

**EVALUASI ALGORITMA *TEXTRANK* PADA PERINGKASAN TEKS  
BERBAHASA INDONESIA**

**SKRIPSI**

**ABDURROHMAN**

**121402104**



**PROGRAM STUDI TEKNOLOGI INFORMASI  
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI  
UNIVERSITAS SUMATERA UTARA**

**MEDAN**

**2018**

EVALUASI ALGORITMA *TEXTRANK* PADA PERINGKASAN TEKS BERBAHASA

INDONESIA

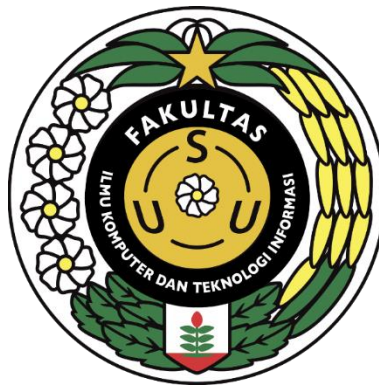
SKRIPSI

Diajukan untuk melengkapi tugas dan memenuhi syarat memperoleh ijazah

Sarjana Teknologi Informasi

ABDURROHMAN

121402104



PROGRAM STUDI TEKNOLOGI INFORMASI

FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UNIVERSITAS SUMATERA UTARA

MEDAN

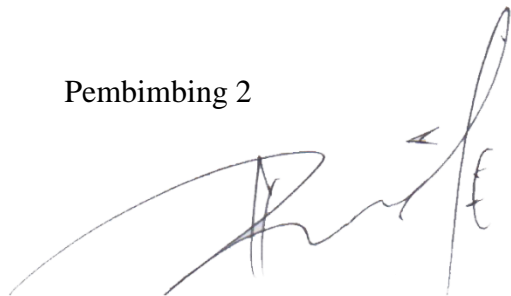
2018

**PERSETUJUAN**

Judul : EVALUASI ALGORITMA *TEXTRANK* PADA  
PERINGKASAN TEKS BERBAHASA INDONESIA  
Kategori : SKRIPSI  
Nama : ABDURROHMAN  
Nomor Induk Mahasiswa : 121402104  
Program Studi : SARJANA (S1) TEKNOLOGI INFORMASI  
Departemen : TEKNOLOGI INFORMASI  
Fakultas : FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI  
INFORMASI

Komisi Pembimbing :

Pembimbing 2



Romi Fadillah Rahmat, B.Comp.Sc., M.Sc  
NIP. 19860303 201012 1 004

Pembimbing 1

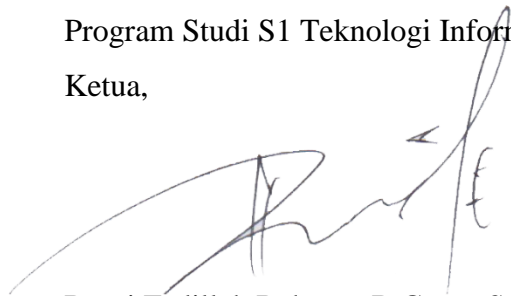


Dani Gunawan, ST., M.T  
NIP. 19820915 201212 1 002

Diketahui/disetujui oleh

Program Studi S1 Teknologi Informasi

Ketua,



Romi Fadillah Rahmat, B.Comp.Sc., M.Sc  
NIP. 19860303 201012 1 004

**PERNYATAAN****EVALUASI ALGORITMA *TEXTRANK* PADA PERINGKASAN TEKS BERBAHASA  
INDONESIA****SKRIPSI**

Saya mengakui bahwa skripsi ini adalah hasil karya saya sendiri, kecuali beberapa kutipan dan ringkasan yang masing – masing telah disebutkan sumbernya.

Medan, 27 Juli 2018



Abdurrohman

121402104

## UCAPAN TERIMA KASIH

Puji dan syukur penulis sampaikan kehadiran Allah SWT yang telah memberikan rahmat dan izin-Nya sehingga penulis dapat menyelesaikan skripsi ini sebagai syarat untuk memperoleh gelar sarjana Komputer pada Program Studi S1 Teknologi Informasi Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Sumatera Utara. Selama dalam penyelesaian tugas akhir ini, banyak sekali bantuan dan dukungan serta doa dari berbagai pihak sehingga penulis ingin menyampaikan ucapan terima kasih sedalam-dalamnya dan penghargaan kepada :

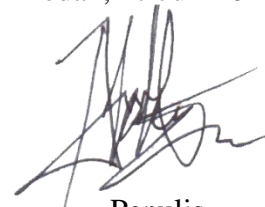
1. Kedua orang tua penulis, Ayahanda Chaerudin, SMI. dan Ibunda Desmawati, S.Pd. yang telah membesarkan penulis dengan sabar dan penuh kasih sayang serta memberikan doa dan dukungan moril maupun materil serta memberikan motivasi terbesar kepada penulis sehingga penulis dapat memberikan yang terbaik untuk menyelesaikan skripsi ini. Serta kepada seluruh anggota keluarga penulis yang disayangi abang Hudzaifah, ST., Muhammad Hilmi, Hanina Fauziah, Muhammad Mush'ab, Salsabila Roshada, dan Ahmad Hamzah.
2. Bapak Dani Gunawan, ST. M.T. selaku pembimbing pertama dan Bapak Romi Fadillah Rahmat, B.Comp.Sc, M.Sc. selaku pembimbing kedua yang telah meluangkan waktu dan pikiran beliau untuk membimbing penulis dalam menyelesaikan penelitian dan skripsi ini. Tanpa motivasi dan arahan yang diberikan kedua pembimbing, tentu penulis tidak akan dapat menyelesaikan skripsi ini dengan baik.
3. Ibu Sarah Purnamawati, ST., M.Sc. sebagai dosen pembimbing pertama dan Bapak Ivan Jaya, S.Si., M.Kom. sebagai dosen pembimbing kedua yang telah memberikan masukan dan kritikan yang membangun dan bermanfaat dalam penulisan skripsi ini.
4. Semua dosen serta pegawai dilingkungan Fakultas Ilmu Komputer dan Teknologi Informasi yang telah membantu dan membimbing penulis selama proses perkuliahan.
5. Sahabat penulis abangda Zuwarbi Wiranda Hsb, S.Kom, abangda Abdi Pratama, S.Kom, Arsandi Saputra, S.Kom, Muhammad Yusuf, S.Kom, Rasyid Sihombing,

Furqan Alatas, S.Kom, Joko Kurnianto, S.Kom, Andika Mulia Utama, S.Kom, Rifai Muda Harahap, S.Hut, dr.Dedi Kurniawan, Rahmad Rahandi, S.E, Muhammad Rizal, S.K.M., Adanan Purba, S.Pt, Hafizan, S.Farm.,Apt, Fazrin Ashari, Agung Hernowo, Bagus Kusumo dan Imam Kurniawan yang telah memberikan motivasi terbaik sehingga penulis dapat menyelesaikan skripsi dengan lancar.

6. Teman – teman Bung Tomo, Sultan USU, Baiti Jannati, Kost31, JOSH, Sarang Komodo yang telah memberikan pengalaman terbaik selama belajar di Universitas Sumatera Utara.
7. Seluruh jajaran Rumah Kepemimpinan Regional Medan selaku pembina Bapak Dr. Rudi Hartono, S.Hut, M.Si selaku manager Bang Andi Pranata, S.Si. selaku supervisor Bang Abdullah Nasution, S.Si. yang telah memberikan pelatihan dan bimbingan selama menjadi mahasiswa penerima beasiswa Rumah Kepemimpinan.
8. Teman – teman dan sahabat seperjuangan mahasiswa Teknologi Informasi USU yang telah memberikan dukungan dan bantuan selama masa perkuliahan yang Namanya tidak dapat penulis sebutkan satu persatu.

Semoga Allah SWT melimpahkan berkah kepada semua pihak yang telah memberikan bantuan, perhatian serta dukungan dan motivasinya kepada penulis dalam menyelesaikan skripsi ini.

Medan, 27 Juli 2018



Penulis

## ABSTRAK

Peringkasan teks otomatis biasa digunakan untuk meringkas sebuah teks menggunakan komputer untuk mendapatkan ringkasan dari teks tersebut. Metode peringkasan teks otomatis menggunakan ekstraktif dikarenakan metode ini mengambil informasi penting dari sebuah teks tanpa mengubah atau memodifikasi informasi tersebut. Salah satu algoritma yang dapat digunakan untuk meringkas teks otomatis adalah dengan menggunakan algoritma TextRank. Keunggulan dari algoritma TextRank yaitu tidak memerlukan pengetahuan mendalam tentang sebuah bahasa dan tidak memerlukan data latih untuk dapat meringkas teks. Cara kerja algoritma ini adalah dengan merepresentasikan kalimat pada teks ke dalam graf, lalu menghitung nilai tiap kalimat dengan menggunakan kesamaan (*similarity*) antar kalimat untuk menentukan hasil ringkasan. Selain menggunakan *similarity* untuk menentukan kalimat penting, pada penelitian ini juga menggunakan TextRank modifikasi yaitu dengan menggunakan *levenshtein distance* untuk menghitung ringkasan dengan membandingkan kemiripan antar string dengan cara memasukkan, menghapus, atau mengganti karakter string. Peringkasan teks otomatis menggunakan TextRank dilakukan dengan meringkas 100 teks berbahasa Indonesia yang selanjutnya akan dievaluasi menggunakan *ROUGE*. Evaluasi *ROUGE* bekerja dengan membandingkan hasil ringkasan dari TextRank dengan ringkasan manual oleh ahli pada bidang bahasa Indonesia. Hasil yang didapat algoritma TextRank mendapatkan rata – rata nilai *F-Score* 0.439 pada *ROUGE-1* dan 0,3186 pada *ROUGE-2*. Sedangkan TextRank modifikasi mendapatkan rata – rata nilai *F-Score* 0,3999 pada *ROUGE-1* dan 0,2805 pada *ROUGE-2*. Jika dibandingkan dengan hasil ringkasan dengan algoritma TextRank berbahasa Inggris yang mendapatkan *F-Score* rata – rata 0,4708 maka hasilnya tidak jauh berbeda. Artinya algoritma TextRank dapat digunakan untuk meringkas teks berbahasa Indonesia.

**Kata Kunci :** Peringkasan Teks Otomatis, *Textrank*, *Sentence Extraction*, *Levenshtein Distance*, *ROUGE*

## EVALUATION TEXTRANK ALGORITHM IN INDONESIAN TEXT SUMMARY

### ABSTRACT

Automatic text summaries are commonly used to summarize text using computer program to get a summary of the text. Usually automatic text summary use extractives method because this method retrieves important information from a text without altering or modifying the information. One of the algorithms that can be used to automatic text summarize is TextRank algorithm. The advantage of using TextRank algorithm is it does not require depth knowledge of a language and does not require data training to be able summarize text. This algorithm works with represent the sentence in text into the graph, then calculate the value of each sentences using similarity to determine important sentences, in this research we also use TextRank modification that is using *levenshtein distance* to calculate summary by comparing similarities between strings by using enter, delete, or replacing string characters. Automatic text summaries using TextRank is done by summarizing 100 Indonesian texts which will be evaluated using ROUGE. ROUGE evaluation works by comparing summary results from TextRank with manual summaries by experts Indonesian linguist. The result obtained by the TextRank algorithm get an average *F-Score* of 0,439 on ROUGE-1 and 0,3186 on ROUGE-2. While TextRank modification gets an average *F-Score* 0,3999 on ROUGE-1 and 0,2805 on ROUGE-2. When compared with the results of a summarize English text using TextRank algorithm that gets an average *F-Score* of 0,4708, the results are not much different. This mean that TextRank algorithm can be used to summarize Indonesian text.

**Kata Kunci :** *Text Summarization, Textrank, Levenshtein Distance, Sentence Extraction, ROUGE*



## DAFTAR ISI

	Hal.
PERSETUJUAN	i
PERNYATAAN	ii
UCAPAN TERIMA KASIH	iii
ABSTRAK	v
<i>ABSTRACT</i>	vi
DAFTAR ISI	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
 BAB 1 PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah	3
1.5 Manfaat Penelitian	3
1.6 Metodologi Penelitian	3
1.7 Sistematikan Penulisan	4
 BAB 2 LANDASAN TEORI	
2.1 Bahasa Indonesia	6
2.2 Berita	7
2.3 Peringkasan Teks Otomatis	8
2.4 Algoritma <i>TextRank</i>	8
2.5 <i>Levenshtein Distance</i>	11
2.6 <i>ROUGE</i>	12

2.6.1 <i>ROUGE -N</i>	13
2.7 Penelitian Terdahulu	13
BAB 3 ANALISA DAN PERANCANGAN SISTEM	
3.1 Data	16
3.2 Arsitektur Umum	16
3.2.1 Input	18
3.2.2 Proses	19
3.2.3 Output	32
3.3 Perancangan Sistem	33
3.3.1 Diagram Aktifitas untuk Peringkasan Teks Otomatis	33
3.3.2 Perancangan Antarmuka Sistem	35
3.3 Evaluasi	37
BAB 4 IMPLEMENTASI DAN PENGUJIAN SISTEM	
4.1 Implementasi Sistem	38
4.1.1 Spesifikasi Hardware dan Software yang digunakan	38
4.1.2 Pengambilan Survei	38
4.1.3 Implementasi Perancangan Antarmuka	39
4.2 Pengujian Sistem	39
4.3 Evaluasi Hasil Ringkasan	40
BAB 5 KESIMPULAN DAN SARAN	
5.1 Kesimpulan	44
5.2 Saran	44
DAFTAR PUSTAKA	45

## DAFTAR TABEL

	Hal.
Tabel 2.1 Hasil Pengujian Algoritma TextRank dengan sistem lain	10
Tabel 2.2 Representasi Matriks untuk Menghitung Jarak Antar String	11
Tabel 2.3 Penelitian Terdahulu	14
Tabel 3.1 Tahapan <i>Sentence Splitting</i>	20
Tabel 3.2 Tahapan <i>Case Folding</i>	23
Tabel 3.3 Tahapan <i>Tokenizing</i>	24
Tabel 3.4 <i>Word Overlap</i> pada vertex ke-7 dan vertex ke-12	29
Tabel 3.5 <i>Word Overlap</i> pada vertex ke-7 dan <i>vertex</i> ke-2	30
Tabel 3.6 Contoh hasil <i>scoring</i> pada tiap <i>vertex</i>	31
Tabel 3.7 Contoh pemringkatan kalimat	32
Tabel 4.1 Hasil Peringkasan Teks Otomatis	41
Tabel 4.2 Hasil Evaluasi Peringkasan ROUGE-1 <i>Textrank</i> (a) dan <i>Textrank Levenshtein Distance</i> (b)	42
Tabel 4.3 Hasil Evaluasi Peringkasan ROUGE-2 <i>Textrank</i> (a) dan <i>Textrank Levenshtein Distance</i> (b)	43
Tabel 4.4 Hasil Rata – Rata Evaluasi <i>F-Score</i>	45

## DAFTAR GAMBAR

	Hal.
Gambar 3.1 Arsitektur Umum	17
Gambar 3.2 Contoh Teks Berbahasa Indonesia	18
Gambar 3.2 Contoh Teks Berbahasa Indonesia ( <i>Lanjutan</i> )	19
Gambar 3.3 <i>Text Cleaning</i>	21
Gambar 3.3 <i>Text Cleaning (Lanjutan)</i>	22
Gambar 3.4 Representasi Graf	28
Gambar 3.5 Contoh Output Peringkasan Menggunakan <i>Textrank</i>	33
Gambar 3.6 Diagram Aktivitas Peringkasan Teks Otomatis	34
Gambar 3.7 Rancangan Tampilan Halaman Survei	35
Gambar 3.8 Rancangan Tampilan Halaman Program	36
Gambar 4.1 Tampilan Halaman Survei	39
Gambar 4.2 Tampilan Halaman Utama	40

## BAB 1

### PENDAHULUAN

#### 1.1 Latar Belakang

Tren *Bring Your Own Device (BYOD)* menjadi salah satu *lifestyle* di kota – kota maju dimana perusahaan memperbolehkan pegawainya untuk membawa *gadget* mereka. Dengan tren tersebut setiap orang dapat dengan mudah mendapatkan akses informasi. Sumber informasi yang didapat juga beragam mulai dari media sosial, forum diskusi, *website*, berita *online* dan masih banyak lagi. Salah satu sumber informasi yang sangat diminati saat ini adalah berita *online* dikarenakan informasi yang diberikan haruslah berdasarkan fakta, aktual, lengkap, akurat, menarik, berimbang, dan mudah dipahami.

Berita yang muncul pada berita *online* sangatlah banyak sehingga memerlukan waktu yang lebih untuk dapat memahami isi berita. Salah satu cara untuk memudahkan dalam memahami isi berita adalah dengan menggunakan peringkasan teks otomatis (*automatic text summarization*). Peringkasan teks adalah mengutip bagian penting dari satu atau lebih teks untuk menghasilkan teks yang lebih singkat yang berukuran kurang dari setengah teks asli dan tetap terdiri dari beberapa informasi penting yang berasal dari teks asli. Dengan peringkasan teks otomatis dapat menjadi alternatif untuk dapat memahami isi berita dengan waktu yang cepat.

Pada tahun 2012 Jayashree R, Basavaraj Anami, dan Srikantha Murthy K pernah membuat sistem peringkasan teks dimana sistem tersebut dapat mengategorikan teks dalam bahasa Kannada. Penelitian tersebut dilakukan dengan menggunakan algoritma GSS (Galavotti, Sebastiani, Simi) digabungkan dengan tf-idf. Sistem tersebut bekerja dengan cara mengekstraksi kata kunci dan tulisan untuk digunakan pada peringkasan teks. Dimana kata kunci yang diambil haruslah memiliki bobot atau hubungan yang tinggi antara satu dengan yang lain. Hasil yang didapat dari penelitian tersebut menunjukkan peringkasan teks yang dilakukan oleh manusia lebih efektif dibandingkan dengan komputer hal tersebut terjadi karena peringkasan teks yang dilakukan manual oleh manusia memiliki koheren antar kalimat lebih tinggi antar kalimat dibandingkan dengan komputer.

Nguyen Quang, Pham Tuan Anh, Truong Cong, Nguyen Xuan dari Vietnam pada tahun 2012 juga melakukan peringkasan teks otomatis. Dimana mereka menggunakan *genetic programming* untuk mengekstraksi kalimat penting pada teks berbahasa Vietnam. *Genetic programming* memiliki 2 tahapan yaitu *training* dan *testing*. Tujuan dari *training* adalah melatih sistem untuk meningkatkan kemampuan menemukan kalimat penting dalam teks. Sedangkan pada tahapan *testing* data yang paling bagus dari *training* akan digunakan untuk meningkatkan hasil peringkasan. Penelitian ini bisa dibilang cukup sukses dikarenakan dari proses evaluasi yang dilakukan dengan membandingkan antara kinerja peringkasan teks oleh ahli bidang linguistik, siswa, dan *genetic programming* hasil dari sistem 20% lebih baik dari yang dilakukan oleh siswa, walaupun sistem memiliki hasil lebih rendah dari ahli bidang linguistik.

Tidak hanya peringkasan teks berbahasa asing penelitian tentang peringkasan teks otomatis juga dilakukan pada teks berbahasa Indonesia oleh Anwar Pasaribu pada tahun 2016. Penelitian dilakukan dengan menggunakan algoritma *TextTeaser*. Cara kerja dari algoritma ini adalah dengan memanfaatkan judul teks, posisi kalimat, panjang kalimat, dan frekuensi *keyword* untuk menentukan apakah kalimat termasuk ringkasan atau tidak. Hasil yang didapat bahwa 60,11% dari data yang diuji memiliki F-Score  $\geq 0.5$  yang artinya *TextTeaser* memiliki tingkat presisi yang cukup tinggi.

Berdasarkan latar belakang ini maka penulis akan mengajukan penelitian dengan judul “EVALUASI ALGORITMA *TEXTRANK* PADA PERINGKASAN TEKS OTOMATIS BERBAHASA INDONESIA”.

## 1.2 Rumusan Masalah

Kalimat penting dalam sebuah paragraf biasanya mewakili keseluruhan isi paragraf. Untuk menemukan gagasan utama dalam sebuah teks biasanya dapat dilakukan dengan menemukan kalimat penting dalam setiap paragraf. Dengan memanfaatkan sistem untuk meringkas teks maka akan mempercepat proses mendapatkan informasi penting. *TextRank* merupakan salah satu algoritma yang sangat berguna untuk digunakan pada peringkasan teks otomatis karena algoritma ini bekerja tanpa perlu menggunakan data training sehingga dapat diaplikasikan untuk berbagai bahasa tanpa perlu mengetahui secara mendalam bahasa yang digunakan (Mihalcea, 2005). Penelitian mengenai algoritma *TextRank* yang dilakukan Mihalcea pada tahun 2005 menggunakan teks berbahasa Inggris, sehingga diperlukan evaluasi algoritma tersebut untuk dapat

mengetahui seberapa efektif algoritma *TextRank* untuk meringkas teks berbahasa Indonesia.

### 1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk mengevaluasi hasil peringkasan dengan algoritma *TextRank* pada teks berbahasa Indonesia.

### 1.4 Batasan Masalah

- Data yang digunakan sebagai data uji adalah teks berita yang diambil dari berbagai website berita online berbahasa Indonesia.
- Teks berita yang dipilih hanya berdasarkan isi berita.
- Kategori teks berita yang diambil antara lain bidang teknologi, kesehatan, umum, olahraga, ekonomi, gaya hidup, nasional, sains, *travel*, otomotif.
- Peringkasan yang dilakukan dengan metode ekstraktif.
- Peringkasan yang dilakukan adalah dengan meringkas teks berbahasa Indonesia berdasarkan kalimat.
- Evaluasi dilakukan dengan membandingkan hasil ringkasan algoritma *TextRank* dengan hasil ringkasan dari ahli menggunakan *ROUGE*.

### 1.5 Manfaat Penelitian

- Mengembangkan sistem yang dapat digunakan untuk meringkas teks berita berbahasa Indonesia.
- Mempermudah memahami isi teks berita.
- Mengevaluasi algoritma *TextRank* dalam meringkas teks berbahasa Indonesia.
- Mengetahui akurasi algoritma *TextRank* dalam meringkas teks berbahasa Indonesia

### 1.6 Metodologi Penelitian

Dalam menyelesaikan penelitian ini ada beberapa tahapan yang dilakukan oleh penulis. Berikut adalah tahapan penelitian yang dilakukan :

### 1. Studi Literatur

Tahap ini penulis mencari informasi terkait dengan masalah yang akan diteliti dengan cara mengumpulkan beberapa referensi yang diperoleh dari berbagai jurnal, penelitian, *paper*, buku, skripsi dan sumber informasi lainnya.

### 2. Analisis dan Pengumpulan Data

Pada tahapan ini dilakukan analisis dari berbagai informasi mengenai peringkasan teks otomatis dengan hanya memilih berdasarkan kalimat. Salah satu metode yang dapat digunakan dalam meringkas teks berdasarkan kalimat adalah dengan menggunakan algoritma *TextRank*. Selanjutnya dilakukan pengumpulan data berupa teks berita yang didapat dari berbagai situs berita *online* untuk diuji dalam sistem.

### 3. Perancangan Sistem

Tahap ini dilakukan perancangan sistem berupa arsitektur umum, perancangan data, dan tampilan antarmuka sistem.

### 4. Implementasi Sistem

Tahapan selanjutnya yaitu pembuatan perancangan sistem kedalam kode program sistem aplikasi untuk menggunakan bahasa pemrograman *Python* sebagai *back-end* untuk memproses algoritma *TextRank* dan menampilkannya kedalam tampilan *website* sebagai *front-end*.

### 5. Pengujian Sistem

Pada tahapan ini dilakukan pengujian terhadap sistem yang telah dibangun untuk menemukan kelemahan ataupun kesalahan yang dapat diperbaiki dan menguji sistem yang telah dibuat dengan hasil manual ekstraksi kata kunci dari ahli yang telah dipilih.

### 6. Dokumentasi dan Penyusunan Laporan

Tahapan akhir ini dilakukan dokumentasi terhadap sistem yang dibuat dan penyusunan dari hasil akhir penelitian yang dilakukan.

## 1.7 Sistematika Penulisan

Sistematika penulisan skripsi ini dibagi menjadi beberapa bagian masing – masing bagian dijelaskan secara singkat sebagai berikut:



**BAB 1 : Pendahuluan**

Pada bab ini akan dijelaskan mengenai latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian, manfaat, metodologi penelitian serta sistematika penulisan

**BAB 2 : Landasan Teori**

Bab ini berisi landasan teori yang digunakan untuk mendukung penelitian yang dilakukan agar lebih mudah dalam memahami dan mendapatkan solusi terhadap permasalahan yang ada.

**BAB 3 : Analisis dan Perancangan Sistem**

Dalam bab ini berisi analisis terkait algoritma *TextRank* dan perancangan sistem yang akan dibuat mulai dari arsitektur umum, rancangan program, dan antar muka aplikasi.

**BAB 4 : Implementasi dan Pengujian Sistem**

Pada bab ini berisi tentang implementasi algoritma ke dalam sistem untuk menyelesaikan permasalahan yang menjadi tujuan utama penelitian. Serta dilanjutkan dengan pengujian sistem untuk menemukan kelemahan ataupun kesalahan yang terjadi pada sistem sehingga dapat diperbaiki.

**BAB 5 : Kesimpulan dan Saran**

Bab ini berisikan kesimpulan berdasarkan penelitian yang dilakukan serta saran yang bertujuan memberikan masukan untuk pengembangan penelitian lebih lanjut.

## BAB 2

### LANDASAN TEORI

#### 2.1 Bahasa Indonesia

Bahasa Indonesia yang pada mulanya berasal dari bahasa Melayu dijadikan sebagai bahasa persatuan bangsa Indonesia. Dari sudut pandang linguistik, Bahasa Indonesia adalah salah satu dari banyak ragam bahasa Melayu. Hingga saat ini, Bahasa Indonesia adalah bahasa yang hidup, dan terus menghasilkan kata-kata baru, baik melalui penciptaan maupun penyerapan dari bahasa daerah dan bahasa asing.

Fokus dari penelitian ini adalah untuk mengetahui apakah algoritma *TextRank* efektif untuk meringkas teks berbahasa Indonesia, dikarenakan bahasa Inggris dan bahasa Indonesia sangatlah berbeda dalam struktur, antara lain :

a. Perbedaan waktu penggunaan

Waktu penggunaan yang berbeda merupakan salah satu contoh perbedaan mendasar dari bahasa Indonesia dan bahasa Inggris. Bahasa Indonesia tidak memiliki aturan khusus untuk kapan pemakaian suatu bahasa, namun bahasa Inggris memilikinya.

b. Perbedaan mengenai kata kerja

Dalam hal kata kerja, Bahasa Indonesia tidak memiliki perbedaan, lampau, sekarang atau kapanpun waktunya tetap menggunakan kata kerja yang sama. Hal tersebut berbeda dengan bahasa Inggris, yang memiliki perbedaan kata kerja bergantung dari waktu terjadinya dan masuk kedalam jenis kata tersebut, regular ataupun irregular verb.

c. Perbedaan bunyi atau cara baca

Bahasa Indonesia dan bahasa Inggris memiliki bunyi yang berbeda atau pengejaan yang berbeda. Jika bahasa Indonesia memiliki sistem pengejaan yang sama dengan apa yang di tulis, maka bahasa Inggris belum tentu demikian, apa yang di tulis belum tentu sama dengan cara mengejanya.

d. Perbedaan perubahan kata ganti

Bahasa Indonesia, tidak mengalami perubahan kata ganti, baik ia sebagai subjek ataupun sebagai objek. Hal tersebut berbeda dengan bahasa Inggris, yang memiliki perubahan pada kata ganti orang ketiga tunggal nya saat ia menjadi subjek ataupun objek.

e. Perbedaan susunan kalimat

Antara Bahasa Indonesia dan bahasa Inggris memiliki susunan kalimat yang berbeda, perhatikan contoh dibawah ini:

- Gadis cantik itu membeli sayuran di pasar.
- *That beautiful girl buys vegetable in the market.*

## 2.2 Berita

Salah satu cara untuk mendapatkan informasi mengenai sesuatu yang sedang terjadi adalah dengan memanfaatkan berita. Dikarenakan berita sangat mudah untuk diakses karena berita disajikan lewat bentuk cetakan, siaran, internet, atau dari mulut ke mulut kepada orang ketiga sehingga mudah untuk didapatkan. Menurut Uskali (2007) berita sebagai media memiliki peran fungsi sebagai jembatan komunikasi antar berbagai perbedaan budaya dan pola pikir yang menyebar di masyarakat. Saat ini pemberitaan di seluruh dunia semakin rumit atas berbagai kompleks masalah yang semakin melebar mulai dari isu SARA, *hoax* dan lain sebagainya, maka media berita dituntut untuk semakin bekerja secara profesional dengan menjaga kualitas berita yang disampaikan. Berikut adalah beberapa syarat agar berita yang disampaikan terjaga kualitasnya :

1. Berita harus merupakan fakta
2. Berita harus merupakan aktual (kejadian terkini)
3. Berita harus seimbang (tidak memihak)
4. Berita harus lengkap (mengandung semua unsur berita)
5. Berita harus menarik dan bermanfaat
6. Berita harus disusun dengan sistematis

Sedangkan dalam penyusunan sebuah berita maka perlu diperhatikan unsur yang harus dipenuhi agar berita tersebut dapat dikatakan lengkap. Syarat berita dikatakan lengkap adalah jika sebuah memiliki 5 W + 1 H, yaitu :

- What, yaitu berita harus memuat mengenai apa yang sedang terjadi
- Who, yaitu berita harus memuat informasi mengenai siapa saja yang terlibat dalam kejadian tersebut
- Why, yaitu berita harus memuat latar belakang kejadian tersebut terjadi
- When, yaitu berita harus memuat informasi mengenai kapan yang diberikan terjadi
- Where, yaitu berita harus memberikan informasi tempat dan lokasi kejadian
- How, yaitu berita harus memuat mengenai bagaimana kejadian tersebut bisa terjadi

### 2.3 Peringkasan Teks Otomatis

Peringkasan teks otomatis saat ini lazim digunakan pada program komputer yang memiliki tujuan untuk mengurangi dokumen teks dengan membuat ringkasan dari sebuah dokumen. Dengan memanfaatkan peringkasan teks otomatis maka dapat menghemat waktu untuk menemukan data relevan yang terdapat dalam sebuah teks (Chatterje et al, 2012). Terdapat dua cara dalam meringkas teks, yaitu :

#### 1. Peringkasan Abstraktif

Dengan menggunakan peringkasan secara abstraktif maka hasil ringkasan akan lebih kuat dikarenakan teknik ini mengandalkan pengetahuan yang mendalam dari suatu topik berita dan dapat mengutip dari berbagai sumber dengan topik bahasan yang sama, sehingga hasil ringkasan menjadi lebih menarik.

#### 2. Peringkasan Ekstraktif

Teknik ekstraktif hanya mengambil informasi paling penting oleh sistem untuk ringkasan (misalnya kata, kalimat, paragraf). Sehingga setiap ringkasan hanya terdiri dari potongan atau bagian dari teks tersebut.

### 2.4 Algoritma *TextRank*

*TextRank* merupakan algoritma berbasis graf bekerja dengan cara memberikan peringkat pada teks dengan cara merepresentasikan objek dalam teks (Mihalcea & Tarau, 2004). Dengan algoritma ini mengandalkan bentuk graf maka untuk dapat menentukan sebuah

kalimat penting atau tidak digunakan simpul (*node/verteks*). Dalam menerapkan algoritma ini, langkah pertama yang dilakukan adalah membangun sebuah graf yang terdiri dari titik (*verteks*) mewakili tiap kalimat. Setiap kalimat akan dihubungkan berdasarkan kesamaan (*similarity*) antar kalimat biasa disebut dengan *edge*. Sedangkan untuk menghitung bobot dari tiap kalimat Mihalcea & Tarau (2004) menggunakan rumus *weighted graph* ditentukan dengan Persamaan 2.1 berikut :

$$WS(V_i) = (1 - d) + d * \sum_{V_j \in In(V_i)} \frac{W_{ji}}{\sum_{V_k \in Out(V_j)} W_{jk}} WS(V_j) \quad (2.1)$$

Dimana  $d$  sebagai faktor yang bisa di beri nilai antara 0 dan 1 biasanya di set dengan 0.85. Untuk vertex  $V_i$ ,  $In(V_i)$  dijadikan simpul yang mengarah ke simpul lainnya (*predecessors*) dan  $Out(V_i)$  dijadikan simpul yang mengarah ke vertex  $V_i$  (*successors*) sedangkan  $WS(V_i)$  merupakan jumlah skor simpul  $V_i$ .

Metode *TextRank* dapat digunakan untuk menemukan *keyword* dan peringkasan teks secara otomatis dengan memilih kalimat penting (*sentence extraction*). Untuk dapat memilih kalimat penting pada teks maka digunakan rumus persamaan *similarity* yang dapat dilihat pada Persamaan 2.2 (Mihalcea, et al.2004) berikut :

$$Similarity(S_i, S_j) = \frac{|\{W_k | W_k \in S_i \& W_k \in S_j\}|}{\log(|S_i|) + \log(|S_j|)} \quad (2.2)$$

Secara keseluruhan tiap kalimat dalam teks akan dibagi ke dalam *token*. Dimana  $S_i$  dan  $S_j$  adalah simpul pada dua kalimat yang akan dihitung dan direpsentasikan dengan  $N_i$  yaitu kata yang muncul pada kedua kalimat misalnya :  $S_i = w_1^i, w_2^i, \dots w_{N_i}^i$  dan begitu juga kalimat untuk  $S_j$ . Selanjutnya  $\log(|S_i|)$  atau  $\log(|S_j|)$  akan dihitung dengan persamaan di atas.

Menurut Mihalcea & Tarau (2004), implementasi algoritma yang berbasis graf pada pemrosesan text bahasa alami memiliki 4 tahapan yaitu:

- Identifikasi unit teks yang paling sesuai untuk dijadikan simpul pada graf
- Pemberian sisi antar simpul unit teks baik dengan bobot atau tidak, berarah atau tidak berarah.

- Proses menggunakan algoritma hingga objek graf bertemu satu dengan lainnya (*convergence*).
- Urutkan simpul berdasarkan skornya. Nilai skor selanjutnya dimasukkan kedalam simpul pada saat proses algoritma.

Evaluasi dari algoritma *TextRank* dengan ekstraksi kalimat pada dokumen menggunakan 567 artikel berita dengan cara menguji dengan metode *recall-oriented understudy for gisting evaluation* (ROUGE). Metode ROUGE sendiri pada dasarnya merupakan metode pengembangan Ngram statistik yang dikembangkan oleh (Lin dan Hovy, 2003) yang bekerja dengan cara membandingkan kerja algoritma dengan manusia. Dari hasil evaluasi ternyata algoritma *TextRank* berada pada peringkat lima besar dari lima belas sistem yang dapat dilihat pada tabel 2.1.

**Tabel 2.1 Hasil Pengujian Algoritma *TextRank* dengan sistem lain**

Sistem	ROUGE score – Ngram (1,1)		
	<i>basic</i> (a)	<i>Stemmed</i> (b)	<i>stemmed no-stopwords</i> (c)
S27	0,4814	0,5011	0,4405
S31	0,4715	0,4914	0,4160
<b><i>TextRank</i></b>	<b>0,4708</b>	<b>0,4904</b>	<b>0,4229</b>
S28	0,4703	0,4890	0,4346
S21	0,4683	0,4869	0,4222
S29	0,4502	0,4681	0,4019

Hasil yang didapat dari pengujian algoritma *TextRank* pada data di atas cukup bagus karena berhasil mengenali kalimat penting dalam kalimat dengan hanya mengandalkan kesamaan antar kalimat. Menurut Anwar (2015) keunggulan dari algoritma *TextRank* antara lain :

- Unsupervised*, artinya algoritma ini bekerja tanpa perlu data *training* untuk memproses data.
- Language independent*, Dikarenakan *TextRank* hanya mengandalkan kata, kalimat ataupun paragraf yang terdapat dalam teks sehingga pengetahuan mendalam akan sebuah bahasa tidak begitu diandalkan.

## 2.5 Levenshtein Distance

*Levenshtein Distance* biasa digunakan untuk mengukur nilai kesamaan atau kemiripan antara dua buah kata (*string*). Jarak *Leveinshtein* diperoleh dengan mencari cara termudah untuk mengubah suatu string. Secara umum, operasi mengubah yang diperbolehkan untuk keperluan ini adalah :

- Memasukkan karakter ke dalam string
- Menghapus sebuah karakter dari suatu string
- Mengganti karakter string dengan karakter lain

Untuk menghitung jarak, digunakan matriks  $(n+1) \times (m+1)$  di mana  $n$  adalah panjang string  $s_1$  dan  $m$  adalah panjang string  $s_2$ . Dua buah string yang akan digunakan sebagai contoh adalah BARU dengan BATUK. Jika dilihat sekilas, kedua string tersebut memiliki jarak 2. Berarti untuk mengubah string BARU menjadi BATUK diperlukan 2 operasi yaitu:

1. Mensubtitusikan R dengan T (BARU  $\rightarrow$  BATU)
2. Menambahkan huruf K (BARU  $\rightarrow$  BATUK)

Proses perhitungan antara 2 string dapat dilakukan dengan cara menghitung matriks. Dengan menggunakan representasi matriks dapat dilihat pada tabel 2.2 di bawah.

**Tabel 2.2 Representasi Matriks untuk Menghitung Jarak Antar String**

		B	A	R	U
B	0	1	2	3	4
A	1	0	1	2	3
T	2	1	0	1	2
U	3	2	1	1	2
K	4	3	2	2	2
	5	4	3	3	2

Pada Tabel 2.2 elemen baris 1 kolom 1 ( $M[1,1]$ ) adalah jumlah operasi yang diperlukan untuk mengubah *substring* dari kata BARU yang diambil mulai dari karakter awal sebanyak 1(B) ke *substring* dari kata BATUK yang diambil juga dari karakter awal sebanyak 1 yaitu B. Dengan demikian maka pada  $M[2,4]$  adalah jumlah operasi yang dilakukan dari kata BA (*substring* yang diambil dari awal karakter sebanyak 2) dengan BATU (*substring* yang diambil dari awal karakter sebanyak 4). Hal ini disimpulkan bahwa elemen  $M[p,q]$  adalah jumlah operasi antar substring kata pertama yang dimulai dari awal sebanyak p dengan substring kata kedua yang dimulai dari awal sebanyak q. Elemen terakhir(kanan bawah) adalah elemen yang nilainya menyatakan jarak antara kedua *string* yang dibandingkan. Untuk menghitung nilai kemiripan dapat menggunakan Persamaan 2.3 berikut.

$$Sim = 1 - \left( \frac{Dis}{MaxLength} \right) \quad (2.3)$$

Dimana *Sim* adalah nilai kemiripan antara kedua kalimat. Sedangkan *Dis* adalah jarak *Levenshtein* dan *MaxLength* adalah string terpanjang. Jika nilai *similarity* adalah 1, maka kedua string yang dibandingkan sama sebaliknya jika nilainya 0, maka kedua string yang dibandingkan tidak sama dan perlu dilakukan operasi perubahan, penambahan atau pengurangan.

## 2.6 ROUGE

*ROUGE* atau biasa dikenal dengan *Recall-Oriented Understudy for Gisting Evaluation* adalah serangkaian metrik dan perhitungan komputer yang biasanya digunakan untuk mengevaluasi peringkasan teks otomatis ataupun terjemahan otomatis yang dilakukan oleh komputer dan sangat berguna bagi fokus ilmu *natural language processing*. Algoritma ini digunakan untuk menentukan kualitas dari peringkasan teks otomatis yang dilakukan oleh sistem dengan membandingkannya dengan ringkasan ideal yang dibuat oleh manusia. Langkah yang dilakukan dalam pengerjaan algoritma ini adalah dengan menghitung tiap unit yang berulang dan membandingkannya ke dalam beberapa aspek penilaian antara lain *n*-gram, urutan kata, dan kesamaan antar unit yang hasilnya akan dievaluasi langsung oleh peringkasan yang dilakukan oleh manusia Chin-Yew Lin (2004).



### 2.6.1 ROUGE – N

*ROUGE – N* adalah perhitungan yang dilakukan dengan merepresentasikan ringkasan kedalam  $n$  – gram dan dihitung berdasarkan *recall* yang selanjutnya akan dibandingkan dengan referensi peringkasan yang telah dibuat oleh ahli. Rumus perhitungannya dapat dilihat pada Persamaan 2.4 berikut :

$$ROUGE - N = \frac{\sum_{S \in \{ReferenceSummaries\}} \sum_{gram_n \in S} Count_{match}(gram_n)}{\sum_{S \in \{ReferenceSummaries\}} \sum_{gram_n \in S}} \quad (2.4)$$

Dimana  $n$  adalah panjang  $n$  – gram dan  $Count_{match}(gram_n)$  adalah nilai maksimum dari  $n$ -gram yang diberikan untuk dihitung berdasarkan referensi yang telah disediakan. Nilai dari  $n$  ditetapkan berdasarkan 2 hal yaitu:

1. *ROUGE – 1* : Mengacu pada unit yang dibandingkan berdasarkan *1-gram* (tiap kata) antara sistem dan referensi.
2. *ROUGE – 2* : Mengacu pada *bigrams* keseluruhan isi dari sistem dengan referensi.

## 2.7 Penelitian Terdahulu

Peringkasan teks otomatis menjadi salah satu alternatif untuk mendapatkan hasil yang terbaik dalam mengekstraksi kalimat penting atau *keyword* dari teks yang panjang. Manfaat lain dari peringkasan teks otomatis adalah untuk menemukan fakta dan kalimat utama dari sebuah teks dengan cepat.

Yazan A.Jaradat *et al.*(2016) meneliti tentang peringkasan teks otomatis menggunakan genetik algoritma pada teks berbahasa Arab. Dimana keunggulan dari sistem yang dibuat adalah dengan menggunakan algoritma genetika dapat mengekstrak kalimat dengan lebih baik karena dibantu dalam hal *scoring*. Selain itu dengan adanya algoritma genetika memecahkan masalah peringkasan berbasis graf yang biasa dilakukan dengan normal graf. Tetapi kekurangan dari algoritma genetika adalah sistem memerlukan pemahaman tentang bahasa yang akan diringkas agar hasilnya lebih optimal.

Nguyen Quang, Pham Tuan Anh, Truong Cong, Nguyen Xuan dari Vietnam pada tahun 2012 juga melakukan peringkasan teks otomatis. Dimana mereka menggunakan genetic programming untuk mengekstraksi kalimat penting pada teks berbahasa Vietnam. Genetic programming memiliki 2 tahapan yaitu training dan testing. Tujuan dari training

adalah melatih sistem untuk meningkatkan kemampuan menemukan kalimat penting dalam teks. Sedangkan pada tahapan testing data yang paling bagus dari training akan digunakan untuk meningkatkan hasil peringkasan. Penelitian ini bisa dibilang cukup sukses dikarenakan dari proses evaluasi yang dilakukan dengan membandingkan antara kinerja peringkasan teks oleh ahli bidang linguistik, siswa, dan genetic programming hasil dari sistem 20% lebih baik dari yang dilakukan oleh siswa, walaupun sistem memiliki hasil lebih rendah dari ahli bidang linguistik.

Pada tahun 2012 Jayashree R, Basavaraj Anami, dan Srikanta Murthy K pernah membuat sistem peringkasan teks dimana sistem tersebut dapat mengategorikan teks dalam bahasa Kannada. Penelitian tersebut dilakukan dengan menggunakan algoritma GSS (Galavotti, Sebastiani, Simi) digabungkan dengan tf-idf. Sistem tersebut bekerja dengan cara mengekstraksi kata kunci dan tulisan untuk digunakan pada peringkasan teks. Dimana kata kunci yang diambil haruslah memiliki bobot atau hubungan yang tinggi antara satu dengan yang lain. Hasil yang didapat dari penelitian tersebut menunjukkan peringkasan teks yang dilakukan oleh manusia lebih efektif dibandingkan dengan komputer hal tersebut terjadi karena peringkasan teks yang dilakukan manual oleh manusia memiliki koheren antar kalimat lebih tinggi antar kalimat dibandingkan dengan komputer. Sudah banyak penelitian yang dilakukan untuk mencari metode terbaik dalam meringkas teks yang dilakukan oleh peneliti sebelumnya yang dapat dilihat pada tabel 2.3 berikut :

**Tabel 2.3 Penelitian Terdahulu**

No	Judul	Peneliti	Tahun	Keterangan
1	Sistem Peringkasan Berita Online Otomatis Menggunakan Algoritma TextTeaser	Anwar Pasaribu	2016	Penelitian ini menggunakan Algoritma <i>TextTeaser</i> berbasis mobile dievaluasi dengan dibandingkan dengan algoritma <i>TextRank</i>
2.	TextRank : Bringing Order into Texts	Rada Mihalcea dan Paul Tarau	2004	Penelitian ini menggunakan algoritma <i>TextRank</i> dalam meringkas teks abstrak dengan cara <i>word extraction</i> dan <i>sentence extraction</i> pada teks berbahasa Inggris.

Tabel 2.3 Penelitian Terdahulu (Lanjutan)

No.	Judul	Peneliti	Tahun	Keterangan
3.	A Language Independent Algorithm for Single and Multiple Document Summarization	Rada Mihalcea dan Paul Tarau	2005	Penelitian ini menggunakan algoritma HITS, <i>PageRank</i> , dan <i>TextRank</i> untuk mengevaluasi peringkasan teks otomatis untuk banyak atau satu dokumen.
4.	Hybrid-based Arabic Single – Document Text Summarization Approach Using Genetic Algorithm	Yazan A.Jaradat dan Ahmad T. Al - Taani	2016	Penelitian ini menggunakan gabungan antara Bayes dan <i>Genetic Algorithm</i> untuk meringkas dokumen tunggal berbahasa Arab.
5.	Categorized Text Document Summarization in the Kannada Language by Sentence Ranking	Jayashree R dan Srikanta Murthy K	2016	Penelitian ini meneliti bahasa Kannada yang biasa digunakan oleh masyarakat India Selatan. Teks yang diteliti menggunakan algoritma Tf - idf.
6.	A Study on the Use of Genetic Programming for Automatic Text Summarization	Nguyen Quang Uy, Pham Tuan Anh, Truong Cong Doan dan Nguyen Xuan Hoai	2012	Penelitian ini meneliti <i>Genetic Programming</i> untuk meringkas teks berbahasa vietnam dimana hasil uji dengan membandingkan yang dilakukan oleh sistem, siswa dan ahli bidang linguistik

## BAB 3

### ANALISIS DAN PERANCANGAN SISTEM

Bab ini akan menjelaskan tentang perancangan sistem mulai dari arsitektur umum yang akan dibuat, menyiapkan data yang dibutuhkan, sampai tampilan antar muka yang dibentuk. Analisis terhadap algoritma *TextRank* sangat dibutuhkan agar dapat diimplementasi ke dalam sistem dalam meringkas teks berita.

#### 3.1 Data

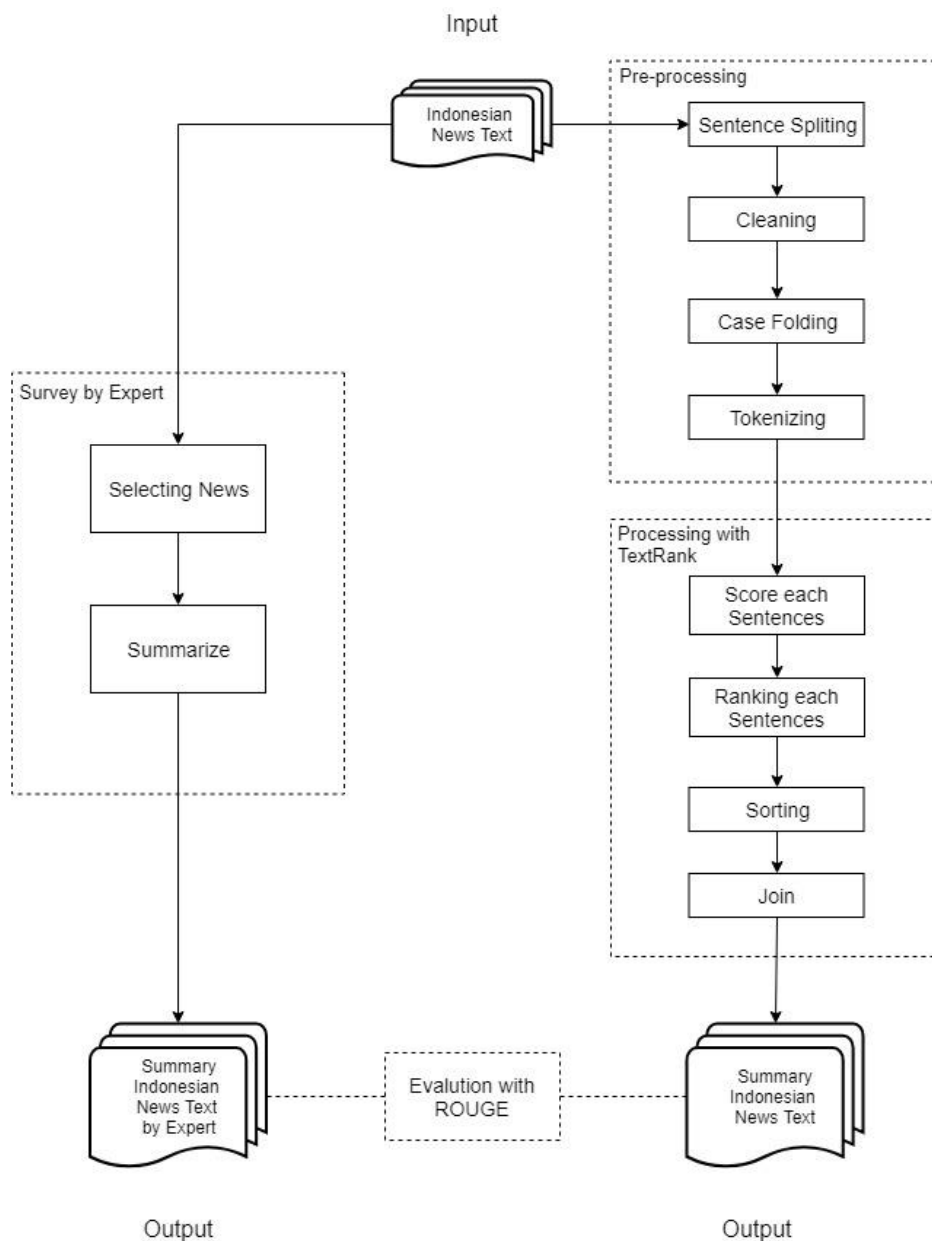
Data yang digunakan pada penelitian ini adalah teks berita *online* yang sudah diambil dari halaman berita *online* antara lain liputan6.com, kompas.com dan detik.com. Terdapat sepuluh kategori teks berita yaitu teknologi, kesehatan, umum, olahraga, ekonomi, gaya hidup, nasional, sains, *travel* dan otomotif. Secara keseluruhan terdapat 100 berita yang akan digunakan pada penelitian ini dan setiap teks berita memiliki panjang rata-rata 400 sampai 900 kata per berita.

#### 3.2 Arsitektur Umum

Sistem peringkasan otomatis biasa digunakan untuk dapat mengekstraksi kalimat penting pada teks berita sehingga mendapatkan kalimat inti atau intisari dari keseluruhan isi teks berita. Sistem ini berjalan dengan cara mengolah seluruh teks berita dan menghasilkan beberapa kalimat penting. Peringkasan teks otomatis bermanfaat untuk dapat menemukan fakta atau data-data yang terdapat pada sebuah teks berita.

Terdapat 3 tahapan utama yang terjadi pada sistem peringkasan teks otomatis menggunakan algoritma *TextRank* yaitu *preprocessing*, *processing*, dan *postprocessing*. Tahapan *preprocessing* dimulai dengan pembersihan kalimat dengan cara mengubah seluruh kalimat menjadi huruf kecil, menghilangkan tanda baca, dan memberikan nilai awal pada setiap kalimat. Tahapan ini dilakukan agar kalimat yang akan diringkas memiliki bobot nilai awal yang sama sebelum nanti diproses oleh sistem. Dilanjutkan dengan tahapan *processing*, pada tahapan ini setiap kalimat yang sudah dibersihkan dan diberikan nilai awal akan masuk

diproses oleh sistem. Proses yang dilakukan antara lain yaitu pembobotan antar kalimat untuk menemukan kesamaan (*similarity*) untuk dapat diketahui apakah kalimat tersebut dapat dikategorikan sebagai kalimat penting atau tidak. Setelah melewati beberapa perhitungan antar kalimat maka tahapan yang terakhir yaitu *postprocessing* dimana beberapa kalimat yang sudah dihitung bobot dan nilai *similarity* akan digabungkan menjadi kalimat yang utuh yang nantinya menjadi akhir dari ringkasan kalimat menggunakan algoritma *TextRank*. Secara lebih detail arsitektur umum dari sistem yang dibuat dapat dilihat pada gambar 3.1 berikut :



**Gambar 3.1. Arsitektur Umum**

Metode yang dilakukan pada penelitian ini adalah dengan membandingkan hasil peringkasan teks yang dilakukan oleh sistem menggunakan algoritma *Textrank* dengan hasil yang dilakukan oleh ahli dalam bidang linguistik. Metode ini dilakukan untuk dapat menemukan tingkat presisi algoritma tersebut terhadap peringkasan teks berbahasa Indonesia. Tahapan secara keseluruhan proses peringkasan teks otomatis berbahasa Indonesia akan dijelaskan sebagai berikut :

### 3.2.1 Input

Tahapan awal yang dilakukan pada penelitian ini adalah dengan memilih teks berita bahasa Indonesia berdasarkan judul berita. Terdapat 300 judul berita yang dapat dipilih, dimana masing-masing judul telah diambil dari 10 kategori yaitu teknologi, kesehatan, umum, olahraga, ekonomi, gaya hidup, nasional, sains, travel dan otomotif. Teks berita yang dipilih nantinya akan ditampilkan berdasarkan isi dari berita seperti pada gambar 3.2 berikut :

#### **Input Teks :**

JAKARTA, KOMPAS. com - Di sela-sela kunjungannya ke kantor Menteri Komunikasi dan Informatika ( Menkominfo) Rudiantara, founder Google Sergey Brin mengaku kagum dengan aplikasi Go-Jek dan tertarik memasang aplikasinya. Brin pun langsung ingin menginstal-nya saat menteri memberi tahu keberadaan fitur GoMassage. GoMassage merupakan salah satu fitur di aplikasi Gojek yang memungkinkan pengguna memesan jasa layanan pijat. Kunjungan Brin pada Senin (28/12/2015) sore itu singkat saja. Satu jam lebih dia berada di Kemenkominfo, dengan lebih dari 30 menit dihabiskan berbincang-bincang dengan Rudiantara di ruang tertutup. Setelah keluar dari ruangan, keduanya segera menggelar jumpa pers. Rudiantara mengatakan bahwa selama di dalam ruangan tadi, dirinya mengungkapkan ambisi untuk menciptakan 1000 startup lokal berkualitas dalam lima tahun ke depan, mengembangkan ekonomi digital dan roadmap e-commerce. Menteri yang akrab disapa Chief RA itu juga meminta keterlibatan Google di Indonesia untuk mewujudkan ambisi tersebut. Brin menanggapi permintaan dengan hangat. Bos laboratorium Google X itu mengungkapkan bahwa mereka memiliki program untuk mengembangkan startup di Indonesia, yaitu Launchpad Accelerator. "Saya senang

**Gambar 3.2. Contoh Teks Berbahasa Indonesia**

berada di Indonesia, negara digital ekonomi terbesar di regional Asia. Punya 300 juta pelanggan seluler lebih. Indonesia sudah lompat ke next generation technology. Tentu saja Google akan senang hati mau membantu program digital ekonomi Indonesia," ujarnya. Di sela-sela pembicaraan tersebut, Brin sempat melontarkan pujian untuk Go-Jek. Menurutnya, usaha rintisan Nadiem Makarim menjadi fenomena yang terdengar hingga mancanegara. "Melihat entrepreneurship seperti Go-Jek membuat kami ingin berada di sini (pengembangan startup)," pungkasnya. Rudiantara menanggapi pernyataan tersebut. Dia memandang ojek berbasis aplikasi itu merupakan model bisnis baru yang sanggup mengguncang dunia transportasi. "Go-Jek itu bukan cuma transport orang, tapi bikin malas. Apapun diantarkan, misalnya kalau ingin makanan atau massage. Membuat malas tapi juga memberi efisiensi. Ini masa," ujarnya. Namun sebelum menteri melanjutkan kata-katanya, Brin keburu memotong, "Wait, what?! Bisa pesan massage juga? Kalau begitu saya install Go-Jek sekarang." Menteri dan para hadirin pun sontak tertawa. Dia kemudian menambahkan bahwa aplikasi seperti Go-Jek merupakan masa depan ekonomi digital Indonesia dan berharap akan lebih banyak lagi jenisnya

### **Gambar 3.2. Contoh Teks Berbahasa Indonesia (Lanjutan)**

#### **3.2.2 Proses**

Dari tahapan input maka selanjutnya teks yang telah dipilih akan diproses oleh sistem. Proses dilakukan oleh sistem dengan cara memberikan nilai pada tiap kalimat yang nantinya akan dihitung berdasarkan kesamaan antar kalimat. Kalimat dengan nilai tertinggi akan dijadikan sebagai hasil ringkasan. Untuk lebih jelasnya maka tahapan proses akan dijelaskan sebagai berikut, dimana pada tahap proses terdapat 2 bagian, yaitu :

##### **A. *Pre-processing***

##### **i. *Sentence Splitting***

*Sentence splitting* adalah tahapan dimana teks berita yang akan diinputkan akan dipecah menjadi kalimat per kalimat. Setiap kalimat yang dipecah akan mempunyai nilai tersendiri yang disebut token. Berdasarkan gambar 3.2 di atas maka didapatkan hasil *sentence splitting* yang dapat dilihat pada tabel 3.1 sebagai berikut :

**Tabel 3.1. Tahapan *Sentence Splitting***

No.	Kalimat
1.	JAKARTA, KOMPAS. com - Di sela-sela kunjungannya ke kantor Menteri Komunikasi dan Informatika ( Menkominfo) Rudiantara, founder Google Sergey Brin mengaku kagum dengan aplikasi Go-Jek dan tertarik memasang aplikasinya.
2.	Brin pun langsung ingin menginstal-nya saat menteri memberi tahu keberadaan fitur GoMassage.
3.	GoMassage merupakan salah satu fitur di aplikasi Gojek yang memungkinkan pengguna memesan jasa layanan pijat.
4.	Kunjungan Brin pada Senin (28/12/2015) sore itu singkat saja.
5.	Satu jam lebih dia berada di Kemenkominfo, dengan lebih dari 30 menit dihabiskan berbincang-bincang dengan Rudiantara di ruang tertutup.
6.	Setelah keluar dari ruangan, keduanya segera menggelar jumpa pers.
7.	Rudiantara mengatakan bahwa selama di dalam ruangan tadi, dirinya mengungkapkan ambisi untuk menciptakan 1000 startup lokal berkualitas dalam lima tahun ke depan, mengembangkan ekonomi digital dan roadmap e-commerce.
8.	Menteri yang akrab disapa Chief RA itu juga meminta keterlibatan Google di Indonesia untuk mewujudkan ambisi tersebut.Brin menanggapi permintaan dengan hangat.
9.	Bos laboratorium Google X itu mengungkapkan bahwa mereka memiliki program untuk mengembangkan startup di Indonesia, yaitu Launchpad Accelerator.
10.	"Saya senang berada di Indonesia, negara digital ekonomi terbesar di regional Asia. Punya 300 juta pelanggan seluler lebih. Indonesia sudah lompat ke next generation technology. Tentu saja Google akan senang hati mau membantu program digital ekonomi Indonesia," ujarnya.Di sela-sela pembicaraan tersebut, Brin sempat melontarkan pujian untuk Go-Jek.Menurutnya, usaha rintisan Nadiem Makarim menjadi fenomena yang terdengar hingga mancanegara.
11.	"Melihat entrepreneurship seperti Go-Jek membuat kami ingin berada di sini (pengembangan startup)," pungkasnya.Rudiantara menanggapi pernyataan tersebut.



**Tabel 3.1. Tahapan *Sentence Splitting* (Lanjutan)**

No.	Kalimat
12.	Dia memandang ojek berbasis aplikasi itu merupakan model bisnis baru yang sanggup mengguncang dunia transportasi.
13.	"Go-Jek itu bukan cuma transport orang, tapi bikin malas. Apapun diantarkan, misalnya kalau ingin makanan atau massage. Membuat malas tapi juga memberi efisiensi. Ini masa," ujarnya. Namun sebelum menteri melanjutkan kata-katanya, Brin keburu memotong, "Wait, what?! Bisa pesan massage juga? Kalau begitu saya install Go-Jek sekarang."
14.	Menteri dan para hadirin pun sontak tertawa.
15.	Dia kemudian menambahkan bahwa aplikasi seperti Go-Jek merupakan masa depan ekonomi digital Indonesia dan berharap akan lebih banyak lagi jenisnya.

ii. *Cleaning*

*Cleaning* adalah tahapan dimana teks berita akan dibersihkan terlebih dahulu dengan cara menghilangkan tanda baca, dan angka yang terdapat pada kalimat. Berdasarkan tabel 3.1 di atas berikut adalah *Cleaning* yang dilakukan seperti terlihat pada gambar 3.3.

**Input Teks :**

JAKARTA KOMPAS com Di selasela kunjungannya ke kantor Menteri Komunikasi dan Informatika Menkominfo Rudiantara founder Google Sergey Brin mengaku kagum dengan aplikasi GoJek dan tertarik memasang aplikasinya Brin pun langsung ingin menginstalnya saat menteri memberi tahu keberadaan fitur GoMassage GoMassage merupakan salah satu fitur di aplikasi Gojek yang memungkinkan pengguna memesan jasa layanan pijat Kunjungan Brin pada Senin sore itu singkat saja Satu jam lebih dia berada di Kemenkominfo dengan lebih dari menit dihabiskan berbincang bincang dengan Rudiantara di ruang tertutup Setelah keluar dari ruangan keduanya segera menggelar jumpa pers Rudiantara mengatakan bahwa selama di dalam ruangan tadi dirinya

**Gambar 3.3. *Text Cleaning***

mengungkap ambisi untuk menciptakan startup lokal berkualitas dalam lima tahun ke depan mengembangkan ekonomi digital dan roadmap e commerce Menteri yang akrab disapa Chief RA itu juga meminta keterlibatan Google di Indonesia untuk mewujudkan ambisi tersebut Brin menanggapi permintaan dengan hangat Bos laboratorium Google X itu mengungkap bahwa mereka memiliki program untuk mengembangkan startup di Indonesia yaitu Launchpad Accelerator Saya senang berada di Indonesia negara digital ekonomi terbesar di regional Asia Punya juta pelanggan seluler lebih Indonesia sudah lompat ke next generation technology Tentu saja Google akan senang hati mau membantu program digital ekonomi Indonesia ujarnya Di sela-sela pembicaraan tersebut Brin sempat melontarkan pujian untuk Go Jek. Menurutnya usaha rintisan Nadiem Makarim menjadi fenomena yang terdengar hingga mancanegara Melihat entrepreneurship seperti Go Jek membuat kami ingin berada di sini pengembangan startup pungkasnya. Rudiantara menanggapi pernyataan tersebut Dia memandang ojek berbasis aplikasi itu merupakan model bisnis baru yang sanggup mengguncang dunia transportasi Go Jek itu bukan cuma transport orang tapi bikin malas Apapun diantarkan, misalnya kalau ingin makanan atau massage Membuat malas tapi juga memberi efisiensi Ini masa ujarnya Namun sebelum menteri melanjutkan kata katanya Brin keburu memotong Wait what Bisa pesan massage juga Kalau begitu saya install Go Jek sekarang Menteri dan para hadirin pun sontak tertawa Dia kemudian menambahkan bahwa aplikasi seperti Go Jek merupakan masa depan ekonomi digital Indonesia dan berharap akan lebih banyak lagi jenisnya

**Gambar 3.3. Text Cleaning (Lanjutan)**

iii. *Case Folding*

*Case Folding* adalah tahapan selanjutnya setelah tahapan *cleaning* dimana sistem mengubah semua huruf dalam teks menjadi huruf kecil. Karena tidak semua kata dalam teks menggunakan huruf kapital oleh karena itu peran *case folding* dibutuhkan untuk mengkonversi keseluruhan teks menjadi sebuah bentuk standar sehingga memudahkan untuk melanjutkan ke tahapan *tokenizing* selanjutnya. Hasil dari tahapan *case folding* dapat dilihat pada tabel 3.2 berikut :

**Tabel 3.2. Tahapan Case Folding**

No.	Kalimat
1.	jakarta, kompas com di selasela kunjungannya ke kantor menteri komunikasi dan informatika menkominfo rudiantara, founder google sergey brin mengaku kagum dengan aplikasi go jek dan tertarik memasang aplikasinya
2.	brin pun langsung ingin menginstal-nya saat menteri memberi tahu keberadaan fitur gomassage gomassage merupakan salah satu fitur di aplikasi gojek yang memungkinkan pengguna memesan jasa layanan pijat
3.	gomassage merupakan salah satu fitur di aplikasi gojek yang memungkinkan pengguna memesan jasa layanan pijat
4.	kunjungan brin pada senin sore itu singkat saja
5.	satu jam lebih dia berada di kemenkominfo, dengan lebih dari menit dihabiskan berbincang-bincang dengan rudiantara di ruang tertutup.
6.	setelah keluar dari ruangan, keduanya segera menggelar jumpa pers.
7.	rudiantara mengatakan bahwa selama di dalam ruangan tadi, dirinya mengungkap ambisi untuk menciptakan startup lokal berkualitas dalam lima tahun ke depan, mengembangkan ekonomi digital dan roadmap ecommerce
8.	menteri yang akrab disapa chief ra itu juga meminta keterlibatan google di indonesia untuk mewujudkan ambisi tersebut.brin menanggapi permintaan dengan hangat
9.	bos laboratorium google x itu mengungkap bahwa mereka memiliki program untuk mengembangkan startup di indonesia yaitu launchpad accelerator
10.	saya senang berada di indonesia, negara digital ekonomi terbesar di regional asia. punya juta pelanggan seluler lebih indonesia sudah lompat ke next generation technology tentu saja google akan senang hati mau membantu program digital ekonomi Indonesia ujarnya.di sela-sela pembicaraan tersebut brin sempat melontarkan pujian untuk gojek.menurutnya, usaha rintisan nadiem makarim menjadi fenomena yang terdengar hingga mancanegara

**Tabel 3.2. Tahapan Case Folding (Lanjutan)**

No.	Kalimat
11.	melihat entrepreneurship seperti go jek membuat kami ingin berada di sini pengembangan startup pungkasnya.rudiantara menanggapi pernyataan tersebut
12.	dia memandang ojek berbasis aplikasi itu merupakan model bisnis baru yang sanggup mengguncang dunia transportasi
13.	go jek itu bukan cuma transport orang tapi bikin malas apapun diantarkan, misalnya kalau ingin makanan atau massage membuat malas tapi juga memberi efisiensi ini masa ujarnya.namun sebelum menteri melanjutkan kata-katanya brin keburu memotong wait what bisa pesan massage juga kalau begitu saya install go jek sekarang
14.	menteri dan para hadirin pun sontak tertawa
15.	dia kemudian menambahkan bahwa aplikasi seperti go jek merupakan masa depan ekonomi digital indonesia dan berharap akan lebih banyak lagi jenisnya

iv. *Tokenizing*

*Tokenizing* adalah tahap pemotongan string input berdasarkan tiap kalimat yang menyusunnya. Tokenisasi secara garis besar memecah kalimat dalam suatu teks ke dalam bentuk sekumpulan satuan kalimat untuk dihitung hasilnya nanti. Proses tokenisasi akan semakin sulit dikarenakan harus memperhatikan struktur bahasa. Tahapan *Tokenizing* dapat dilihat seperti pada tabel 3.3 berikut :

**Tabel 3.3 Tahapan Tokenizing**

No.	Text	Token	Score
1.	JAKARTA, KOMPAS.	jakarta kompas	-1
2.	com - Di sela-sela kunjungannya ke kantor Menteri Komunikasi dan Informatika ( Menkominfo) Rudiantara, founder Google Sergey Brin mengaku kagum	com di sela sela kunjungannya ke kantor menteri komunikasi dan informatika menkominfo rudiantara founder google sergey brin mengaku kagum dengan aplikasi go jek dan tertarik	-1

	<p>dengan aplikasi Go-Jek dan tertarik memasang aplikasinya. Brin pun langsung ingin menginstal-nya saat menteri memberi tahu keberadaan fitur GoMassage. GoMassage merupakan salah satu fitur di aplikasi Gojek yang memungkinkan pengguna memesan jasa layanan pijat. Kunjungan Brin pada Senin (28/12/2015) sore itu singkat saja. Satu jam lebih dia berada di Kemenkominfo, dengan lebih dari 30 menit dihabiskan berbincang-bincang dengan Rudiantara di ruang tertutup.</p>	<p>memasang aplikasinya brin pun langsung ingin menginst nya saat menteri memberi tahu keberadaan fitur gomassage gomassage merupakan salah satu fitur di aplikasi gojek yang memungkinkan pengguna memesan jasa layanan pijat kunjungan brin pada senin sore itu singkat saja satu jam lebih dia berada di kemenkominfo dengan lebih dari menit dihabiskan berbincang bincang dengan rudiantara di ruang tertutup</p>	
3.	<p>Setelah keluar dari ruangan, keduanya segera menggelar jumpa pers. Rudiantara mengatakan bahwa selama di dalam ruangan tadi, dirinya mengungkap ambisi untuk menciptakan 1000 startup lokal berkualitas dalam lima tahun ke depan, mengembangkan ekonomi digital dan roadmap e-commerce. Menteri yang akrab disapa Chief RA itu juga meminta keterlibatan Google di</p>	<p>setelah keluar dari ruangan keduanya segera menggelar jumpa pers rudiantara mengatakan bahwa selama di dalam ruangan tadi dirinya mengungkap ambisi untuk menciptakan startup lokal berkualitas dalam lima tahun ke depan mengembangkan ekonomi digit dan roadmap e commerce menteri yang akrab disapa chief ra itu juga meminta keterlibatan google di indonesia untuk</p>	-1

	Indonesia untuk mewujudkan ambisi tersebut. Brin menanggapi permintaan dengan hangat. Bos laboratorium Google X itu mengungkapkan bahwa mereka memiliki program untuk mengembangkan startup di Indonesia, yaitu Launchpad Accelerator.	mewujudkan ambisi tersebut brin menanggapi permintaan dengan hangat bos laboratorium googl x itu mengungkap bahwa mereka memiliki program untuk mengembangkan startup di indonesia yaitu launchpad acceler	
4.	Saya senang berada di Indonesia, negara digital ekonomi terbesar di regional Asia.	saya senang berada di indonesia negara digit ekonomi terbesar di region asia	-1
5.	Punya 300 juta pelanggan seluler lebih.	punya juta pelanggan selul lebih	-1
6.	Indonesia sudah lompat ke next generation technology.	indonesia sudah lompat ke generation technology	-1
7.	Tentu saja Google akan senang hati mau membantu program digital ekonomi Indonesia, ujarnya. Di sela-sela pembicaraan tersebut, Brin sempat melontarkan pujian untuk Go-Jek. Menurutnya, usaha rintisan Nadiem Makarim menjadi fenomena yang terdengar hingga mancanegara. "Melihat entrepreneurship seperti Go-Jek membuat kami ingin berada di sini (pengembangan	tentu saja googl akan senang hati mau membantu program digit ekonomi indonesia ujarnya di sela sela pembicaraan tersebut brin sempat melontarkan pujian untuk jek menurutnya usaha rintisan nadiem makarim menjadi fenomena yang terdengar hingga mancanegara melihat entrepreneurship seperti jek membuat kami ingin berada di sini pengembangan startup pungkasnya rudiantara menanggapi pernyataan tersebut	-1

	startup),pungkasnya.Rudiantara menanggapi pernyataan tersebut.Dia memandang ojek berbasis aplikasi itu merupakan model bisnis baru yang sanggup mengguncang dunia transportasi."Go-Jek itu bukan cuma transport orang, tapi bikin malas.	dia memandang ojek berbasis aplikasi itu merupakan model bisni baru yang sanggup mengguncang dunia transportasi jek itu bukan cuma transport orang tapi bikin malas	
8.	Apapun diantarkan, misalnya kalau ingin makanan atau massage.	apapun diantarkan misalnya kalau ingin makanan atau massag	-1
9.	Membuat malas tapi juga memberi efisiensi.	membuat malas tapi juga memberi efisiensi	-1
10.	Ini masa," ujarnya.Namun sebelum menteri melanjutkan kata-katanya, Brin keburu memotong, "Wait, what?!	ini masa ujarnya namun sebelum menteri melanjutkan kata katanya brin keburu memotong wait	-1
11.	Bisa pesan massage juga?	bisa pesan massage juga	-1
12.	Kalau begitu saya install Go-Jek sekarang."Menteri dan para hadirin pun sontak tertawa.Dia kemudian menambahkan bahwa aplikasi seperti Go-Jek merupakan masa depan ekonomi digital Indonesia dan berharap akan lebih banyak lagi jenisnya	kalau begitu saya instal jek sekarang menteri dan para hadirin pun sontak tertawa dia kemudian menambahkan bahwa aplikasi seperti jek merupakan masa depan ekonomi digit indonesia dan berharap akan lebih banyak lagi jenisnya	-1

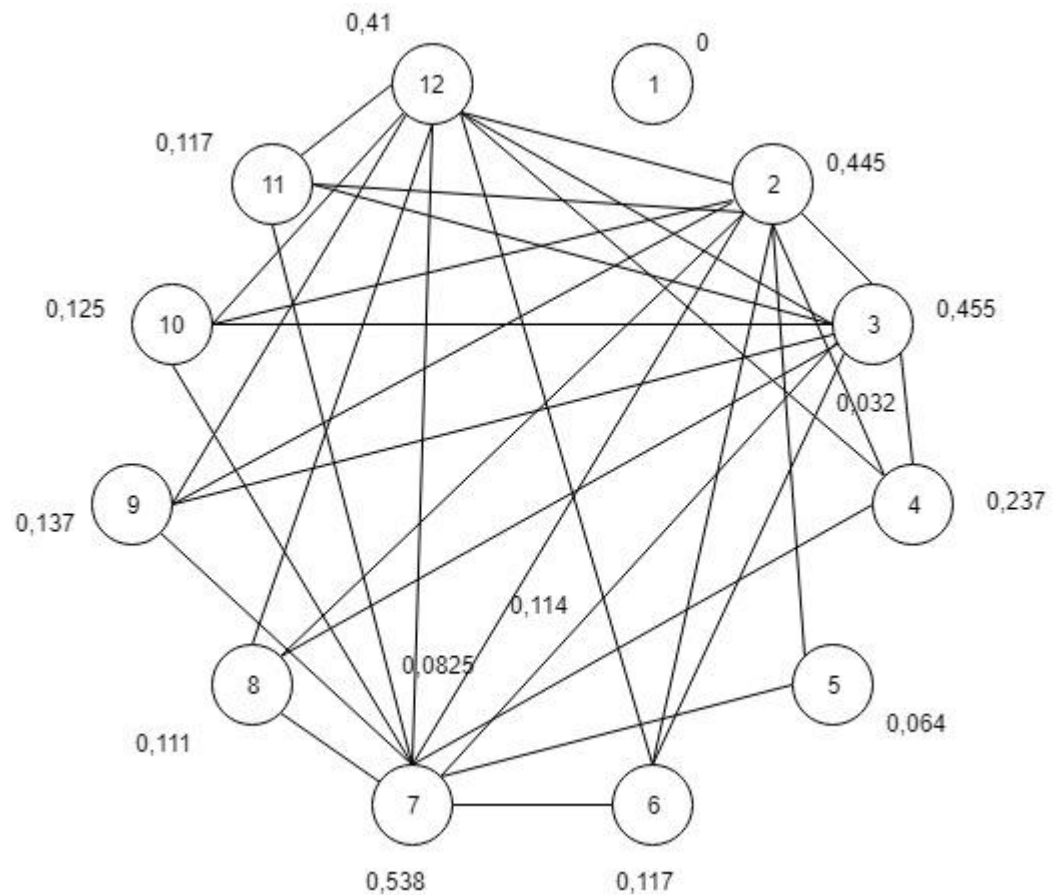
Pada proses awal *tokenizing* setiap token akan diberikan nilai awal sebesar -1 yang nantinya akan digunakan untuk proses perhitungan selanjutnya.

## B. Processing

Pada tahapan ini setiap token yang sudah diberikan nilai awal akan dihitung nilai kesamaan untuk dapat menentukan kalimat yang penting berdasarkan *score* yang diperoleh. Sebelum menghitung nilai antar kalimat maka beberapa hal yang dilakukan antara lain :

### i. Merepresentasikan graf

Setiap token yang sudah terpilih akan mewakili 1 bagian(*vertex*) pada graf yang nantinya akan dihubungkan dengan setiap token yang ada untuk mendapatkan nilai *similarity* (*edge*). Representasi graf dapat dilihat berdasarkan gambar 3.4 berikut :



**Gambar 3.4 Representasi Graf**

Pada Gambar 3.4 di atas dapat dilihat bahwa setiap token pada kalimat direpresentasi sebagai *vertex* pada graf dan *similarity* (kesamaan) antar kalimat direpresentasi sebagai *edge* atau garis. Dan proses perhitungan *edge* akan dijelaskan pada bagian perhitungan di bawah.



ii. Menghitung *similarity* antar kalimat

Menghitung *similarity* dapat dilakukan dengan beberapa syarat antara lain tiap *vertex* memiliki kata yang sama (*word overlap*) dan dinormalisasi dengan membagi jumlah *word overlap* dengan panjang tiap kalimat, sedangkan untuk *vertex* yang tidak memiliki sama sekali *word overlap* maka nilai *similarity*-nya adalah 0 sehingga tidak ada *edge* antar kedua *vertex*. Berikut adalah contoh perhitungan *similarity* yang hasil *edge* dapat dilihat pada gambar 3.4 di atas.

- Menghitung *similarity* antara *vertex* ke-7 dengan *vertex* ke-12

Perlu diperhatikan bahwa antara *vertex* ke-7 dan *vertex* ke-12 harus memiliki *word overlap* sehingga proses perhitungan dapat dilanjutkan. Pemeriksaan antar kalimat dapat dilihat pada tabel 3.4 berikut :

**Tabel 3.4 Word Overlap pada *vertex* ke-7 dan *vertex* ke-12**

<i>Vertex</i> ke-7	<i>Vertex</i> ke-12
tentu saja google akan senang hati mau membantu program <b>digital</b> <b>ekonomi indonesia</b> ujanya di sela sela pembicaraan tersebut brin sempat melontarkan pujian untuk <b>gojek</b> menurutnya usaha rintisan nadiem makarim menjadi fenomena yang terdengar hingga mancanegara melihat entrepreneurship <b>seperti gojek</b> membuat kami ingin berada di sini pengembangan startup pungkasnya rudiantara menanggapi pernyataan tersebut <b>dia</b> memandang ojek berbasis <b>aplikasi</b> itu <b>merupakan</b> model bisni baru yang sanggup mengguncang dunia transportasi <b>gojek</b> itu bukan cuma transport orang tapi bikin malas	kalau begitu saya install <b>gojek</b> sekarang menteri dan para hadirin pun sontak tertawa <b>dia</b> kemudian menambahkan bahwa <b>aplikasi seperti gojek merupakan</b> masa depan <b>ekonomi digital indonesia</b> dan berharap akan lebih banyak lagi jenisnya

Berdasarkan tabel 3.4 di atas maka dapat dilihat bahwa *word overlap*( $W_k$ ) pada *vertex* ke-7 dan *vertex* ke-12 berjumlah 9 dan  $\log(|S_7|) + \log(|S_{12}|)$  adalah panjang kata pada tiap *vertex* tersebut. Dengan demikian maka *edge* dapat dihitung berdasarkan rumus *similarity* sesuai persamaan 2.2 berikut :

$$\text{Similarity}(S_7, S_{12}) = \frac{|\{W_k | W_k \in S_7 \& W_k \in S_{12}\}|}{\log(|S_7|) + \log(|S_{12}|)}$$

$$\text{Similarity}(S_7, S_{12}) = \frac{9}{76 + 33} = \frac{9}{109} = 0,082$$

- Menghitung *similarity* antara *vertex* ke-2 dengan *vertex* ke-4

*Word overlap* *vertex* ke-2 dan *vertex* ke-4 dapat dilihat berdasarkan tabel 3.5 berikut :

**Tabel 3.5 Word Overlap pada vertex ke-7 dan vertex ke-2**

Vertex ke-2	Vertex ke-4
com <b>di</b> sela sela kunjungannya ke kantor menteri komunikasi dan informatika menkominfo rudiantara founder google sergey brin mengaku kagum dengan aplikasi gojek dan tertarik memasang aplikasinya brin pun langsung ingin menginst nya saat menteri memberi tahu keberadaan fitur gomassag gomassag merupakan salah satu fitur <b>di</b> aplikasi gojek yang memungkinkan pengguna memesan jasa layanan pijat kunjungan brin pada senin sore itu singkat saja satu jam lebih dia <b>berada</b> di kemenkominfo dengan lebih dari menit dihabiskan berbincang bincang dengan rudiantara di ruang tertutup	saya senang <b>berada</b> <b>di</b> indonesia negara digital ekonomi terbesar <b>di</b> region asia

Berdasarkan tabel 3.5 maka dapat dilihat bahwa *word overlap*(*Wk*) pada *vertex* ke-2 dan *vertex* ke-4 berjumlah 3 dan  $\log(|S_7|) + \log(|S_{12}|)$  adalah panjang kata pada tiap *vertex* tersebut. Dengan demikian maka *edge* dapat dihitung berdasarkan rumus *similarity* sesuai persamaan 2.2 berikut :

$$\text{Similarity}(S_7, S_{12}) = \frac{|\{Wk | Wk \in S_7 \& Wk \in S_{12}\}|}{\log(|S_7|) + \log(|S_{12}|)}$$

$$\text{Similarity}(S_7, S_{12}) = \frac{3}{81 + 12} = \frac{9}{93} = 0,032$$

iii. Menambahkan *score* pada tiap kalimat

Setelah nilai *similarity* antar kalimat telah didapat hasilnya, selanjutnya adalah dilakukan pemberian *score* akhir pada tiap kalimat (*vertex*). Dengan cara menjumlahkan semua *edge* yang mengarah ke *vertex* tersebut seperti terlihat pada gambar 3.4. Hasil dari *score* tiap *vertex* dapat dilihat pada tabel 3.6 berikut :

**Tabel 3.6 Contoh hasil *scoring* pada tiap *vertex***

No.	Kalimat ke( <i>vertex</i> )	Nilai akhir textrank
1	V1	0
2	V2	0,445
3	V3	0,455
4	V4	0,237
5	V5	0,064
6	V6	0,117
7	V7	0,538
8	V8	0,111
9	V9	0,137
10	V10	0,125
11	V11	0,117
12	V12	0,41

iv. Memilih kalimat yang paling penting

Setelah setiap *vertex* diketahui nilainya maka proses selanjutnya adalah memberikat peringkat sesuai dengan besarnya nilai seperti terlihat pada tabel 3.7 berikut :

**Tabel 3.7 Contoh pemringkatan kalimat**

No.	Kalimat ke( <i>vertex</i> )	Nilai akhir textrank
1	V7	0,538
2	V3	0,455
3	V2	0,445
4	V12	0,41
5	V4	0,237
6	V9	0,137
7	V10	0,125
8	V6	0,117
9	V11	0,117
10	V8	0,111
11	V5	0,064
12	V1	0

### 3.2.3 Output

Setelah mendapatkan peringkat setiap *vertex* dengan nilai tertinggi maka selanjutnya dilakukan pemilihan kalimat yang paling penting yaitu dengan memilih berdasarkan peringkat teratas seperti yang terlihat pada tabel 3.7 di atas. Sehingga output yang dihasilkan dari sistem peringkasan teks otomatis yang diharapkan adalah seperti yang terlihat pada gambar 3.5 berikut :

**Output Teks :**

Setelah keluar dari ruangan, keduanya segera menggelar jumpa pers. Rudiantara mengatakan bahwa selama di dalam ruangan tadi, dirinya mengungkapkan ambisi untuk menciptakan 1000 startup lokal berkualitas dalam lima tahun ke depan, mengembangkan ekonomi digital dan roadmap e-commerce. Menteri yang akrab disapa Chief RA itu juga meminta keterlibatan Google di Indonesia untuk mewujudkan ambisi tersebut. Brin menanggapi permintaan dengan hangat. Bos laboratorium Google X itu mengungkapkan bahwa mereka memiliki program untuk mengembangkan startup di Indonesia, yaitu Launchpad Accelerator.

Tentu saja Google akan senang hati mau membantu program digital ekonomi Indonesia, ujarnya. Di sela-sela pembicaraan tersebut, Brin sempat melontarkan pujian untuk Go-Jek. Menurutnya, usaha rintisan Nadiem Makarim menjadi fenomena yang terdengar hingga mancanegara. "Melihat entrepreneurship seperti Go-Jek membuat kami ingin berada di sini (pengembangan startup)," pungkasnya. Rudiantara menanggapi pernyataan tersebut. Dia memandang ojek berbasis aplikasi itu merupakan model bisnis baru yang sanggup mengguncang dunia transportasi. "Go-Jek itu bukan cuma transport orang, tapi bikin malas.

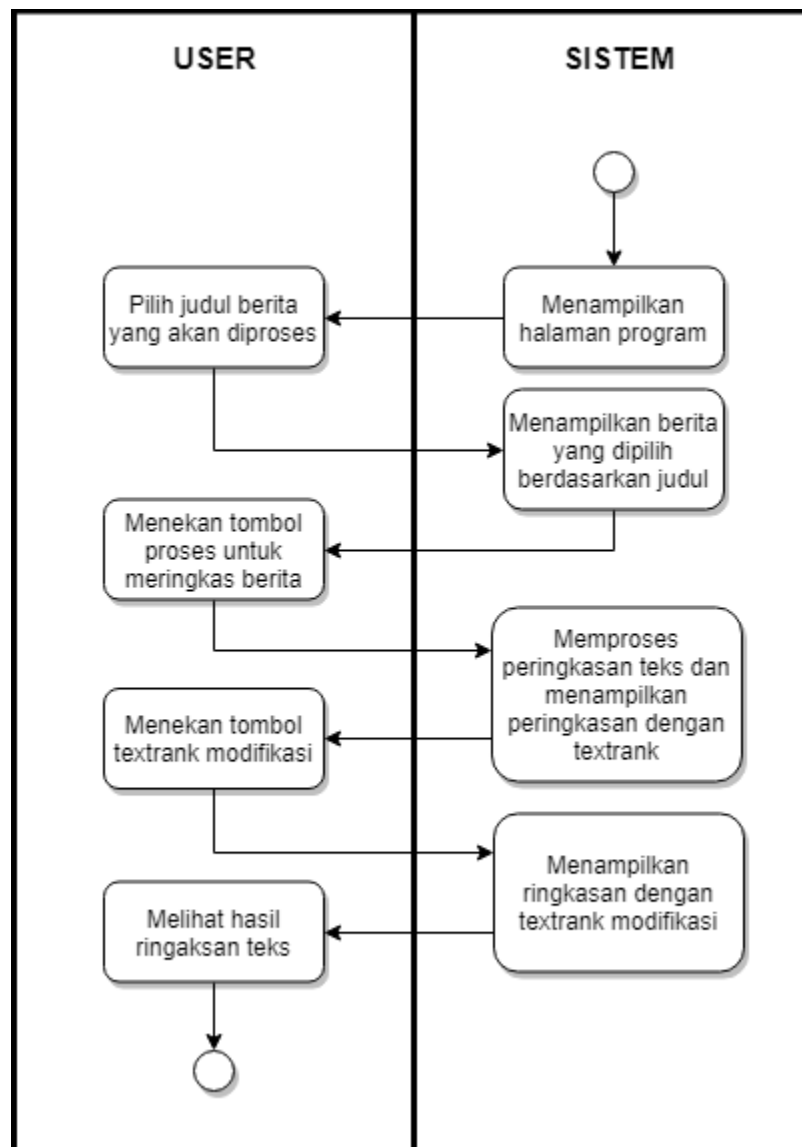
**Gambar 3.5. Contoh Output Peringkasan Menggunakan *Textrank***

### 3.3 Perancangan Sistem

Sistem yang dibuat adalah dengan membuat peringkasan teks otomatis pada teks berita *online* yang telah dipilih. Untuk dapat dilihat dan diuji hasilnya berdasarkan evaluasi yang telah ditetapkan. Proses perancangan sistem terdiri atas beberapa tahapan yaitu :

#### 3.3.1 Diagram Aktifitas untuk Peringkasan Teks Otomatis

Dengan menggunakan diagram *activity* maka alur pertukaran informasi antara *user* dan sistem dapat terjadi sesuai prosedur. Pada gambar 3.6 berikut akan digambarkan proses terjadi peringkasan teks otomatis yang dilakukan oleh sistem dengan menerima *input* dari *user*.



**Gambar 3.6 Diagram Aktivitas Peringkasan Teks Otomatis**

Proses terjadinya peringkasan teks otomatis dimulai saat sistem menampilkan halaman program. Selanjutnya *user* dapat memilih judul berita sesuai dengan yang diminati. Dengan memilih judul berita yang diinginkan maka sistem akan menampilkan teks berita sesuai dengan berita yang dipilih. *User* menekan tombol proses yang selanjutnya sistem akan memproses teks berita yang sudah dipilih untuk dapat mendapatkan hasil ringkasan berdasarkan algoritma *textrank*. Setelah proses selesai dilakukan, selanjutnya sistem akan menampilkan teks yang sudah melalui proses peringkasan. *User* dapat memilih peringkasan lainnya yang menggunakan algoritma *textrank* modifikasi.

### 3.3.2 Perancangan Antarmuka Sistem

Sistem peringkasan berita yang dibangun terdiri dari server aplikasi xampp yang telah diisi beberapa berita *online* yang dipilih oleh penulis.

- Rancangan tampilan halaman *survei*

Pada rancangan halaman *survei* akan digunakan sebagai proses pengambilan data berupa kuesioner yang akan diisi oleh responden. Rancangan tampilan halaman *survei* dapat dilihat pada gambar 3.7 berikut :

**Gambar 3.7 Rancangan Tampilan Halaman *Survei***

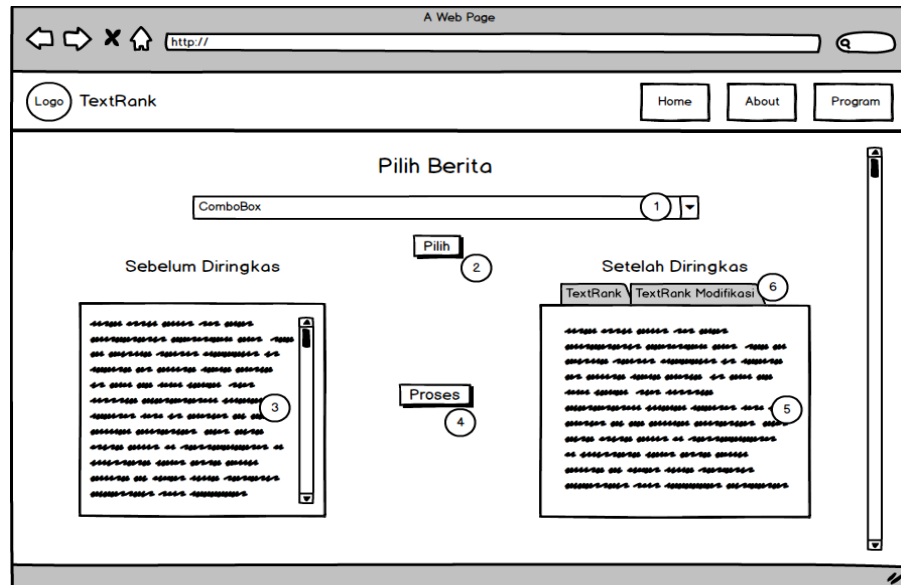
Keterangan dari gambar 3.7 akan dijelaskan sebagai berikut :

- 1) Menampilkan informasi mengenai teks berita yang akan dilakukan survei terdiri dari ID Artikel, Judul Artikel, dan Bidang.
- 2) Survei dilakukan dengan cara responden memilih kalimat penting dari beberapa kalimat yang telah ditampilkan.
- 3) Apabila seluruh kalimat penting telah dipilih pada sebuah artikel maka responden mengkonfirmasi pilihannya dengan menekan tombol Submit.

Survei akan berlanjut ke halaman selanjutnya sampai 100 berita telah terisi surveinya dengan memilih beberapa kalimat penting pada setiap berita.

- Rancangan tampilan halaman program

Rancangan tampilan halaman program berisi pilihan berita yang akan diringkas dan hasil peringkasan yang diinginkan berdasarkan algoritma. Rancangan tampilan halaman program akan ditampilkan pada gambar 3.8 berikut :



**Gambar 3.8 Rancangan Tampilan Halaman Program**

Keterangan dari gambar 3.8 akan dijelaskan sebagai berikut :

- 1) *Combobox* berfungsi untuk melihat berbagai macam judul berita yang akan dipilih untuk diringkas.
- 2) Tombol pilih berfungsi untuk memilih judul berita yang sudah dilihat untuk nantinya diproses oleh sistem.
- 3) Berita yang sudah dipilih akan ditampilkan ke dalam *textarea* yang telah disediakan. Berita yang ditampilkan adalah berita asli.
- 4) Tombol proses berfungsi untuk memproses berita asli ke bentuk yang lebih ringkas.
- 5) *Textarea* menampilkan berita yang telah diproses oleh sistem untuk mendapatkan hasil ringkasannya.
- 6) Terdapat pilihan tampilan ringkasan yang ingin dilihat berdasarkan algoritma yang dipilih yaitu *textrank* atau *textrank* modifikasi dengan *levenshtein distance*.



### 3.4 Evaluasi

Proses evaluasi algoritma *TextRank* dilakukan dengan membandingkan hasil peringkasan oleh sistem dengan hasil yang dilakukan oleh manusia. Oleh karena itu pada proses evaluasi diperlukan responden yang dapat memberikan jawabannya sesuai dengan kapasitas ilmu yang diperlukan, terutama pada sebuah berita maka diperlukan ilmu bidang linguistik. Penelitian ini menggunakan seorang responden berupa ahli pada bidang linguistik yang biodatanya dapat dilihat sebagai berikut :

1. Nama Lengkap : Dr. Dra Liesna Andriany, M.Pd
2. Umur : 55 Tahun
3. Pekerjaan : Dosen
4. Keahlian : Linguistik
5. Alamat : Medan
6. E- mail : liesna.adriany63@gmail.com
7. Beberapa Penelitian :
  - ESTABLISHMENT OF STUDENT'S CHARACTER THROUGH WRITING RELIGIOUS POETRY LEARNING BY USING TYPE COOPERATIVE METHOD THINK PAIR SHARE (TPS). Proceedings of The 3rd Annual International Conference Syiah Kuala University (AIC Unsyiah) 2013. In conjunction with The 2nd International Conference on Multidisciplinary Research (ICMR) 2013. October 2-4, 2013, Banda Aceh, Indonesia.
  - Pemanfaatan Bahan Bekas Sebagai Media Pembelajaran. Portal Garuda Indonesian Publication Index. ISSN(Online) : 2337 – 618X Vol.2 No.1 Hal.277 – 288. Januari -Juni 2014
  - Pembentukan Karakter Melalui Pembelajaran PAKEM. Portal Garuda Indonesian Publication Index. ISSN : 2337 – 6198 Volume 1 No.1 Hal.11 – 18.Januari – Juni 2013

## BAB 4

### IMPLEMENTASI DAN PENGUJIAN SISTEM

Pada bab ini akan dijelaskan tentang proses implementasi algoritma *TextRank*, berdasarkan yang telah dibahas pada bab 3 serta menguji dan mengevaluasi sistem yang telah dibuat.

#### 4.1 Implementasi Sistem

##### 4.1.1 *Spesifikasi hardware dan software yang digunakan*

Adapun spesifikasi perangkat keras yang dibangun untuk membuat sistem pada penelitian ini adalah sebagai berikut :

- a. *Processor Intel ® Core™ i5 – 8250M CPU 1.80 GHz*
- b. *Kapasitas hard disk 1.000 GB*
- c. *Memori RAM 4.00 GB*
- d. *Sistem operasi Windows 10 Home 64-bit*
- e. *Python 2.7.13*

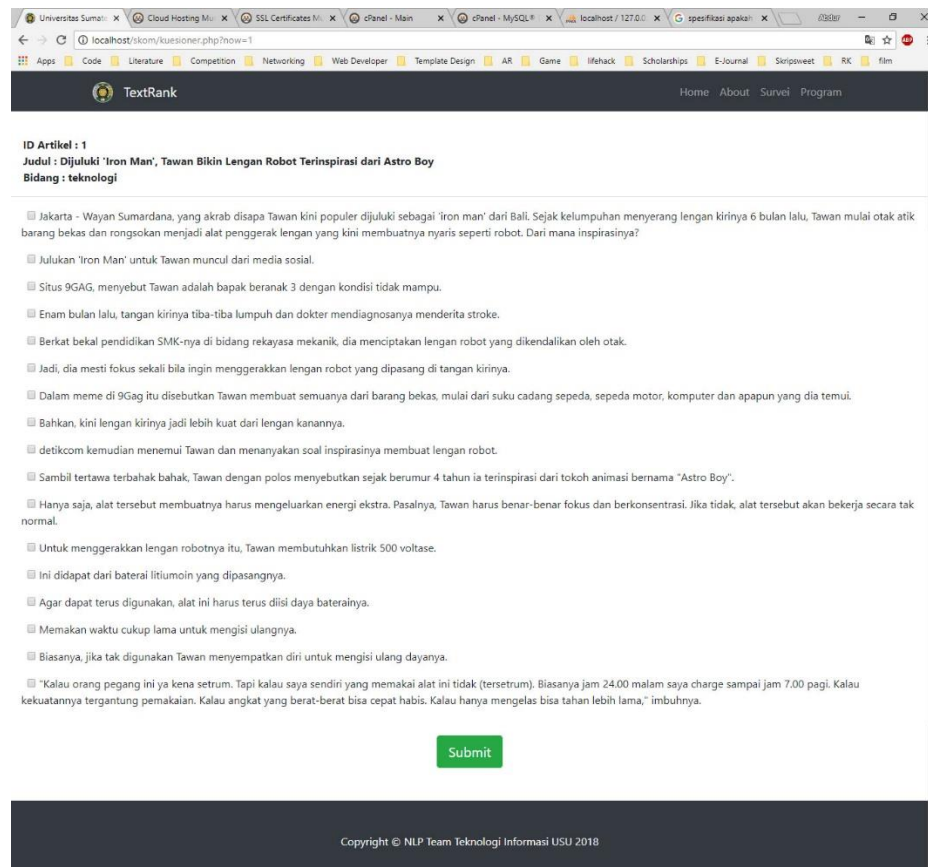
##### *Spesifikasi database server*

Untuk spesifikasi database server yang digunakan dalam pengambilan data melalui responden digunakan seperti sebagai berikut :

- a. *Server : Localhost via UNIX socket*
- b. *Nama Server : sg20.dewaweb.com*
- c. *Nama Domain : www.textrank.web.id*
- d. *CPU Power 1 Core*
- e. *Memory 1 GB*
- f. *Bandwith Unlimited*

##### 4.1.2 *Pengambilan Survei*

Proses pengambilan data melalui survei dilakukan dengan menyajikan halaman *website* yang telah diimplementasi dari rancangan yang telah dibuat sebelumnya melalui *hosting* pada halaman website. Perancangan halaman survei dapat dilihat pada gambar 4.1 berikut :



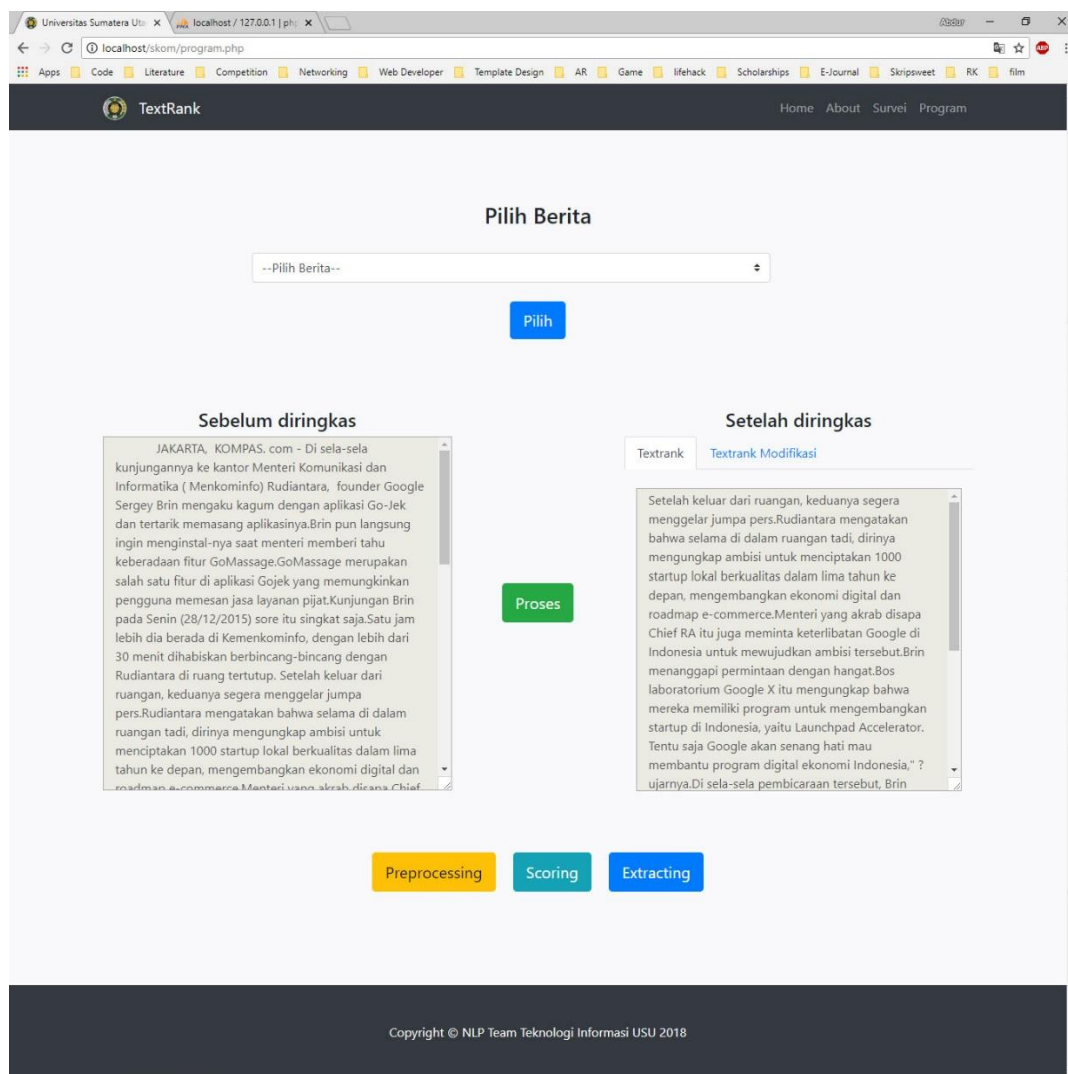
**Gambar 4.1 Tampilan Halaman Survei**

Pada gambar 4.1 dapat dilihat bahwa proses pengambilan survei dilakukan dengan cara responden menekan tombol survei. Selanjutnya halaman survei akan menampilkan berita berdasarkan urutan mulai dari 1 – 100 berita per halaman. Responden memilih beberapa kalimat penting yang ada pada setiap berita yang disajikan. Survei selesai jika keseluruhan 100 berita telah dipilih kalimat pentingnya oleh responden.

#### 4.1.3 Implementasi Perancangan Antar Muka

##### Tampilan Halaman Utama

Pada tampilan halaman utama *user* dapat memilih judul berita yang akan diringkas dengan cara memilih judul pada *combobox* yang telah disediakan. Halaman utama dapat dilihat pada gambar 4.2 berikut dimana setelah *user* telah memilih judul berita yang diinginkan untuk diringkas lalu menekan tombol proses sehingga sistem dapat meringkas berita ke dalam ringkasan yang lebih pendek. Selanjutnya *user* dapat memilih peringkasan yang disajikan menggunakan algoritma *textrank* atau *textrank* modifikasi.



**Gambar 4.2 Tampilan Halaman Utama**

## 4.2 Pengujian Sistem

Sistem yang telah dibuat akan diuji dengan cara meringkas teks berita asli ke bentuk yang lebih ringkas. Sebagai contoh pengujian sistem akan disajikan beberapa contoh teks yang dapat dilihat pada Lampiran A. Dengan menggunakan sistem peringkasan teks otomatis maka hasil ringkasan yang dilakukan mendapatkan hasil yang lebih ringkas seperti dapat terlihat pada tabel 4.1 berikut :

**Tabel 4.1 Hasil Peringkasan Teks Otomatis**

No.	Jumlah Teks Asli (kata)	Frekuensi Berita	Rata - rata Ringkasan (kata)	Persentase Hasil Ringkasan (%)	Ringkasan Responden (kata)	Persentase Hasil Ringkasan Responden (%)
1	>800	22	373	46%	256	30%
2	700 - 800	12	341	45%	203	27%
3	600 - 700	11	304	46%	178	27%
4	500 - 600	22	269	48%	123	22%
5	>500	33	143	39%	78	17%

Dari tabel 4.1 dapat dilihat dengan menggunakan algoritma *TextRank* pada peringkasan teks otomatis menghasilkan ringkasan dengan rata – rata ringkasan 44,8% lebih ringkas dari teks asli, sedangkan responden berhasil meringkas dengan rata – rata 24,6%. Pada pengujian terhadap 100 teks berita *TextRank* hanya dapat meringkas 67 berita lebih ringkas, sedangkan sisanya menghasilkan ringkasan yang sama dengan teks asli ini terjadi dikarenakan pada saat bobot penghitungan nilai *similarity* pada *TextRank* mendapatkan bobot hasil yang sama antar tiap *token* dengan demikian seluruh kalimat pada teks dianggap penting. Sedangkan pengujian dengan *TextRank* dengan perhitungan jarak melalui *levenshtein distance* (*TextRank* modifikasi) mendapatkan 98 hasil berita lebih ringkas dari teks asli.

#### 4.3 Evaluasi Hasil Ringkasan

Evaluasi pada penelitian ini dilakukan pada 100 berita dengan menggunakan metode *Recall-Oriented Understudy for Gisting Evaluation (ROUGE)*. Metode *ROUGE* melakukan evaluasi dengan cara membandingkan hasil ringkasan yang dibuat oleh sistem dengan ringkasan hasil dari manusia. *ROUGE* sendiri memiliki beberapa perhitungan, namun pada penelitian ini hanya menggunakan *ROUGE-N(n-gram)* yang terdiri dari *ROUGE-1* dan *ROUGE-2*. Perbedaan dari *ROUGE-1* dan *ROUGE-2* terletak pada cara mengambil perhitungan akurasi. Jika pada *ROUGE-1* membandingkan kata per kata untuk dibandingkan dengan ringkasan yang dibuat oleh manusia, sebaliknya *ROUGE-2* membandingkan keseluruhan ringkasan yang dilakukan sistem dengan keseluruhan ringkasan yang dilakukan oleh manusia. Pada penelitian ini, peneliti berhasil mendapatkan ringkasan teks dari seorang responden yang ahli dalam bidang linguistik yang hasilnya akan dibandingkan dengan ringkasan oleh sistem. Hasil yang diharapkan

pada penelitian ini adalah untuk mengetahui seberapa akurat algoritma *textrank* dan algoritma *textrank* menggunakan *levenshtein distance* (*Textrank* modifikasi) untuk meringkas teks berbahasa Indonesia dengan menilai dari hasil *F-Score ROUGE* yang telah dihitung menggunakan persamaan 2.4 di atas. Jika hasil *F-Score* yang didapat 0(nol) maka ringkasan tidak sama, sebaliknya jika bernilai 1 maka ringkasan sama persis dengan ringkasan yang dilakukan oleh ahli. Untuk mengetahui hasilnya maka dapat melihat pada tabel 4.2 berikut :

**Tabel 4.2 Hasil Evaluasi Peringkasan *ROUGE-1 Textrank*(a) dan *Textrank Levenshtein Distance*(b)**

Rentang F - Score	F - Score	Frekuensi Berita	Total Frekuensi (%)	Rentang F - Score	F - Score	Frekuensi Berita	Total Frekuensi (%)
0,7 - 1	0,81	2	5%	0,7 - 1	0,74	1	1%
	0,79	1					
	0,76	1					
	0,73	1					
0,5 - 0,699	0,67	1	31%	0,5 - 0,699	0,68	1	22%
	0,66	1			0,66	1	
	0,65	3			0,62	1	
	0,64	1			0,61	2	
	0,61	1			0,6	2	
	0,6	1			0,58	3	
	0,59	3			0,57	2	
	0,58	3			0,56	2	
	0,57	1			0,55	1	
	0,55	1			0,54	2	
	0,54	2			0,52	1	
	0,52	5			0,51	3	
	0,51	4			0,5	1	
	0,5	4			0,4 - 0,499	0,49	
0,49	4	0,48	4				
0,47	3	0,47	3				
0,46	2	0,46	2				
0,45	1	0,44	2				
0,44	3	0,45	1				
0,43	3	0,43	2				
0,4 - 0,499	0,42	3	23%	0,4 - 0,499	0,42	5	19%
	0,41	2			0,41	2	
	0,4	2			0,4	6	
					0,39	1	
					0,38	3	
			0,37	1			

(a)

(b)

**Tabel 4.2 Hasil Evaluasi Peringkasan *ROUGE-1 Textrank(a)* dan *Textrank Levenshtein Distance(b)* (Lanjutan)**

Rentang F - Score	F - Score	Frekuensi Berita	Total Frekuensi (%)
0,3 - 0,399	0,39	5	20%
	0,38	3	
	0,37	1	
	0,36	5	
	0,35	3	
	0,32	1	
	0,31	1	
	0,3	1	
< 0,3	....	21	21%

(a)

Rentang F - Score	F - Score	Frekuensi Berita	Total Frekuensi (%)
0,3 – 0,399	0,36	1	
	0,35	4	
	0,34	1	
	0,33	2	
	0,32	5	
	0,3	1	
< 0,3	....	29	29%

(b)

Hasil peringkasan *ROUGE-1* yang dilakukan pada tabel 4.1 membandingkan tiap kata yang ada pada sistem dengan referensi ringkasan ideal. Dapat dilihat bahwa peringkasan teks dengan algoritma *textrank* (a) memiliki hasil terbanyak sebesar 31% berita yang memiliki nilai *F-Score* antara 0,5 – 0,699 artinya peringkasan lebih mendekati 1, dimana hasil ini lebih baik dibanding dengan peringkasan yang dilakukan dengan algoritma *textrank* modifikasi(b) dimana rata – rata 29% peringkasan memiliki nilai *F-Score* antara 0,4 – 0,499 dan lebih kecil dari 0,3 yang artinya sangat jauh berbeda dari ringkasan ideal. Selanjutnya akan dibandingkan hasil peringkasan menggunakan *ROUGE-2* yang hasilnya dapat dilihat pada tabel 4.3 sebagai berikut :

**Tabel 4.3 Hasil Evaluasi Peringkasan *ROUGE-2 Textrank(a)* dan *Textrank Levenshtein Distance(b)***

Rentang F - Score	F - Score	Frekuensi Berita	Total Frekuensi (%)
0,7 - 1	0,74	1	3%
	0,72	1	
	0,71	1	
0,5 - 0,699	0,67	1	8%
	0,66	1	
	0,59	1	

(a)

Rentang F - Score	F - Score	Frekuensi Berita	Total Frekuensi (%)
0,7 - 1	0,74	-	1%
	0,72	-	
	0,7	1	
0,5 - 0,699	0,62	1	8%
	0,6	1	
	0,57	1	

(b)

**Tabel 4.3 Hasil Evaluasi Peringkasan *ROUGE-2 Textrank(a)* dan *Textrank Levenshtein Distance(b)* (Lanjutan)**

Rentang F - Score	F - Score	Frekuensi Berita	Total Frekuensi (%)	Rentang F - Score	F - Score	Frekuensi Berita	Total Frekuensi (%)
0,7 - 1	0,74	1	3%	0,7 - 1	0,74	-	1%
	0,72	1			0,72	-	
	0,71	1			0,7	1	
0,5 - 0,699	0,67	1	8%	0,5 - 0,699	0,62	1	8%
	0,66	1			0,6	1	
	0,59	1			0,57	1	
	0,58	1			0,55	1	
	0,56	1			0,53	2	
	0,55	2			0,52	1	
	0,54	1			0,5	1	
0,4 - 0,499	0,49	3	22%	0,4 - 0,499	0,49	3	12%
	0,48	3			0,47	1	
	0,47	3			0,46	1	
	0,45	2			0,45	1	
	0,44	1			0,43	3	
	0,43	3			0,42	1	
	0,42	4			0,41	1	
	0,41	1			0,4	1	
	0,4	2					
0,3 - 0,399	0,39	2	20%	0,3 - 0,399	0,39	1	22%
	0,38	3			0,38	1	
	0,37	5			0,37	4	
	0,36	1			0,35	3	
	0,34	1			0,34	1	
	0,33	2			0,33	2	
	0,32	2			0,32	4	
	0,31	3			0,31	4	
	0,3	1			0,3	2	
< 0,3	....	47	47%	< 0,3	....	57	57%

(a)

(b)

Pada tabel 4.3 evaluasi peringkasan teks dengan menggunakan *ROUGE-2* atau dengan membandingkan seluruh ringkasan dengan ringkasan ideal yang dilakukan oleh manusia maka algoritma *textrank* (a) memiliki nilai *F-Score* lebih kecil dari 0,3 sebanyak 47% jumlah ini lebih sedikit dibandingkan dengan algoritma *Textrank* modifikasi (b) dimana nilai *F-Score* lebih kecil dari 0,3 sebanyak 57%. Artinya algoritma *TextRank* maupun *TextRank* modifikasi memiliki hasil yang sangat jauh dari ringkasan ideal jika



dibandingkan dari keseluruhan ringkasan yang telah dibuat. Dari hasil perhitungan *ROUGE-1* maupun *ROUGE-2* maka dapat diambil nilai rata – rata dari *F-Score* seperti pada tabel 4.4 berikut :

**Tabel 4.4 Hasil Rata – Rata Evaluasi *F-Score***

	<i>Textrank</i>	<i>TextRank</i> Modifikasi
<i>ROUGE -1</i>	0,439	0,3999
<i>ROUGE -2</i>	0,3186	0,2805

## BAB 5

### KESIMPULAN DAN SARAN

Pada bab ini akan dijelaskan terkait kesimpulan dan saran sesuai dengan hasil yang telah didapat dan diuji pada penelitian ini yaitu peringkasan teks otomatis menggunakan algoritma *Textrank*.

#### 5.1 Kesimpulan

- Menggunakan metode *Textrank* dalam meringkas kalimat penting pada teks berbahasa Indonesia dapat dilakukan dengan mudah tanpa memerlukan data *training* dengan menghasilkan nilai rata – rata *F-Score* 0,439 pada *ROUGE -1* dan 0,3186 pada *ROUGE-2*. Hasil ini terbilang cukup baik karena berdasarkan hasil uji *TextRank* pada teks berbahasa Inggris (dapat dilihat pada tabel 2.1) mendapatkan hasil rata – rata 0,4708. Artinya tidak terlalu jauh perbedaan hasil menggunakan teks berbahasa Indonesia maupun teks berbahasa Inggris.
- Nilai *F-Score* pada *ROUGE-2* cukup rendah diakibatkan karena pada saat *preprocessing* tepatnya *tokenizing* terdapat kata dalam Bahasa Indonesia yang dianggap seperti Bahasa Inggris sebagai contoh setiap akhiran s pada suatu kata maka akan dianggap sebagai *plural* dan perlu untuk dihilangkan s pada akhiran kata tersebut, sehingga mengubah makna dari kata.
- Jika dibandingkan *Textrank* dengan *Textrank* menggunakan *Levenshtein Distance*(*Textrank* modifikasi) maka akurasi *Textrank* lebih tinggi untuk meringkas sebuah berita, tetapi pada saat meringkas 100 berita metode. Akan tetapi *Textrank* modifikasi dapat meringkas 98 berita jika dibandingkan dengan *Textrank* yang hanya dapat meringkas 63 berita.

#### 5.2 Saran

- Memodifikasi algoritma *Textrank* agar akurasi yang dihasilkan dalam meringkas teks berbahasa Indonesia mendekati ringkasan ideal, terutama tahapan *preprocessing*.
- Mengimplementasikan algoritma *Textrank* tidak hanya pada peringkasan teks melainkan pada penerjemah teks dari bahasa asing ke bahasa Indonesia.

## Daftar Pustaka

- Balas, Valentine E. dan Banerjee, Soumya. 2014. Seeded Text Auto-Summarization : An Experience with Simplified Statistical and Fuzzy Ranking Algorithm. *15<sup>th</sup> IEEE International Symposium on Computational Intelligence and Informatics*. Budapest, Hungary. pp.143 – 146.
- Chatterjee, N. Mittal, A. dan Goyal, S. 2012. Single Document Extractive Text Summarization Using Genetic Algorithms. *Third International Conference on Emerging Applications of Information Technology (EAIT)*. India. pp. 19 – 23.
- Farizi, Salman Al. 2015. Rekomendasi Tag Pada Berita Online Menggunakan Tf-idf Dan Collaborative Tagging. Skripsi. Universitas Sumatera Utara.
- Gunawan, D., Amalia, A., & Charisma, I. 2016. Automatic extraction of multiword expression candidates for Indonesian language. Control System, Computing and Engineering (ICCSCE), 2016 6th IEEE International Conference, pp. 304-309.
- Gunawan, D., Pasaribu, A., Rahmat, R. F., & Budiarto, R. 2017. Automatic Text Summarization for Indonesian Language Using TextTeaser. IOP Conference Series: Materials Science and Engineering Vol. 190, No. 1, p. 012048. IOP Publishing.
- Gunawan, D., Amalia, A., & Charisma, I. 2017. Clustering Articles in Bahasa Indonesia Using Self-Organizing Map. *International Conference on Electrical Engineering and Informatics (ICELTICs)*. Banda Aceh, Indonesia.
- Jaradat, Yazan A. 2016. Hybrid-based Arabic single-document text summarization approach using genetic algorithm. *2016 7<sup>th</sup> International Conference on Information and Communication Systems (ICICS)*. Irbid, Jordan.

- Jayashree, R. dan Murthy, S. Categorized Text Document Summarization in the Kannada Language by Sentence Ranking. *2012 12<sup>th</sup> International Conference on Intelligent Systems Design and Applications (ISDA)*. India. pp.776 – 781.
- Lin, C.-Y., dan Hovy E. Automatic Evaluation of Summaries Using N-gram Co-Occurrence *In Proceedings of the Human Technology Conference 2003 (HLT-NAACL-2003)*, Edmonton, Canada.
- Lin, C.-Y., 2004. Rouge: A package for automatic evaluation of summaries. Workshop on Text Summarization Branches Out, Post-Conference Workshop of ACL 2004, Barcelona, Spain.
- Massandy, Danang Tri. dan Khodra, Mesayu Leylia. 2014. Guided Summarization for Indonesian News Articles. *2014 International Conference of Advance Informatics: Concept, Theory, and Application (ICAICTA)*. Institut Teknologi Bandung, Bandung. pp. 7 – 10.
- Mihalcea, R. dan Tarau, P. 2004. TextRank : Bringing Order into Texts. *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, Barcelona.
- Mihalcea, R. dan Tarau, P. 2005. A Language Independent Algorithm for Single and Multiple Document Summarization. University of North Texas, USA.
- Nguyen, Q.U., Pham T.A., Truong C.D. dan Nguyen X.H. 2012. A Study On The Use of Genetic Programming for Automatic Text Summarization. *2012 4<sup>th</sup> International Conference on Knowledge and Systems Engineering*. Vietnam, pp.93 – 98.
- Pasaribu, Anwar. 2016. Sistem Peringkat Berita Online Menggunakan Algoritma TextTeaser. Skripsi. Universitas Sumatera Utara.
- Ramadhiana, Ranti. 2017. Ekstraksi Kata Kunci Otomatis Teks Berbahasa Indonesia Menggunakan Metode TextRank. Skripsi. Universitas Sumatera Utara.

**Lampiran A**  
**CONTOH BERITA**

<b>ID : 1</b>
<b>Judul :</b> <i>Dijuluki 'Iron Man', Tawan Bikin Lengan Robot Terinspirasi dari Astro Boy</i>
<b>Ringkasan TextRank :</b> <p>Dari mana inspirasinya? Julukan 'Iron Man' untuk Tawan muncul dari media sosial. Situs 9GAG, menyebut Tawan adalah bapak beranak 3 dengan kondisi tidak mampu. Enam bulan lalu, tangan kirinya tiba-tiba lumpuh dan dokter mendiagnosanya menderita stroke. Berkat bekal pendidikan SMK-nya di bidang rekayasa mekanik, dia menciptakan lengan robot yang dikendalikan oleh otak. Jadi, dia mesti fokus sekali bila ingin menggerakkan lengan robot yang dipasang di tangan kirinya.</p> <p>Dalam meme di 9Gag itu disebutkan Tawan membuat semuanya dari barang bekas, mulai dari suku cadang sepeda, sepeda motor, komputer dan apapun yang dia temui. Bahkan, kini lengan kirinya jadi lebih kuat dari lengan kanannya.</p> <p> detikcom kemudian menemui Tawan dan menanyakan soal inspirasinya membuat lengan robot. Sambil tertawa terbahak-bahak, Tawan dengan polos menyebutkan sejak berumur 4 tahun ia terinspirasi dari tokoh animasi bernama "Astro Boy". Menginjak usia dewasa, tak disangka kekaguman Tawan pada Astro Boy malah menjadi inspirasi utama dirinya menjadi manusia robot. Kondisi fisik mendesaknya untuk bergerak memaksa diri menghadirkan wujud tokoh animasi Astro Boy dalam dirinya.</p> <p>"Dari umur 4 tahun saya sudah ngefans sekali sama Astro Boy. Nggak nyangka saking senengnya sama Astro Boy malah jadi inspirasi saya untuk bikin tangan robot ini," ujar Tawan ketika berbincang dengan detikcom, Kamis (21/1/2016).</p> <p>Komponen alat buatan Tawan mulai dari besi penopang, tabung hidrolis, hingga beberapa kabel terlihat bekas. Tawan merancang, merakit dan menyambung sendiri seluruh item tersebut. Tak hanya itu, dia juga menggunakan CPU komputer, dinamo, tuning potensio, sensor ultasonik, sensor infra merah dan sensor jumlah putaran dinamo.</p> <p>"Alat ini saya ciptakan untuk aktivitas saya. Dan saya rangkai sendiri," ucapnya.</p> <p>Satu komponen kunci, yakni sensor EEG (Electro Encephalo Graphi) dibelinya online dari Amerika Serikat (AS) senilai Rp 4,7 juta. Ia menjelaskan, alat ini berfungsi melalui sensor otak yang dipasang di kepala yang mengendalikan arah gerak ke tangan kirinya melalui alat yang dipasang di punggung dan tangan kirinya.</p> <p>Untuk memuluskan gerak alat tersebut, Tawan memasang satu unit CPU komputer pun dipasang di bagian belakang tubuhnya. Jika tidak, alat tersebut akan bekerja secara tak normal. Untuk menggerakkan lengan robotnya itu, Tawan membutuhkan listrik 500 voltase. Ini didapat dari baterai litiumoin yang dipasangnya. Agar dapat terus digunakan, alat ini harus terus diisi daya baterainya. Memakan waktu cukup lama untuk mengisi ulangannya. Biasanya, jika tak digunakan Tawan menyempatkan diri untuk mengisi ulang dayanya. "Kalau orang pegang ini ya kena setrum."</p>

**Ringkasan Ahli :**

Jakarta - Wayan Sumardana, yang akrab disapa Tawan kini populer dijuluki sebagai 'iron man' dari Bali. Sejak kelumpuhan menyerang lengan kirinya 6 bulan lalu, Tawan mulai otak atik barang bekas dan rongsokan menjadi alat penggerak lengan yang kini membuatnya nyaris seperti robot. Dari mana inspirasinya?

Berkat bekal pendidikan SMK-nya di bidang rekayasa mekanik, dia menciptakan lengan robot yang dikendalikan oleh otak. Dalam meme di 9Gag itu disebutkan Tawan membuat semuanya dari barang bekas, mulai dari suku cadang sepeda, sepeda motor, komputer dan apapun yang dia temui.

Menginjak usia dewasa, tak disangka kekaguman Tawan pada Astro Boy malah menjadi inspirasi utama dirinya menjadi manusia robot. Tawan merancang, merakit dan menyambung sendiri seluruh item tersebut. Ia menjelaskan, alat ini berfungsi melalui sensor otak yang dipasang di kepala yang mengendalikan arah gerak ke tangan kirinya melalui alat yang dipasang di punggung dan tangan kirinya.

Untuk menggerakkan lengan robotnya itu, Tawan membutuhkan listrik 500 voltase. Agar dapat terus digunakan, alat ini harus terus diisi daya baterainya.

**Teks Asli :**

Jakarta - Wayan Sumardana, yang akrab disapa Tawan kini populer dijuluki sebagai 'iron man' dari Bali. Sejak kelumpuhan menyerang lengan kirinya 6 bulan lalu, Tawan mulai otak atik barang bekas dan rongsokan menjadi alat penggerak lengan yang kini membuatnya nyaris seperti robot. Dari mana inspirasinya?

Julukan 'Iron Man' untuk Tawan muncul dari media sosial. Situs 9GAG, menyebut Tawan adalah bapak beranak 3 dengan kondisi tidak mampu. Enam bulan lalu, tangan kirinya tiba-tiba lumpuh dan dokter mendiagnosanya menderita stroke. Berkat bekal pendidikan SMK-nya di bidang rekayasa mekanik, dia menciptakan lengan robot yang dikendalikan oleh otak. Jadi, dia mesti fokus sekali bila ingin menggerakkan lengan robot yang dipasang di tangan kirinya.

Dalam meme di 9Gag itu disebutkan Tawan membuat semuanya dari barang bekas, mulai dari suku cadang sepeda, sepeda motor, komputer dan apapun yang dia temui. Bahkan, kini lengan kirinya jadi lebih kuat dari lengan kanannya.

detikcom kemudian menemui Tawan dan menanyakan soal inspirasinya membuat lengan robot. Sambil tertawa terbahak bahak, Tawan dengan polos menyebutkan sejak berumur 4 tahun ia terinspirasi dari tokoh animasi bernama "Astro Boy". Menginjak usia dewasa, tak disangka kekaguman Tawan pada Astro Boy malah menjadi inspirasi utama dirinya menjadi manusia robot. Kondisi fisik mendesaknya untuk bergerak memaksa diri menghadirkan wujud tokoh animasi Astro Boy dalam dirinya.

"Dari umur 4 tahun saya sudah ngefans sekali sama Astro Boy. Nggak nyangka saking senengnya sama Astro Boy malah jadi inspirasi saya untuk bikin tangan robot ini," ujar Tawan ketika berbincang dengan detikcom, Kamis (21/1/2016).

Komponen alat buatan Tawan mulai dari besi penopang, tabung hidrolis, hingga beberapa kabel terlihat bekas. Tawan merancang, merakit dan menyambung sendiri seluruh item tersebut. Tak hanya itu, dia juga menggunakan CPU komputer, dinamo, tuning potensio, sensor ultasonik, sensor infra merah dan sensor jumlah putaran dinamo.

"Alat ini saya ciptakan untuk aktivitas saya. Ada sebagian yang saya buat dan sebagian saya beli. Dan saya rangkai sendiri," ucapnya.

Satu komponen kunci, yakni sensor EEG (Electro Encephalo Graphi) dibelinya online dari Amerika Serikat (AS) senilai Rp 4,7 juta. Ia menjelaskan, alat ini berfungsi melalui sensor otak yang dipasang di kepala yang mengendalikan arah gerak ke tangan kirinya melalui alat yang dipasang di punggung dan tangan kirinya.

Untuk memuluskan gerak alat tersebut, Tawan memasang satu unit CPU komputer pun dipasang di bagian belakang tubuhnya. Fungsinya sebagai penggerak dari sensor di kepala. Tuning potensio merupakan rangkaian pengolah input dan output mikro kontroler.

"Infra merah, sensor ultrasonik, dan sensor jumlah putaran dinamo ini adalah rangkaian penguat power. Ada pula EEG. Semuanya tersambung ke dinamo agar dapat bekerja secara maksimal," jelasnya.

"Lalu ada drone, elektroda dan lainnya. Posisinya ditempel di kepala sebagai penangkap sinyal, alpa, delta, beta dan teta," tambahnya.

Hanya saja, alat tersebut membuatnya harus mengeluarkan energi ekstra. Pasalnya, Tawan harus benar-benar fokus dan berkonsentrasi. Jika tidak, alat tersebut akan bekerja secara tak normal.

Untuk menggerakkan lengan robotnya itu, Tawan membutuhkan listrik 500 voltase. Ini didapat dari baterai litiumoin yang dipasangnya. Agar dapat terus digunakan, alat ini harus terus diisi daya baterainya. Memakan waktu cukup lama untuk mengisi ulang. Biasanya, jika tak digunakan Tawan menyempatkan diri untuk mengisi ulang dayanya.

"Kalau orang pegang ini ya kena setrum. Tapi kalau saya sendiri yang memakai alat ini tidak (tersetrum). Biasanya jam 24.00 malam saya charge sampai jam 7.00 pagi. Kalau kekuatannya tergantung pemakaian. Kalau angkat yang berat-berat bisa cepat habis. Kalau hanya mengelap bisa tahan lebih lama," imbuhnya.

**ID : 2**

**Judul Berita :** Ada Apa di Balik Penggratisan WhatsApp?

**Ringkasan *Textrank* :**

Lalu yang paling penting, "Anda bisa yakin bahwa tak akan ada iklan yang mengganggu."Selama hampir dua tahun pasca akuisisi Facebook, WhatsApp memegang teguh janjinya dan menghadirkan pengalaman yang konsisten dan "bersih" untuk pengguna.

Tak ada iklan, game, stiker, kanal jualan, atau segala macam hal lain yang umum merecoki pengguna layanan sejenis di luar WhatsApp. Di luar tambahan fitur seperti voice call dan akses dari desktop, layanan pesan instan yang memulai hidup sebagai alat update status ini bisa dibilang setia pada bentuknya yang lama. Pun demikian dengan jalur pemasukan WhatsApp yang dari dulu hanya bersumber dari iuran tahunan. Namun kini telah terjadi perubahan.

Pada 18 Januari 2016, WhatsApp mengumumkan bakal mencabut biaya pemakaian sebesar 1 dollar AS per tahun. Artinya, layanan ini sepenuhnya tersedia gratis. Kebijakan tersebut sudah mulai berlaku tak lama kemudian, pada 20 Januari 2016, kala pengguna WhatsApp disambut pemberitahuan bahwa aplikasi pesan instan itu bisa dipakai secara cuma-cuma "seumur hidup".

Uang dari mana?

Dengan memangkas iuran tahunan, WhatsApp kehilangan satu-satunya sumber pendapatan. Lalu dari mana perusahaan ini akan mendapatkan uang?

Lagi-lagi, Koum menampik kekhawatiran bahwa pihaknya bakal melakukan sesuatu yang berpotensi mengganggu kenyamanan pengguna. "Orang-orang mungkin bertanya bagaimana kami menjalankan WhatsApp tanpa menarik biaya langganan, atau khawatir bahwa pengumuman ini bakal menandai kehadiran iklan pihak ketiga," tulis Koum dalam sebuah posting blog. "Jawabannya adalah tidak akan ada iklan." Sebagai ganti pemasukan yang hilang, WhatsApp menyatakan bakal menjajaki kemungkinan menawarkan sejumlah layanan berbayar untuk pengguna korporat.

Layanan ini nantinya bisa dipakai oleh entitas bisnis untuk berkomunikasi dengan pelanggan masing-masing tanpa perlu membayar biaya tambahan lain. Koum mencontohkan komunikasi dalam proses verifikasi transaksi antara konsumen dengan bank, lalu bisa juga pemberitahuan tertundanya penerbangan dari maskapai ke calon penumpang. "Semua pesan tersebut sekarang disampaikan melalui jalur lain, yakni lewat telepon dan SMS.

**Ringkasan Ahli :**

KOMPAS.com - Awal tahun 2014, perusahaan penyedia layanan pesan instan WhatsApp diakuisisi oleh Facebook dengan nilai yang fantastis: 19 miliar dollar AS atau sekitar Rp 223 triliun. Selain membuat para pemilik dan karyawannya kebanjiran duit - CEO WhatsApp, Jan Koum, konon mendapat "jatah" 6,8 miliar dollar AS- ketika itu banyak yang mempertanyakan motif Facebook mencaplok WhatsApp.



Melalui sebuah pernyataan resmi, Jan Koum menyatakan WhatsApp akan tetap keukeuh dengan model bisnisnya yang lama, yakni memungut iuran tahunan sebesar 1 dollar AS setelah pelanggan melalui masa pakai gratis selama setahun.

Pada 18 Januari 2016, WhatsApp mengumumkan bakal mencabut biaya pemakaian sebesar 1 dollar AS per tahun. Lalu dari mana perusahaan ini akan mendapatkan uang?

Saat berbicara di Konferensi Digital Life Design di Munich, Jerman, awal minggu ini, Koum menjelaskan bahwa WhatsApp mencoba menerapkan model bisnis "commercial-participation", sebuah bentuk evolusi model "freemium" di mana hanya pengguna tertentu saja yang dikenai biaya layanan.

Langkah WhatsApp memasuki segmen korporat menimbulkan sedikit kekhawatiran soal keamanan informasi sensitif yang banyak beredar di bidang ini. Soal ini, WhatsApp menyatakan bahwa tiap pesan yang dikirim melalui layanannya dilindungi oleh enkripsi.

WhatsApp juga tak menyimpan pesan pengguna di dalam server miliknya, dengan pengecualian pesan belum terkirim yang akan disimpan di server dengan waktu maksimal 30 hari sebelum dihapus. Namun langkah-langkah di atas tak menjamin seratus persen bahwa informasi pengguna yang tersimpan dan dikirim lewat WhatsApp tak bakal diintip oleh pihak lain. Basis pengguna WhatsApp berikut data-data mereka yang sangat besar jumlahnya boleh jadi akan menjadi magnet abadi yang akan selalu mengundang kekhawatiran bakal dieksploitasi, entah oleh hacker, otoritas, ataupun pengiklan.

#### **Teks Asli :**

KOMPAS.com - Awal tahun 2014, perusahaan penyedia layanan pesan instan WhatsApp diakuisisi oleh Facebook dengan nilai yang fantastis: 19 miliar dollar AS atau sekitar Rp 223 triliun.

Selain membuat para pemilik dan karyawannya kebanjiran duit -CEO WhatsApp, Jan Koum, konon mendapat "jatah" 6,8 miliar dollar AS- ketika itu banyak yang mempertanyakan motif Facebook mencaplok WhatsApp.

Sang raksasa jejaring sosial tersebut dicurigai ingin menguasai basis data pengguna WhatsApp yang pada saat pengumuman akuisisi sudah mencapai angka ratusan juta.

Basis pengguna WhatsApp ibarat peti harta karun buat Facebook, perusahaan yang sebagian besar pemasukannya berasal dari iklan.

Bayangkan duit yang bakal masuk ke rekening Facebook kalau perusahaan tersebut bisa, misalnya, menyalurkan promosi dari klien ke tiap nomor telepon yang menjadi identitas dari para pengguna WhatsApp.

Toh, kekhawatiran itu segera ditampik oleh pihak WhatsApp.

Melalui sebuah pernyataan resmi, Jan Koum menyatakan WhatsApp akan tetap keukeuh dengan model bisnisnya yang lama, yakni memungut iuran tahunan sebesar 1 dollar AS setelah pelanggan melalui masa pakai gratis selama setahun.

"Tak akan ada perubahan apapun bagi pengguna kami," kata Koum ketika itu dalam blog resmi perusahaan.

"WhatsApp akan tetap beroperasi secara independen. Anda akan tetap bisa menggunakannya secara berbayar," tambah dia. Lalu yang paling penting, "Anda bisa yakin bahwa tak akan ada iklan yang mengganggu."

Selama hampir dua tahun pasca akuisisi Facebook, WhatsApp memegang teguh janjinya dan menghadirkan pengalaman yang konsisten dan "bersih" untuk pengguna.

Tak ada iklan, game, stiker, kanal jualan, atau segala macam hal lain yang umum merecoki pengguna layanan sejenis di luar WhatsApp.

Di luar tambahan fitur seperti voice call dan akses dari desktop, layanan pesan instan yang memulai hidup sebagai alat update status ini bisa dibilang setia pada bentuknya yang lama.

Pun demikian dengan jalur pemasukan WhatsApp yang dari dulu hanya bersumber dari iuran tahunan.

Namun kini telah terjadi perubahan. Pada 18 Januari 2016, WhatsApp mengumumkan bakal mencabut biaya pemakaian sebesar 1 dollar AS per tahun. Artinya, layanan ini sepenuhnya tersedia gratis.

Kebijakan tersebut sudah mulai berlaku tak lama kemudian, pada 20 Januari 2016, kala pengguna WhatsApp disambut pemberitahuan bahwa aplikasi pesan instan itu bisa dipakai secara cuma-cuma "seumur hidup".

Uang dari mana?

Dengan memangkas iuran tahunan, WhatsApp kehilangan satu-satunya sumber pendapatan. Lalu dari mana perusahaan ini akan mendapatkan uang?

Lagi-lagi, Koum menampik kekhawatiran bahwa pihaknya bakal melakukan sesuatu yang berpotensi mengganggu kenyamanan pengguna.

"Orang-orang mungkin bertanya bagaimana kami menjalankan WhatsApp tanpa menarik biaya langganan, atau khawatir bahwa pengumuman ini bakal menandai kehadiran iklan pihak ketiga," tulis Koum dalam sebuah posting blog.

"Jawabannya adalah tidak akan ada iklan."

Sebagai ganti pemasukan yang hilang, WhatsApp menyatakan bakal menjajaki kemungkinan menawarkan sejumlah layanan berbayar untuk pengguna korporat.

Layanan ini nantinya bisa dipakai oleh entitas bisnis untuk berkomunikasi dengan pelanggan masing-masing tanpa perlu membayar biaya tambahan lain.

Koum mencontohkan komunikasi dalam proses verifikasi transaksi antara konsumen dengan bank, lalu bisa juga pemberitahuan tertundanya penerbangan dari maskapai ke calon penumpang.

"Semua pesan tersebut sekarang disampaikan melalui jalur lain, yakni lewat telepon dan SMS. Kami ingin menguji aneka tools baru untuk membuat ini lebih mudah dilakukan lewat WhatsApp," terang Koum. Masih belum jelas seperti apa nantinya layanan bisnis yang dimaksud.

Mengapa WhatsApp mengganti model bisnis yang tadinya menggarap konsumen jadi menasar segmen korporat?

Perusahaan yang didirikan pada 2009 ini rupanya merasa bahwa penarikan biaya langganan seperti yang dijalankannya selama bertahun-tahun kurang bisa diandalkan untuk mengisi kas perusahaan.

Sebabnya, menurut Koum, adalah banyaknya pengguna WhatsApp yang tak punya kartu debit atau kredit untuk membayar biaya langganan.

Alhasil, basis pengguna WhatsApp yang kini sudah mencapai kisaran 990 juta, berlipat dua semenjak akuisisi oleh Facebook- pun tak bisa dimanfaatkan dengan optimal. Hingga kemudian diputuskanlah untuk menghapus biaya tersebut sama sekali. Saat berbicara di Konferensi Digital Life Design di Munich, Jerman, awal minggu ini,

Koum menjelaskan bahwa WhatsApp mencoba menerapkan model bisnis "commercial-participation", sebuah bentuk evolusi model "freemium" di mana hanya pengguna tertentu saja yang dikenai biaya layanan.

Dalam hal ini, para pengguna itu adalah kalangan korporat yang bersedia membayar lebih untuk fitur premium atau volume yang lebih besar.

Pihak Facebook sendiri selaku pemilik WhatsApp tak menerapkan "deadline" tertentu bagi layanan chatting tersebut untuk mulai mendulang pendapatan besar.

Dengan demikian, menurut Koum, WhatsApp bisa fokus mendorong pertumbuhan pengguna tanpa takut kehilangan pelanggan yang tak bisa membayar, sambil sekaligus mencoba-coba peruntungan di ranah baru.

#### Keamanan terburuk

Langkah WhatsApp memasuki segmen korporat menimbulkan sedikit kekhawatiran soal keamanan informasi sensitif yang banyak beredar di bidang ini. Apalagi WhatsApp bisa dibilang belum memiliki reputasi mumpuni di bidang keamanan data.

Malah sebaliknya, perusahaan ini kerap tertimpa persoalan sekuriti.

Sebagai contoh, belum lama ini, pada September 2015, muncul kabar bahwa client WhatsApp versi web (desktop) memiliki celah berbahaya yang bisa dimanfaatkan untuk memasang malware di komputer pengguna.

Sebelumnya, sekitar pertengahan tahun lalu, WhatsApp dinobatkan jadi perusahaan teknologi terburuk dalam hal pengamanan privasi pengguna oleh The Electronic Frontier Foundation (EFF).

WhatsApp dinilai tak menerapkan praktik-praktik terbaik menurut standar keamanan industri teknologi. WhatsApp turut dipandang kurang transparan dalam menangani beberapa isu terkait privasi, misalnya penghapusan konten tertentu yang diimbau oleh pemerintah dan kebijakan terkait penyimpanan data pengguna.

Soal ini, WhatsApp menyatakan bahwa tiap pesan yang dikirim melalui layanannya dilindungi oleh enkripsi. WhatsApp juga tak menyimpan pesan pengguna di dalam server miliknya, dengan pengecualian pesan belum terkirim yang akan disimpan di server dengan waktu maksimal 30 hari sebelum dihapus.

Namun langkah-langkah di atas tak menjamin seratus persen bahwa informasi pengguna yang tersimpan dan dikirim lewat WhatsApp tak bakal diintip oleh pihak lain.

Laporan EFF menggarisbawahi absennya kebijakan yang mengatur prosedur penyerahan data pengguna ke pihak otoritas seperti pemerintah sebuah negara di dalam policy WhatsApp. Data pengguna WhatsApp, misalnya, bisa diminta tanpa memerlukan surat perintah dari pengadilan. Pengguna pun tak diberitahu apabila ada permintaan data dari pihak tertentu.

Terlebih, perusahaan yang menaungi WhatsApp, Facebook, dikenal memiliki sejarah mematuhi permintaan data pengguna oleh otoritas.

Urusan sekuriti data sendiri ibarat pedang permata dua. Si satu sisi, otoritas sebuah negara membutuhkan informasi untuk melacak hal-hal terkait keamanan seperti jaringan teroris yang kerap aktif berkomunikasi di media sosial, di balik tameng perlindungan enkripsi dan privasi.

Di sisi lain, pengamanan WhatsApp harus melindungi informasi-informasi sensitif yang tersimpan dan beredar lewat layanan tersebut, yang bisa saja dibobol dan dipakai untuk keperluan lain oleh pihak tertentu, misalnya hacker.

WhatsApp selaku penyedia layanan pesan instan terbesar pastilah sering digunakan untuk bertukar informasi penting oleh hampir semiliar pemakainya.

Sebelum kalangan perbankan dan maskapai ikut nimbrung sekalipun, WhatsApp sudah lazim dipakai ngobrol oleh tokoh penting penyelenggara pemerintahan, termasuk mungkin di Indonesia.

Basis pengguna WhatsApp berikut data-data mereka yang sangat besar jumlahnya boleh jadi akan menjadi magnet abadi yang akan selalu mengundang kekhawatiran bakal dieksploitasi, entah oleh hacker, otoritas, ataupun pengiklan.

<b>ID : 3</b>
<b>Judul Berita : Memulihkan Anak dari Rasa Trauma</b>
<p><b>Ringkasan <i>Textrank</i> :</b></p> <p>KOMPAS.com - Berbagai peristiwa dalam hidup kita, mulai dari bencana alam, kriminalitas, hingga meninggalnya orang terdekat, dapat menimbulkan rasa trauma. Tak terkecuali pada anak-anak.</p> <p>"Anak-anak Indonesia memang sering menghadapi peristiwa traumatik. Bahkan, bangsa kita adalah bangsa yang trauma," kata Psikolog Forensik Nathanael EJ Sumampauw dalam diskusi bersama Forum Ngobrol Bareng Sahabat (Ngobras) di Jakarta (19/1/16).</p> <p>Bukan hanya peristiwa langsung, paparan informasi bertubi-tubi mengenai suatu tragedi juga dapat menyebabkan terjadinya trauma sekunder.</p> <p>"Trauma sekunder itu timbul walau seseorang tidak merasakan langsung atau tidak hadir dalam peristiwa itu tapi efeknya dirasakan," ujar psikolog yang biasa disapa Nael ini.</p> <p>Ia menjelaskan, anak-anak dan orangtua hidup di tempat yang sama, sehingga apa yang terjadi di sekitar kita juga akan dirasakan oleh anak.</p> <p>Trauma pada anak, menurut psikolog Sani Budianti Hermawan, M.Si, bisa terjadi ketika berita yang muncul menimbulkan ketakutan berlebih.</p> <p>"Misalnya melihat gambar korban tanpa sensor.</p> <p><b>Ringkasan Ahli :</b></p> <p>"Anak-anak Indonesia memang sering menghadapi peristiwa traumatik. Trauma pada anak, menurut psikolog Sani Budianti Hermawan, M.Si, bisa terjadi ketika berita yang muncul menimbulkan ketakutan berlebih.</p> <p>Tidak semua anak menunjukkan perilaku yang sama dalam menghadapi kejadian traumatik. Ada yang menunjukkan gejala penghindaran (avoidance), yakni menghindari segala sesuatu yang berkaitan dengan trauma yang ia alami.</p> <p>Gejala kedua adalah mengingat-ingat atau mengulang kejadian yang sudah berlalu (reexperiencing).</p> <p>Lalu yang ketiga adalah ketergugahan fisik yang berlebihan (hyper arousal), misalnya takut mendengar suara keras, dan sebagainya.</p> <p>"Yang harus dikenali orangtua adalah perubahan sikap yang signifikan dari anak. Dengan mengenali perubahan itu orangtua bisa melakukan deteksi dini agar dampak trauma tidak terbawa sampai anak dewasa. Pada anak yang mengalami trauma, ekspresikan kasih sayang dari orangtua, misalnya dengan memberi pelukan, tersenyum, dan mengajaknya beraktivitas.</p>

Bantu anak untuk melakukan rutinitasnya, misalnya kembali bersekolah, bermain bola, atau mengikuti les. Sementara itu, tanda-tanda anak sudah terlepas dari trauma adalah anak sudah bisa bermain, bersekolah, bergaul, atau menikmati hari-harinya dengan baik.

**Teks Asli :**

KOMPAS.com - Berbagai peristiwa dalam hidup kita, mulai dari bencana alam, kriminalitas, hingga meninggalnya orang terdekat, dapat menimbulkan rasa trauma. Tak terkecuali pada anak-anak.

"Anak-anak Indonesia memang sering menghadapi peristiwa traumatik. Bahkan, bangsa kita adalah bangsa yang trauma," kata Psikolog Forensik Nathanael EJ Sumampauw dalam diskusi bersama Forum Ngobrol Bareng Sahabat (Ngobras) di Jakarta (19/1/16).

Bukan hanya peristiwa langsung, paparan informasi bertubi-tubi mengenai suatu tragedi juga dapat menyebabkan terjadinya trauma sekunder.

"Trauma sekunder itu timbul walau seseorang tidak merasakan langsung atau tidak hadir dalam peristiwa itu tapi efeknya dirasakan," ujar psikolog yang biasa disapa Nael ini.

Ia menjelaskan, anak-anak dan orangtua hidup di tempat yang sama, sehingga apa yang terjadi di sekitar kita juga akan dirasakan oleh anak.

Trauma pada anak, menurut psikolog Sani Budianti Hermawan, M.Si, bisa terjadi ketika berita yang muncul menimbulkan ketakutan berlebih.

"Misalnya melihat gambar korban tanpa sensor. Ini bisa menghantui anak sehingga sulit tidur atau bahkan mengganggu konsentrasinya," kata Sani saat dihubungi Kompas.com (17/1/16).

Reaksi trauma, menurut Nael, adalah hal yang wajar. Yang tidak wajar adalah pengalamannya.

Tidak semua anak menunjukkan perilaku yang sama dalam menghadapi kejadian traumatik. Ada yang menunjukkan gejala penghindaran (avoidance), yakni menghindari segala sesuatu yang berkaitan dengan trauma yang ia alami. Misalnya, tidak mau melewati lokasi suatu kejadian.

Gejala kedua adalah mengingat-ingat atau mengulang kejadian yang sudah berlalu (reexperiencing). Lalu yang ketiga adalah ketergugahan fisik yang berlebihan (hyper arousal), misalnya takut mendengar suara keras, dan sebagainya.

"Yang harus dikenali orangtua adalah perubahan sikap yang signifikan dari anak. Tidak selalu anak menjadi pasif atau menarik diri. Ada juga anak trauma yang justru menjadi aktif atau agresif," ujar Nael.

Perubahan sikap yang mungkin terjadi misalnya anak yang tadinya tidak pernah mengompol sekarang jadi sering ngompol, atau anak terus "nempel" dengan orangtua, atau gangguan konsentrasi belajar.

Yang bisa membedakan perubahan itu hanya orangtua atau orang terdekat. Dengan mengenali perubahan itu orangtua bisa melakukan deteksi dini agar dampak trauma tidak terbawa sampai anak dewasa.

#### Atasi trauma

Bila anak menunjukkan tanda trauma, orangtua perlu menjelaskan pada anak mengenai kejadian traumatik yang dialaminya. Ajak anak berdialog untuk menggali sejauh mana pemahaman anak dan bagaimana perasaannya.

"Batasi paparan bermuatan kekerasan lebih lanjut melalui media massa. Risiko anak trauma juga lebih besar jika ia tinggal dengan orang dewasa yang memiliki reaksi berlebihan terhadap peristiwa itu," ujar Nael.

Meski anak perlu dijelaskan mengenai peristiwa tersebut, tapi orangtua juga sebaiknya menumbuhkan optimisme dan harapan pada anak. Misalnya untuk peristiwa kriminalitas atau terorisme, ceritakan aksi heroik polisi yang dengan cepat melumpuhkan terorisnya atau relawan yang menolong korban bencana alam.

Pada anak yang mengalami trauma, ekspresikan kasih sayang dari orangtua, misalnya dengan memberi pelukan, tersenyum, dan mengajaknya beraktivitas.

Bantu anak untuk melakukan rutinitasnya, misalnya kembali bersekolah, bermain bola, atau mengikuti les. "Rutinitas akan memberikan harapan pada anak akan hari esok," kata dosen di Fakultas Psikologi Universitas Indonesia ini.

Sementara itu, tanda-tanda anak sudah terlepas dari trauma adalah anak sudah bisa bermain, bersekolah, bergaul, atau menikmati hari-harinya dengan baik.

"Itu adalah indikator anak sudah adaptif terhadap kejadian yang dialaminya. Berarti ia cukup tangguh (resilience) melampaui masa sulit," ujarnya.