan berita hasil inputan dan ringkasan



3.5.2 Activity Diagram

Activity Diagram menggambarkan workflow (aliran kerja) atau aktivitas dari sebuah sistem atau proses bisnis.

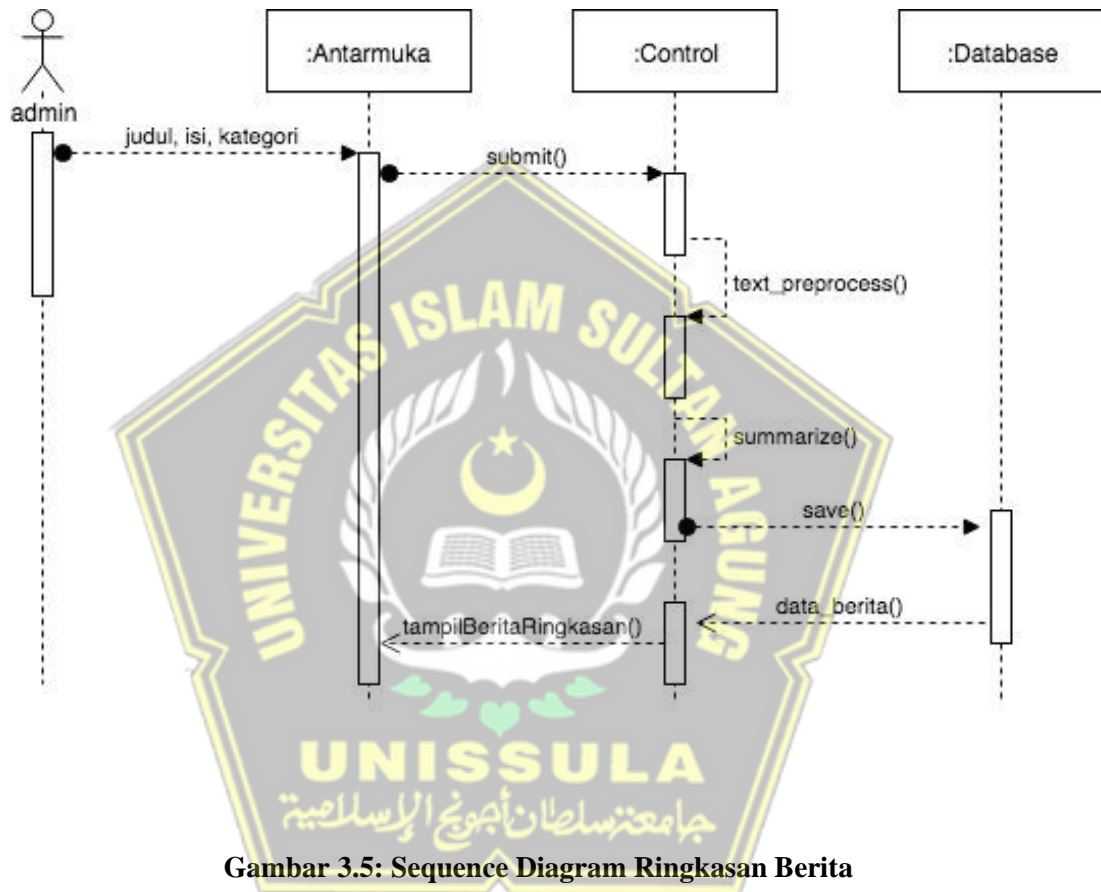


Gambar 3.4: Activity Diagram Aplikasi

Pada gambar 3.4 ditampilkan activity diagram aplikasi peringkat berita local kota semarang. Activity diagram tersebut menggambarkan alur kerja yang dilakukan user terhadap aplikasi. Tahap awal adalah admin masuk kedalam sistem dan input data berita (judul, isi berita dan kategori). Berdasarkan inputan tersebut data isi berita diproses yang melalui tahap text splitting, text cleansing, tokenizing, stopword removal, text similiarity, pembobotan kalimat, pemotongan kalimat dan simpan ke database. Setelah data tersimpan ke database user dapat sewaktu-waktu membaca berita asli dan hasil ringkasan.

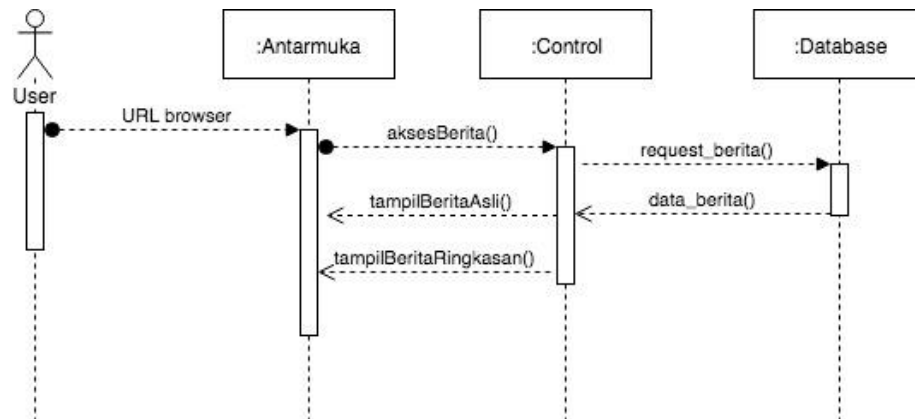
3.5.3 Sequence Diagram

Sequence diagram menggambarkan interaksi antar obyek dan mengindikasikan komunikasi diantara obyek-obyek tersebut. Diagram ini juga menunjukkan serangkaian pesan yang dipertukarkan oleh obyek-obyek yang melakukan suatu tugas atau aksi tertentu.



Gambar 3.5: Sequence Diagram Ringkasan Berita

Pada gambar 3.5 ditampilkan sequence diagram ringkasan berita. Tahap awal dari proses adalah Admin menginput judul, isi dan kategori berita. Kemudian setelah disubmit control akan memproses dengan tahapan `text_preprocess()`, `summarize` dan simpan ke database. Setelah data tersimpan ke database sistem akan mengarahkan ke halaman data berita.

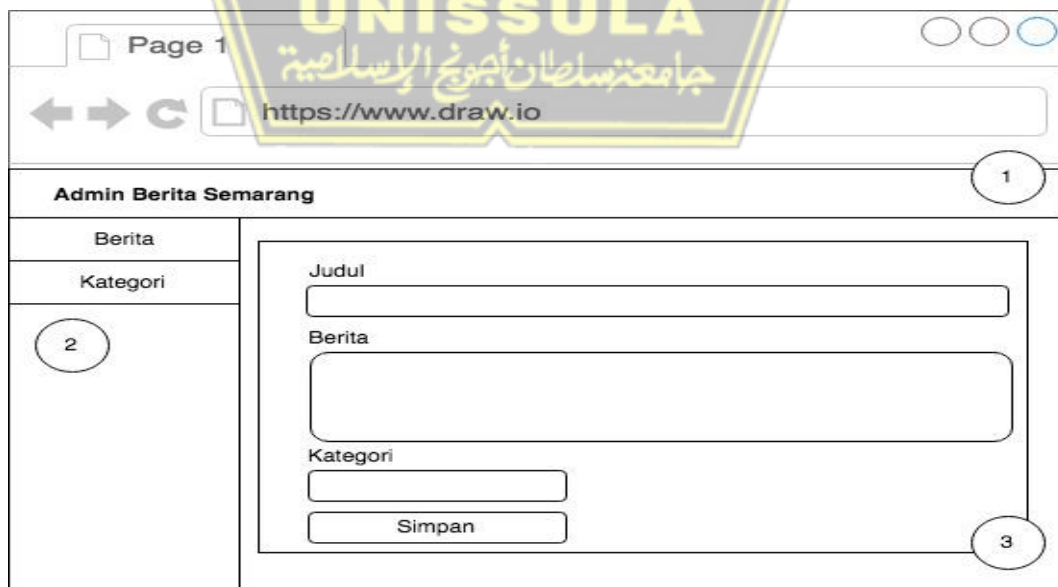


Gambar 3.6: Sequence Diagram Berita Asli

Pada gambar 3.6 ditampilkan sequence diagram user. Tahap awal dari proses adalah User mengakses aplikasi melalui browser pada URL aplikasi. Kemudian tampilan aplikasi akan mengirim request untuk mengakses berita ke control (dengan bahasa pemrograman PHP). Selanjutnya text berita article yang tersimpan di database ditampilkan pada aplikasi.

3.6 Perancangan Antarmuka

Perancangan antarmuka merupakan rancangan awal tampilan sebelum diterapkan pada bahasa pemrograman. Pada perancangan antarmuka menunjukkan tata letak komponen-komponen yang akan ditampilkan pada aplikasi.



Gambar 3.7: Sequence Diagram Berita Asli

Pada gambar 3.7 ditampilkan perancangan antarmuka pada aplikasi. Komponen-komponen aplikasi dijabarkan sebagai berikut:

1. Tampilan header yang merupakan judul dari sistem administrator.
2. Komponen menu bar yang menampilkan menu pada sistem (berita dan kategori).
3. Pada bagian tengah diisi dengan container yang dapat berisi form input atau data dalam bentuk table untuk mengelola data (berita dan kategori).



Gambar 3.8: Perancangan Antarmuka

Pada gambar 3.8 ditampilkan perancangan antarmuka pada aplikasi. Komponen-komponen aplikasi dijabarkan sebagai berikut:

1. Uniform Location Locator (URL) merupakan kolom input yang digunakan untuk mengakses ke alamat website pada browser.
2. Tampilan atas pada aplikasi merupakan header yang menampilkan informasi nama aplikasi dan menu bar.
3. Pada bagian tengah diisi dengan konten berita. Setiap berita ditampilkan pada satu kartu penampung yang berisi informasi tentang judul berita,

ringkasan dan berita asli. User dapat membaca ringkasan beserta sumber berita asli secara langsung dengan menekan tombol tab yang disediakan.

4. Bagian akhir pada aplikasi untuk menampilkan footer dari aplikasi.



BAB 4

HASIL DAN ANALISIS PENELITIAN

4.1.1 Pengujian Aplikasi

Pada tahap ini dilakukan pengujian aplikasi yang ditunjukkan sebagai berikut:

1. Halaman login admin

Halaman login merupakan halaman awal ketika admin mengakses system. Halaman login ditunjukkan pada gambar 4.1.

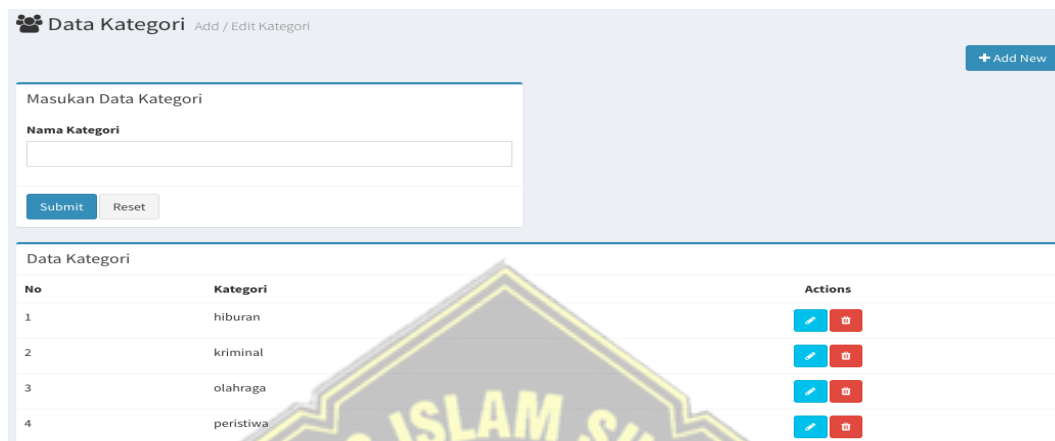


Gambar 4.1: Login Admin

Pada gambar 4.1 ditampilkan halaman login. Halaman login digunakan untuk masuk kedalam sistem administrator. Admin harus menginput username dan password untuk dapat masuk kedalam sistem.

2. Halaman form dan data kategori

Halaman kategori merupakan halaman untuk mengelola data kategori. Halaman kategori ditunjukkan pada gambar 4.2.



Gambar 4.2: Halaman Kategori

Gambar diatas ditunjukkan halaman kategori. Halaman kategori menampilkan data kategori yang tersimpan di database. Selain data pada halaman kategori ditampilkan form untuk menambah data kategori.

3. Halaman data berita

Halaman berita merupakan halaman untuk mengelola data berita. Halaman berita ditunjukkan pada gambar 4.3.

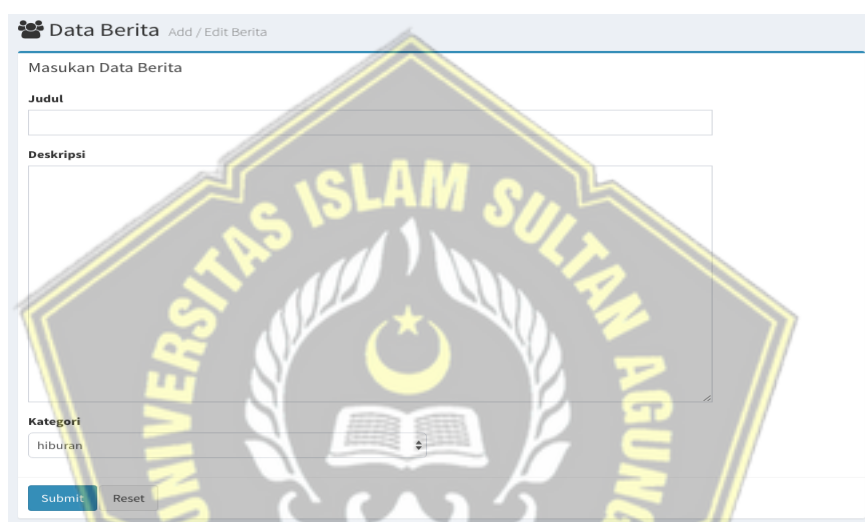


Gambar 4.3: Halaman Data Berita

Gambar diatas ditunjukkan halaman data berita. Halaman berita menampilkan data-data berita yang telah diinput dan tersimpan di database sistem. Pada table data ditampilkan informasi terkait dengan data berita dan tombol untuk mengelola berupa edit dan hapus.

4. Halaman form input berita

Halaman input berita digunakan untuk memasukan berita. Halaman berita ditunjukkan pada gambar 4.4.



The image shows a web form titled "Data Berita" with a sub-header "Add / Edit Berita". The form is titled "Masukan Data Berita" and contains the following fields:

- Judul**: A text input field.
- Deskripsi**: A large text area for entering the news content.
- Kategori**: A dropdown menu with the selected value "hiburan".

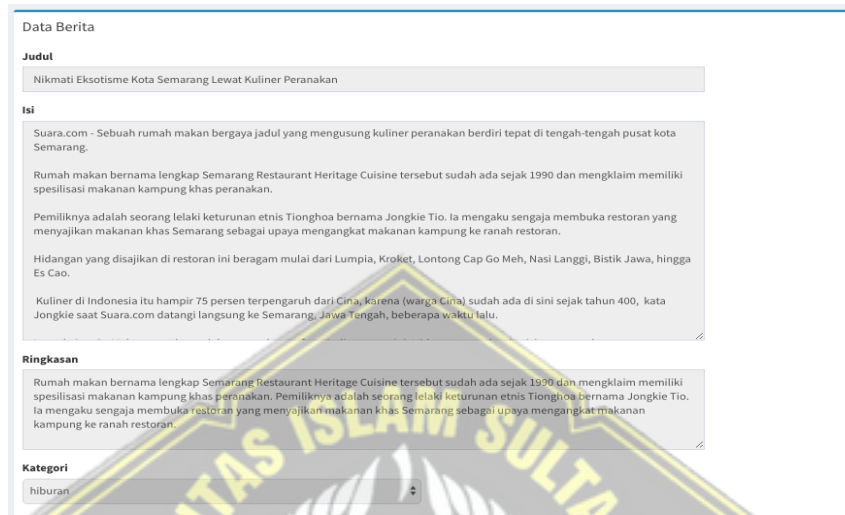
At the bottom of the form, there are two buttons: "Submit" and "Reset". A large watermark of the Universitas Islam Sultan Agung logo is overlaid on the form.

Gambar 4.4: Halaman Input Berita

Gambar diatas menunjukkan halaman input berita. Form input berita terdapat judul, deskripsi atau isi berita dan kategori. Setelah admin menginput dan menekan tombol submit berita tersebut akan diproses menjadi ringkasan dan disimpan kedalam database.

5. Halaman hasil ringkasan

Halaman ringkasan menunjukkan hasil ringkasan dari berita yang diinput. Halaman hasil ditampilkan pada gambar 4.5.

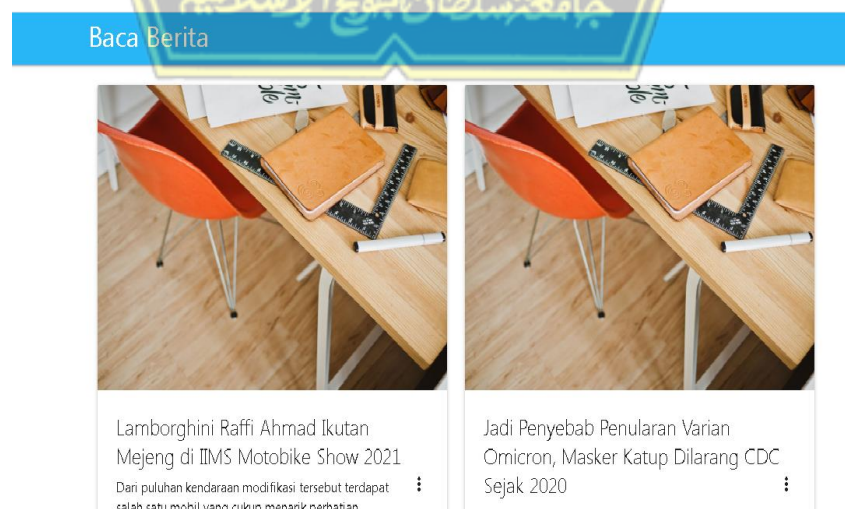


Gambar 4.5: Halaman Hasil Ringkasan

Gambar diatas menunjukkan halaman hasil input data dan hasil ringkasan. Data yang diringkas merupakan isi dari berita asli.

6. Halaman beranda user

Pada saat user mengakses aplikasi halaman pertama yang dimunculkan merupakan halaman beranda. Halaman beranda ditunjukan pada gambar 4.6



Gambar 4.6: Halaman Beranda User

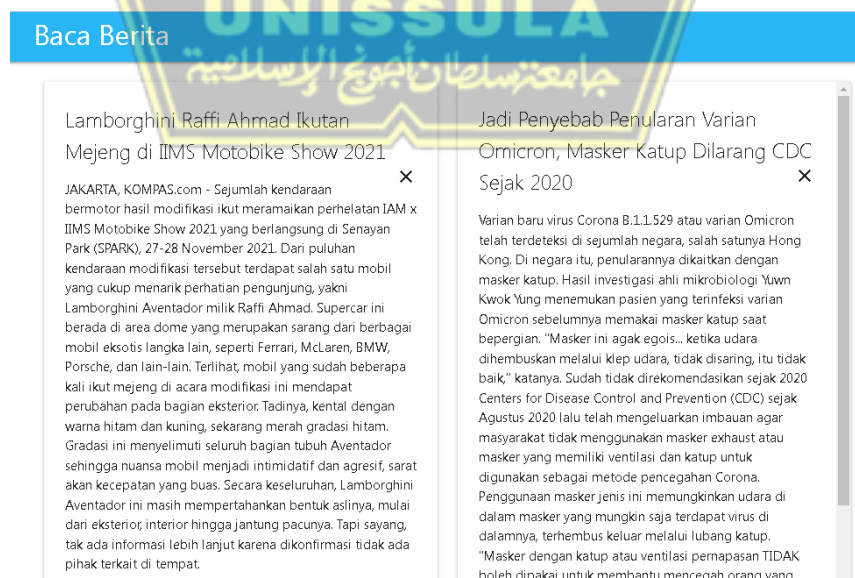
Gambar diatas menunjukkan halaman beranda user. Halaman beranda merupakan halaman awal yang ditampilkan ketika user mengakses aplikasi. Pada halaman aplikasi user terdapat menu yang dapat dilihat dengan menekan tombol pada sisi kiri atas.

7. Halaman berita dan ringkasan

Pada saat user mengakses aplikasi halaman berita ditampilkan setelah halaman beranda. Halaman berita ditunjukkan pada gambar 4.7



Gambar 4.7: Halaman Berita Ringkasan



Gambar 4.8: Halaman Detail Berita Asli

Gambar diatas menunjukkan halaman berita. Pada halaman tersebut berita ditampilkan dalam bentuk urutan secara vertical. Pada setiap berita ditampilkan informasi yaitu judul berita, ringkasan dan teks asli.

4.1.2 Pengujian Fungsional

Pengujian fungsional dilakukan untuk memastikan tanggapan/respons atas suatu event atau masukan pada fungsi yang dibuat akan menjelankan proses yang tepat dan menghasilkan keluaran/output sesuai dengan rancangan yang telah dibuat. Pengujian ditunjukan pada tabel 4.6.

Tabel 4.6: Pengujian Fungsional Admin

Input / Event	Output	Hasil
Login (username dan password) salah	Menampilkan halaman login, pesan error login	Sesuai / OK
Login (username dan password) benar	Menampilkan halaman dashboard admin	Sesuai/ OK
Klik Halaman kategori	Menampilkan data kategori	Sesuai/ OK
Menginput kategori	Menyimpan data kategori dari form	Sesuai/ OK
Klik halaman berita	Menampilkan data berita	Sesuai/ OK
Menginput data berita	Meringkas dan menyimpan berita	Sesuai/ OK

Tabel 4.7: Pengujian Fungsional User

Input / Event	Output	Hasil
Membuka aplikasi	Tampilan aplikasi	Sesuai / OK

Pilih menu berita	Menampilkan	Sesuai/ OK
Baca berita ringkasan	Menampilkan berita hasil scrapping dan ringkasan	Sesuai/ OK
Baca berita asli	Menampilkan hasil berita asli	Sesuai/ OK

4.1.3 Pengujian Ringkasan

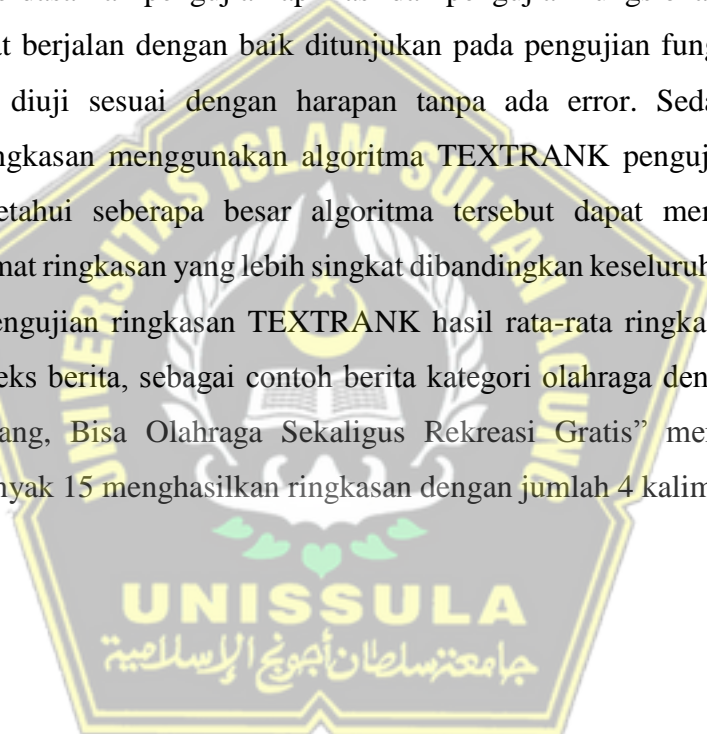
Pengujian ringkasan dilakukan untuk mengetahui hasil dari data teks berita sebelum dan sesudah diproses ringkasan. Metode pengujian yang digunakan metode compression rate yaitu dengan mengetahui prosentase jumlah kalimat hasil ringkasan dibandingkan dengan jumlah kalimat asli. Hasil dari pengujian ringkasan

Tabel 4.8: Pengujian Ringkasan

No	Kategori	Judul	Jumlah Asli	Jumlah Ringkasan	Prosentase Ringkasan (%)
1	hiburan	Jelang Sincia di Semarang: Bergaya Naik Motor Sespan di Kota Lama	24	3	12.50
2	hiburan	Nikmati Eksotisme Kota Semarang Lewat Kuliner Peranakan	33	4	12.12
3	kriminal	Rumah Pak RT Terduga Teroris Tak Jauh dari Kediaman Wali Kota Semarang	26	4	15.38
4	kriminal	Caleg Partai Gerindra Ditangkap saat Isap Sabu di Rumah Relawan	7	5	71.43
5	olahraga	41 Crosser Dunia Ramaikan MXGP Seri 13 Semarang	17	5	29.41
6	olahraga	Di Trilomba Juang, Bisa Olahraga Sekaligus Rekreasi Gratis	15	4	26.67

7	peristiwa	Sebanyak 97 Rumah di Kanal Banjir Timur Semarang Digusur	26	4	15.38
8	peristiwa	Pemungutan Suara Ulang di TPS Kota Semarang Ini Minim Peminat	15	5	33.33
Rata-rata (%)					27

Pengujian telah dilakukan dengan menguji aplikasi, fungsional dan pengujian ringkasan. Berdasarkan pengujian aplikasi dan pengujian fungsional telah diketahui aplikasi dapat berjalan dengan baik ditunjukkan pada pengujian fungsional semua fungsi yang diuji sesuai dengan harapan tanpa ada error. Sedangkan untuk pengujian ringkasan menggunakan algoritma TEXTRANK pengujian dilakukan untuk mengetahui seberapa besar algoritma tersebut dapat meringkas berita menjadi kalimat ringkasan yang lebih singkat dibandingkan keseluruhan teks berita. Hasil dari pengujian ringkasan TEXTRANK hasil rata-rata ringkasan adalah 27 persen dari teks berita, sebagai contoh berita kategori olahraga dengan judul “Di Trilomba Juang, Bisa Olahraga Sekaligus Rekreasi Gratis” memiliki jumlah kalimat sebanyak 15 menghasilkan ringkasan dengan jumlah 4 kalimat.



BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari hasil penelitian dan pengujian Aplikasi peringkasan teks otomatis dengan metode TEXTRANK, pada akhir laporan penulis dapat memberikan kesimpulan yaitu Metode yang diusulkan yaitu TEXTRANK telah berhasil diimplementasikan untuk peringkasan teks otomatis pada aplikasi berita online sekitar kota Semarang. Berdasarkan hasil pengujian menggunakan tingkat kompresi penerapan metode TEXTRANK mempunyai performa ringkasan yang mampu meringkas teks hingga 27 persen dari keseluruhan teks berita. Aplikasi yang dibuat telah diuji menggunakan pengujian blackbox yang menguji keseluruhan fungsi, hasil dari pengujian menunjukkan fungsi-fungsi yang terdapat pada aplikasi berjalan dengan baik dan memiliki output yang sesuai.

5.2 Saran

Berdasarkan kesimpulan dan analisis laporan, saran dari peneliti untuk penelitian lebih lanjut yaitu Penelitian lebih lanjut dapat menerapkan dan melakukan optimasi terhadap metode peringkasan teks otomatis dengan melakukan penambahan metode dalam proses peringkasan. Dikembangkan penelitian yang lebih mendalam dan variasi algoritma untuk lain seperti text clustering, kecerdasan buatan dan lain-lain.

DAFTAR PUSTAKA

- A.S Rossa, S. ..., 2011. *Rekayasa Perangkat Lunak*. Bandung: Modula.
- A., 2018. Evaluasi Algoritma TextRank Pada Peringkasan Teks Berbahasa Indonesia.
- Barus, S. W., 2016. *Jurnalistik: Petunjuk Teknis Menulis Berita*. In: s.l.:s.n.
- Budianto, I., 2016. Rancang Bangun Sistem Peringkasan Teks Otomatis Artikel Berbahasa Indonesia Menggunakan Metode TF-IDF (Term Frequency-Inversed Document Frequency).
- Collobert, R., 2011. Natural language processing (almost) from scratch. *Journal of Machine Learning Research* 12.Aug.
- E., Christianti, V. & Pragantha, J., 2017. Penerapan Algoritma TEXTRANK untuk Automatic Summarization pada Dokumen Berbahasa Indonesia.
- Kurniawan, B., 2012. Klasifikasi Konten Berita Dengan Metode Text Mining.
- Laksana, J. & Purwarianti, A., 2014. *Indonesian Twitter Text Authority Classification For Government in Bandung*. s.l., IEEE.
- Moreno, T. & Manuel, J., 2014. Automatic text summarization.
- Mubarak, A., 2019. Rancang Bangun Aplikasi Web Sekolah Menggunakan Uml (Unified Modeling Language) Dan Bahasa Pemrograman Php (Php Hypertext Preprocessor) Berorientasi Objek. *JIKO (Jurnal Informatika dan Komputer)*.
- Mustaqhfi, M., 2012. Peringkasan Teks Otomatis Berita Berbahasa Indonesia Menggunakan Metode Maximum Marginal Relevance.
- Nugraha, N., 2016. Automatic Summarization Bahasa Indonesia Menggunakan TextRank.
- Pragantha, J. & Eris, E., 2017. Penerapan Algoritma TextRank untuk Automatic Summarization pada Dokumen Berbahasa Indonesia. *Jurnal Ilmu Teknik dan Komputer* 1.1.

- Pressman, R. S., 2005. *Software Engineering*. Somerville: s.n.
- Sovia, R., 2017. Membangun Aplikasi E-Library Menggunakan Html, Php Script, Dan Mysql Database.. *Jurnal Processor*.
- Sutami, D. P., 2020. Pendefinisian objek dan pelengkap dalam kamus besar bahasa indonesia sebagai penunjang literasi. In: *Linguistik Indonesia*. s.l.:s.n.
- Tujni, B. & Hutrianto, H., 2020. Pengembangan Perangkat Lunak Monitoring Wellies Dengan Metode Waterfall Model.. *Jurnal Ilmiah Matrik 22.1*.
- Weiss, S. M., 2010. *Text mining: predictive methods for analyzing unstructured information*.. s.l.:Springer Science & Business Media.
- Widiantoro, A., 2014. Peringkasan Teks Otomatis Pada Dokumen Berbahasa Jawa Menggunakan Metode Tf-Idf.
- Zhai, C. & Lafferty, J., 2017. A study of smoothing methods for language models applied to ad hoc information retrieval. *ACM SIGIR Forum*.

