

Penggunaan Metode TF-IDF Dalam Menentukan Ringkasan Berita Pada Surat Kabar Kompas

Rein Hotlayn Johanes Silitonga¹, Margaretha Yohanna², Harlen Gilbert Manullang³

^{1,2,3}Fakultas Ilmu Komputer, Universitas Methodist Indonesia

Info Artikel

Histori Artikel:

Received, 30 Mei 2026

Revised, 14 Juli 2026

Accepted, 27 Juli 2026

Keywords:

Ringkasan,
Berita Online,
Kompas,
TF-IDF.

ABSTRAK

Perkembangan teknologi informasi mendorong perubahan dalam cara masyarakat mengakses berita, di mana media *online* seperti Kompas menjadi salah satu sumber utama informasi. Namun, narasi yang panjang dan berlebihan dalam penyajian berita seringkali menyulitkan pembaca untuk menangkap inti informasi secara cepat. Untuk mengatasi permasalahan ini, penelitian ini menerapkan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) dalam merangkum berita online Kompas secara otomatis. Proses *text mining* yang dilakukan meliputi *tokenizing* (menghilangkan tanda baca), *filtering* (menghapus kata-kata tidak penting), dan *stemming* (mengubah kata ke bentuk dasar), yang kemudian dilanjutkan dengan pemberian bobot kalimat berdasarkan metode TF-IDF untuk menghasilkan ringkasan yang merepresentasikan isi utama berita. Sistem dirancang menggunakan bahasa pemrograman Python sebagai *backend* dan HTML sebagai *frontend*, sehingga memudahkan pengguna dalam mengakses dan menggunakan aplikasi peringkasan berita. Hasil pengujian menunjukkan bahwa metode TF-IDF mampu menghasilkan ringkasan dengan rata-rata panjang 53,32% dari teks asli dan waktu pemrosesan rata-rata 0,017 detik. Temuan ini membuktikan bahwa TF-IDF efektif dalam menyederhanakan informasi tanpa mengurangi esensi utama berita. Dengan demikian, metode ini dapat dijadikan solusi yang andal untuk membantu pembaca memperoleh informasi secara ringkas, cepat, dan efisien dalam era digital yang serba cepat saat ini.

This is an open access article under the [CC BY-SA](#) license.



Penulis Koresponden:

Rein Hotlayn Johanes Silitonga,
Fakultas Ilmu Komputer,
Universitas Methodist Indonesia, Medan,
Jl. Hang Tuah No.8, Medan - Sumatera Utara.
Email: silitongarein1@gmail.com

1. PENDAHULUAN

Berita merupakan fakta ataupun opini yang membuat banyak orang merasa tertarik untuk mengetahuinya. Berita dapat diperoleh dengan berbagai media seperti koran, surat kabar, televisi, internet dan lain-lain. Pada saat ini, media yang paling sering digunakan untuk memperoleh berita adalah internet. Salah satunya media surat kabar Kompas. Berita *online* Kompas merupakan salah satu sumber informasi yang ada di internet dengan jumlah yang sangat banyak serta topik yang beragam. Topik berita dapat meliputi mengenai politik, ekonomi, olahraga, teknologi, dan masih banyak lagi.

Berita *online* Kompas menyajikan berita-berita dengan keterangan informasi yang berlebihan, hal ini disebabkan karena banyaknya narasi berita yang dibuat agar dapat memenuhi halaman berita. Masalahnya tidak semua pembaca berita *online* Kompas yang akan tertarik membaca seluruh isi berita. Membaca seluruh isi berita *online* Kompas yang memiliki narasi berlebihan memakan waktu untuk menemukan informasi utama berita. Oleh karena itu perlu dilakukan rangkuman atau ringkasan dari berita *online*

Kompas. Ringkasan merupakan suatu teks singkat yang dihasilkan dari kumpulan teks panjang namun tetap menyimpan informasi penting dari teks asalnya. Dengan adanya ringkasan akan membantu dalam membaca dengan waktu yang lebih cepat dibandingkan dengan harus membaca informasi secara utuh [1]. Selain itu, ringkasan juga membantu dalam mengabaikan informasi yang tidak relevan tanpa harus kehilangan makna dari informasi yang dibaca.

Ringkasan otomatis dapat dilakukan dengan teknik *text mining*. Salah satu metode yang dapat digunakan untuk melakukan peringkasan secara otomatis adalah metode *Term Frequency Inverse Document Frequency* (TF-IDF). Metode TF-IDF merupakan metode yang digunakan menentukan seberapa jauh keterhubungan kata (*term*) terhadap dokumen dengan memberikan bobot setiap kata [2]. Metode TF-IDF ini menggabungkan dua konsep yaitu frekuensi kemunculan sebuah kata di dalam sebuah dokumen serta inverse frekuensi dokumen yang mengandung kata tersebut [3].

Penelitian terdahulu yang dilakukan oleh [4] dengan judul “Penerapan Text Mining Dengan Menggunakan Algoritma TF-IDF Untuk Klasifikasi Genre Novel” menyimpulkan bahwa algoritma TF-IDF melakukan klasifikasi dengan membandingkan nilai kemiripan teks dengan data *resorce*. pengujian sistem dilakukan dengan mengambil 10 sinopsis novel dan menghasilkan nilai akurasi 100%. Penelitian lain yang dilakukan oleh [5] dengan judul “Peringkasan Teks Artikel Ilmiah Berbahasa Indonesia dengan Metode Pembobotan Kalimat” menyimpulkan Pembobotan Kalimat dengan metode TF-IDF mampu menghasilkan ringkasan artikel ilmiah yang dipotong sebesar 50% dari jumlah kata artikel.

Berdasarkan keberhasilan penelitian terdahulu, maka penelitian ini menguraikan tentang bagaimana mengurangi kata pada berita Kompas menggunakan metode TF-IDF. Proses yang dilakukan *text mining* adalah *tokenizing* (menghilangkan tanda baca), *filtering* (menghilangkan kata tidak penting) dan *stemming* (mengubah kata turunan menjadi kata dasar). Proses yang dilakukan TF-IDF adalah membandingkan setiap kalimat topik yang telah dihasilkan dari proses *text mining* dengan menghitung bobotnya berdasarkan algoritma TD-IDF

2. METODE PENELITIAN

2.1 Text Mining

Text mining merupakan suatu teknik untuk mengolah dan mengesktraksi pola dari suatu data menjadi sebuah pengetahuan dan informasi yang bermanfaat[6]. *Text mining* mengubah data yang tidak terstruktur menjadi terstruktur dalam bentuk teks [7]. Perbedaan mendasar dari *text mining* dan data mining terletak pada sumber data yang digunakan. Pada *data mining* data yang diekstrak berasal dari pola-pola tertentu dan terstruktur, sedangkan *text mining* sumber data yang digunakan berasal dari teks yang relatif tidak terstruktur karena menggunakan tata Bahasa manusia atau biasa disebut (*natural language*). Hasil dari proses tersebut dapat dipergunakan untuk menggali informasi yang lebih bermanfaat, contohnya dapat dipergunakan sebagai mendapatkan ringkasan kata-kata dari sebuah dokumen [8]. Secara umum basis data didesain untuk program dengan tujuan melakukan pemrosesan secara otomatis, sedangkan teks ditulis untuk dibaca langsung oleh manusia[9].

2.2 Peringkasan Teks

Peringkasan teks merupakan sebuah proses untuk meringkas dokumen teks dengan tidak mengurangi arti penting dari isi dokumen teks secara keseluruhan. Meringkas dokumen teks secara umum dibatasi dengan panjang kalimat kurang lebih dua ratus kata dari keseluruhan dokumen teks yang tersedia. Peringkasan data otomatis adalah bagian dari pembelajaran mesin dan penggalian data. Gagasan utama peringkasan adalah untuk menemukan subset data yang berisi "informasi" dari seluruh set [10].

2.3 TF_IDF

Pembobotan TF-IDF didapat dari menghitung jumlah dari sebuah term pada dokumen hal ini disebut sebagai *term frequency* (TF) dan selanjutnya adalah jumlah dari term yang muncul pada sebuah inverse document yang biasa disebut *inverse document frequency* (IDF) [11][12].

1. Case Folding

Case Folding biasa disebut penyeragaman kata dengan cara mengubah seluruh kata menjadi huruf kecil (*lowercase*). Hanya huruf a sampai z yang dapat diterima karakter selain huruf dihilangkan[13], [14].

Terdapat juga kata-kata tertentu yang harus sesuai dengan kaidah yang tidak bisa dilakukan

penyeragaman kata seperti kata lembaga atau institusi yang selalui diawali huruf kapital dan juga nama gelar seperti halnya ST, M.Psi dan lain sebagainya. Tergantung dari sumber data yang digunakan untuk diproses.

2. Tokenizing

Adalah suatu tahapan pemotongan *string* kata berdasarkan penyusunan kata tersebut.

3. Filtering

Filtering adalah pengambilan kata-kata penting dari hasil *Tokenizing* atau biasa disebut pengeliminasi sebuah kata-kata sesuai dengan kaidahnya. Algoritma *stop-word removal* adalah salah satu yang digunakan untuk melakukan tahapan *filtering*.

4. Stemming

Stemming adalah suatu proses untuk memecah suatu varian-varian kata menjadi kata dasar sesuai dengan kata yang sedang diproses. Jika kata yang diproses adalah Bahasa Indonesia untuk memecah varian kata menjadi kata dasar harus sesuai dengan aturan Bahasa Indonesia.

5. Analyzing

Analyzing adalah suatu tahapan menganalisa data teks yang sedang diproses untuk menentukan kemiripan antar dokumen teks salah satu metode yang digunakan adalah *cosine similarity*.

Pembobotan dapat diperoleh berdasarkan jumlah kemunculan suatu *term* dalam sebuah dokumen *term frequency* (tf) dan jumlah kemunculan *term* dalam koleksi dokumen *inverse document frequency* (idf) [15]. Bobot suatu istilah semakin besar jika istilah tersebut sering muncul dalam suatu dokumen dan semakin kecil jika istilah tersebut muncul dalam banyak dokumen [16]. Nilai idf sebuah term (kata) dapat dihitung menggunakan persamaan sebagai berikut: D adalah jumlah dokumen yang berisi *term* (t) dan dfi adalah jumlah kemunculan (frekuensi) term terhadap D. Adapun algoritma yang digunakan untuk menghitung bobot (W) masing masing dokumen terhadap kata kunci (*query*) [17], yaitu:

$$TF(t,d) = \text{Jumlah kemunculan kata } t \text{ dalam dokumen} - \text{Jumlah kata ringkas} \quad (1)$$

$$IDF(t,D) = \frac{TF(t,d)}{\text{Jumlah kata dalam dokumen}} \quad (2)$$

$$W_{d,t} = IDF_t * 100 \quad (3)$$

Keterangan :

d = dokumen ke-d

t = term ke-t dari kata kunci

tf = term frekuensi/frekuensi kata

W = bobot dokumen ke-d terhadap ter m ke-t.

2.4 Pengumpulan Data

Pengumpulan data dengan mengidentifikasi dan membuat alur perancangan yang akan dilaksanakan, agar dalam proses pencarian data tidak terjadi penyimpangan dalam mengemukakan tujuan yang ingin dicapai, penulis menggunakan metode pengumpulan data observasi. Adapun observasi dilakukan dengan melihat berita *online* olahraga pada link <https://kompas.com/>. Pada penelitian ini diambil 15 berita yang dijadikan data penelitian. Adapun semua data berita olahraga yang berhasil dikompulkan dapat dilihat seperti pada Tabel1.

Tabel 1 Semua Data Berita Olahraga

No	Kategori	Sampel Informasi Berita
1	Bola	Persita mampu memaksakan hasil imbang 3-3 saat melawan Persib. Pendekar Cisadane bangkit setelah sempat tertinggal 02 dan 2-3. Pertandingan pekan ke-31 Liga 1 2023-2024 antara Persita vs Persib di Stadion Kaptan I Wayan Dipta, Gianyar, Bali, Senin (15/4/2024). berkesudahan dengan skor 3-3. Persita kecolongan dan ketinggalan 0-2 pada babak pertama akibat kemasukan gol Febri Hariyadi (5') dan Ciro Alves (24'). Drama terjadi ketika Persita menggebrak pertahanan Persib pada awal babak kedua. Pendekar Cisadane bisa menyamakan skor jadi 22 lewat gol Irsyad Maulana (49') dan penalti Ramiro Fergonzi (58').

No	Kategori	Sampel Informasi Berita
2	Bola	Persib Bandung masih belum bisa menyegel tempat di Championship Series Liga 1 2023-2024. Penyebabnya, Persib kehilangan poin dalam dua pertandingan terakhir. Mulai ditahan imbang Bhayangkara FC (0-0) dan paling menyedihkan dipaksa berbagi angka dengan Persita Tangerang (3-3), Senin (15/4/2024) kemarin, setelah sempat unggul 2-0 dan 3-2 sebelumnya. Persib memang masih berada di peringkat kedua klasemen dengan 56 poin, berselisih sembilan poin dari tim posisi lima PSIS Semarang yang belum memainkan laga pekan ke-31.
...	...	...
15	Bola	Penyerang Timnas Malaysia, Faisal Halim, terancam pensiun dini seussai disiram air keras oleh orang tak dikenal. Faisal Halim mengalami insiden disiram air keras saat dirinya tengah berada di pusat perbelanjaan di Kota Damansara, Distrik Petaling, Selangor, Malaysia, pada Minggu (5/5/2024). Menurut laporan Stadium Astro, Faisal Halim saat ini tengah mendapatkan perawatan intensif di ruang ICU karena menderita luka bakar serius. Jurnalis Astro Arena, Zulhelmi Zainal Azam, bahkan menyebut Faisal Halim terancam gantung sepatu dini karena efek luka bakar di tubuhnya. Baca juga: Pengamat Malaysia Sebut Timnas U23 Indonesia Main Tanpa Rasa Takut "Faisal Halim dilaporkan terancam berisiko pensiun lebih awal dari sepak bola karena luka dari penyiraman air keras itu," kata Zulhelmi Zainal Azam, dikutip dari Stadium Astro.

2.5 Penerapan TF-IDF

Metode *text mining* digunakan pada perancangan aplikasi ringkasan berita surat kabar Kompas dengan melakukan pengklasifikasian untuk menemukan pola yang berbeda di dalam data tekstual yang berjumlah besar. Berdasarkan pada Tabel 1, untuk mempermudah proses *text mining* secara manual, maka penelitian ini hanya mengambil beberapa baris berita untuk diringkas dari berita nomor 1 pada Tabel 1. Adapun sampel berita pada untuk perhitungan manual *text mining* seperti pada Tabel 2.

Tabel 2 Contoh Sampel Beberapa Baris Berita Untuk *Text Mining*

No	Sampel Informasi Berita
1	Persita mampu memaksakan hasil imbang 3-3 saat melawan Persib. Pendekar Cisadane bangkit setelah sempat tertinggal 0-2 dan 2-3. Pertandingan pekan ke-31 Liga 1 2023-2024 antara Persita vs Persib di Stadion Kapten I Wayan Dipta, Gianyar, Bali, Senin (15/4/2024). berkesudahan dengan skor 3-3. Persita kecolongan dan ketinggalan 0-2 pada babak pertama akibat kemasukan gol Febri Hariyadi (5') dan Ciro Alves (24'). Drama terjadi ketika Persita menggebrak pertahanan Persib pada awal babak kedua. Pendekar Cisadane bisa menyamakan skor jadi 2-2 lewat gol Irsyad Maulana (49') dan penalti Ramiro Fergonzi (58').

2.5.1 Case Folding

Case Folding biasa disebut penyeragaman kata dengan cara mengubah seluruh kata menjadi huruf kecil (*lowercase*). Berikut proses *case folding* pada kalimat-kalimat yang terdapat didalam berita seperti pada Tabel 3:

Tabel 3 Proses *Case Folding*

Teks Input	Hasil <i>Case Folding</i>
Persita mampu memaksakan hasil imbang 3-3 saat melawan Persib. Pendekar Cisadane bangkit setelah sempat tertinggal 0-2 dan 2-3. Pertandingan pekan ke-31 Liga 1 2023-2024 antara Persita vs Persib di Stadion Kapten I Wayan Dipta, Gianyar, Bali, Senin (15/4/2024). berkesudahan dengan skor 3-3. Persita kecolongan dan ketinggalan 0-2 pada babak pertama akibat kemasukan gol Febri Hariyadi (5') dan Ciro Alves (24'). Drama terjadi ketika Persita menggebrak pertahanan Persib pada awal babak kedua. Pendekar Cisadane bisa menyamakan skor	persita mampu memaksakan hasil imbang 3-3 saat melawan persib. pendekar cisadane bangkit setelah sempat tertinggal 0-2 dan 2-3. pertandingan pekan ke-31 liga 1 2023-2024 antara persita vs persib di stadion kapten i wayan dipta, gianyar, bali, senin (15/4/2024). berkesudahan dengan skor 3-3. persita kecolongan dan ketinggalan 0-2 pada babak pertama akibat kemasukan gol febri hariyadi (5') dan ciro alves (24'). drama terjadi ketika persita menggebrak pertahanan persib pada

Teks Input	Hasil <i>Case Folding</i>
jadi 22 lewat gol Irsyad Maulana (49') dan penalti Ramiro Fergonzi (58').	awal babak kedua. pendekar cisadane bisa menyamakan skor jadi 22 lewat gol irsyad maulana (49') dan penalti ramiro fergonzi (58').

2.5.2 Tokenizing

Adalah suatu tahapan pemotongan *string* kata berdasarkan penyusunan kata tersebut. Berikut proses *Tokenizing* pada kalimat-kalimat yang terdapat didalam berita seperti pada Tabel 4:

Tabel 4 Proses *Tokenizing*

Hasil <i>Case Folding</i>	Hasil <i>Tokenizing</i>
persita mampu memaksakan hasil imbang 3-3 saat melawan persib. pendekar cisadane bangkit setelah sempat tertinggal 02 dan 2-3. pertandingan pekan ke-31 liga 1 2023-2024 antara persita vs persib di stadion kapten i wayan dipta, gianyar, bali, senin (15/4/2024). berkesudahan dengan skor 3-3. persita kecolongan dan ketinggalan 0-2 pada babak pertama akibat kemasukan gol febri hariyadi (5') dan ciro alves (24'). drama terjadi ketika persita mengebrak pertahanan persib pada awal babak kedua. pendekar cisadane bisa menyamakan skor jadi 22 lewat gol irsyad maulana (49') dan penalti ramiro fergonzi (58').	persita, mampu, memaksakan, hasil, imbang, 3-3, saat, melawan, persib, pendekar, cisadane, bangkit, setelah, sempat, tertinggal, 0-2, dan, 2-3, pertandingan, pekan, ke-31, liga, 1, 2023-2024, antara, persita, vs, persib, di, stadion, kapten, i, wayan, dipta, gianyar, bali, senin, (15/4/2024), berkesudahan, dengan, skor, 3-3, persita, kecolongan, dan, ketinggalan, 0-2, pada, babak, pertama, akibat, kemasukan, gol, febri, hariyadi, (5'), dan, ciro, alves, (24'), drama, terjadi, ketika, persita, mengebrak, pertahanan, persib, pada, awal, babak, kedua, pendekar, cisadane, bisa, menyamakan, skor, jadi, 2-2, lewat, gol, irsyad, maulana, (49'), dan, penalti, ramiro, fergonzi, (58').

2.5.3 Filtering

Tahap *filtering* merupakan tahap menghilangkan kata yang tidak penting (*stoplist*) atau menyimpan kata yang penting (*wordlist*). *Stoplist/stopword* adalah kata-kata yang tidak deskriptif yang dapat dibuang dalam pendekatan *bag-of-words*. Berikut proses *filtering* pada kalimat-kalimat yang terdapat didalam berita seperti pada Tabel 5:

Tabel 5 Proses *Filtering*

Proses <i>Filtering</i>	Hasil <i>Filtering</i>
persita, mampu, memaksakan, hasil, imbang, 3-3, saat, melawan, persib, pendekar, cisadane, bangkit, setelah, sempat, tertinggal, 0-2, dan, 2-3, pertandingan, pekan, ke-31, liga, 1, 2023-2024, antara, persita, vs, persib, di, stadion, kapten, i, wayan, dipta, gianyar, bali, senin, (15/4/2024), berkesudahan, dengan, skor, 3-3, persita, kecolongan, dan, ketinggalan, 0-2, pada, babak, pertama, akibat, kemasukan, gol, febri, hariyadi, (5'), dan, ciro, alves, (24'), drama, terjadi, ketika, persita, mengebrak, pertahanan, persib, pada, awal, babak, kedua, pendekar, cisadane, bisa, menyamakan, skor, jadi, 2-2, lewat, gol, irsyad, maulana, (49'), dan, penalti, ramiro, fergonzi, (58').	persita, memaksakan, imbang, 3-3, saat, melawan, persib, pendekar, cisadane, bangkit, setelah, tertinggal, 0-2, 2-3, pertandingan, pekan, ke-31, liga, 1, 2023-2024, persita, vs, persib, di, stadion, kapten, i, wayan, dipta, gianyar, bali, senin, (15/4/2024), skor, 3-3, persita, kecolongan, 0-2, babak, pertama, gol, febri, hariyadi, (5'), ciro, alves, (24'), persita, mengebrak, babak, kedua, skor, 2-2, lewat, gol, irsyad, maulana, (49'), penalti, ramiro, fergonzi, (58').

2.5.4 Stemming.

Tahap *stemming* merupakan tahap mengubah kata imbuhan menjadi kata dasar atau dengan kata lain adalah mengembalikan kata ke dalam bentuk aslinya. Berikut proses *stemming* pada kalimat-kalimat yang terdapat didalam berita seperti pada Tabel 6:

Tabel 6 Proses *Stemming*

Proses <i>Stemming</i>	Hasil <i>Stemming</i>
persita, memaksakan , imbang, 3-3, saat, melawan, persib, pendekar, cisadane, bangkit, setelah, tertinggal, 0-2, 2-3, pertandingan, pekan, ke-31, liga, 1, 2023-2024, persita, vs, persib, di, stadion, kapten, i, wayan, dipta, gianyar, bali, senin, (15/4/2024), skor, 3-3, persita, kecolongan, 0-2, babak, pertama, gol, febri, hariyadi, (5'), ciro, alves, (24'), persita, menggebrak, babak, kedua, skor, 2-2, lewat, gol, irsyad, maulana, (49'), penalti, ramiro, fergonzi, (58').	persita, maksa, imbang, 3-3, saat, melawan, persib, pendekar, cisadane, bangkit, setelah, tertinggal, 0-2, 2-3, pertandingan, pekan, ke-31, liga, 1, 2023-2024, persita, vs, persib, di, stadion, kapten, i, wayan, dipta, gianyar, bali, senin, (15/4/2024), skor, 3-3, persita, kecolongan, 0-2, babak, pertama, gol, febri, hariyadi, (5'), ciro, alves, (24'), persita, menggebrak, babak, kedua, skor, 2-2, lewat, gol, irsyad, maulana, (49'), penalti, ramiro, fergonzi, (58').

2.5.5 Analyzing

Analyzing adalah suatu tahapan menganalisa data teks yang sedang diproses untuk menentukan kemiripan antar dokumen teks. Berikut adalah hasil ringkasan dengan *text mining* yang telah dikembalikan ke format aslinya seperti pada Tabel 7.

Tabel 7 Hasil *Text Mining*

Hasil <i>Text Mining</i>
Persita maksa imbang 3-3 saat lawan Persib. Pendekar Cisadane bangkit setelah tinggal 0-2 menjadi 2-3. Pertandingan pekan ke-31 Liga 1 2023-2024 Persita vs Persib di Stadion Kapten I Wayan Dipta, Gianyar, Bali, Senin (15/4/2024) skor 3-3. Persita kecolongan 0-2 babak pertama gol Febri Hariyadi (5') Ciro Alves (24'). Persita menggebrak babak kedua, skor 2-2 lewat Gol Irsyad Maulana (49') penalti Ramiro Fergonzi (58').

Berdasarkan pada Tabel 7, Selanjutnya menentukan pengukuran presentasi bobot pengurangan TF- IDF yang hasilnya seperti pada Tabel 8.

Tabel 8 Hasil Pembobotan Ringkasan

No	Jumlah Karakter Asli	Jumlah Karakter Ringkas	TF (t,d)	IDF (t,D)	W d,t (%)
1	607	420	187	0.3081	30.81%
2	520	361	159	0.3058	30.58%
3	465	258	207	0.4452	44.52%
4	549	290	259	0.4718	47.18%
5	874	386	488	0.5584	55.84%
6	531	308	223	0.4200	42.00%
7	798	433	365	0.4574	45.74%
8	633	226	407	0.6430	64.30%
9	819	237	582	0.7106	71.06%
10	683	238	445	0.6515	65.15%
11	674	234	440	0.6528	65.28%
12	676	349	327	0.4837	48.37%
13	795	364	431	0.5421	54.21%
14	892	282	610	0.6839	68.39%
15	823	277	546	0.6634	66.34%

3. HASIL DAN PEMBAHASAN

3.1 Pengujian Sistem

Aplikasi peringkasan berita *online* Kompas ini dilengkapi dengan tampilan yang bertujuan untuk memudahkan penggunaannya. Fungsi dari *interface* (antarmuka) ini adalah untuk memberikan *input* dan menampilkan *output* dari aplikasi. Pada aplikasi ini memiliki *interface* yang digunakan oleh *user* untuk melakukan proses peringkasan teks berita *online* Kompas.

3.1.1 Pengujian Simulasi Ringkasan Teks

Setelah proses tampilan aplikasi, penelitian ini akan mensimulasikan skenario dalam melakukan ringkasan teks berita *online* Kompas. Adapun sampel berita yang digunakan seperti pada Tabel 1 pada bab sebelumnya. Sampel yang diambil terdiri dari 15 berita *online* Kompas.

1. Pengujian Ringkasan 1

Berikut adalah hasil pengujian dari ringkasan berita online Kompas sampel pertama seperti pada Gambar 1.

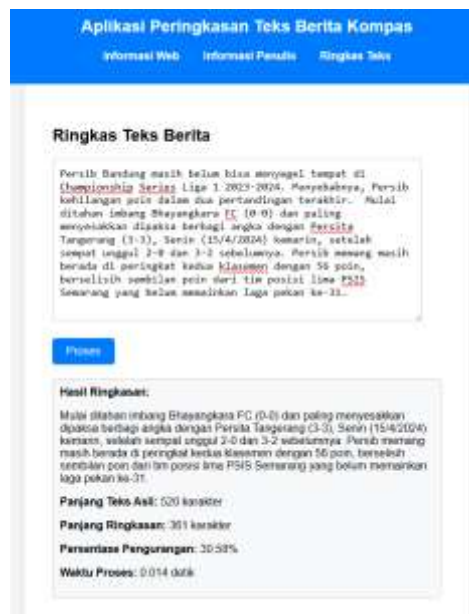


Gambar 1 Tampilan Proses Peringkasan Teks 1

Berdasarkan pada Gambar 1, aplikasi peringkasan teks berita Kompas berhasil melakukan ringkasan terhadap teks dengan panjang 607 karakter dan hasil ringkasannya 420 karakter. Adapun presentasi ringkasan sebanyak 30.81% dengan waktu 0 detik.

2. Pengujian Ringkasan 2

Berikut adalah hasil pengujian dari ringkasan berita online Kompas sampel kedua seperti pada Gambar 2.

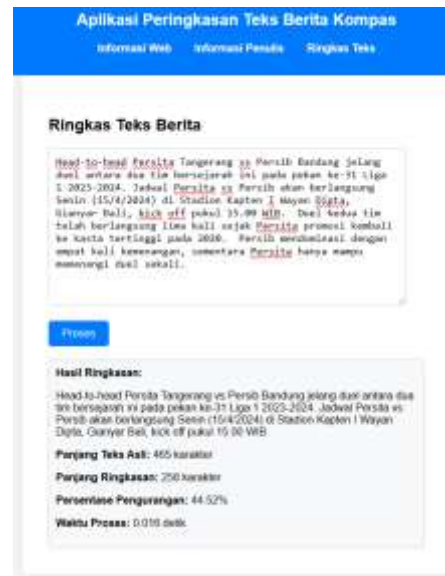


Gambar 2 Tampilan Proses Peringkasan Teks 2

Berdasarkan pada Gambar 2, aplikasi peringkasan teks berita Kompas berhasil melakukan ringkasan terhadap teks dengan panjang 520 karakter dan hasil ringkasannya 361 karakter. Adapun presentasi ringkasan sebanyak 30.58% dengan waktu 0.014 detik.

3. Pengujian Ringkasan 3

Berikut adalah hasil pengujian dari ringkasan berita online Kompas sampel ketiga seperti pada Gambar 3.



Gambar 3 Tampilan Proses Peringkasan Teks 3

Berdasarkan pada Gambar 3, aplikasi peringkasan teks berita Kompas berhasil melakukan ringkasan terhadap teks dengan panjang 465 karakter dan hasil ringkasannya 258 karakter. Adapun presentasi ringkasan sebanyak 44.52% dengan waktu 0.016 detik.

3.2 Pembahasan

Rekap keseluruhan hasil pengujian ringkasan berita *online* Kompas berdasarkan sampel data yang digunakan seperti pada Tabel 9.

Tabel 9 Hasil Rekap Ringkasan Berita

No	Panjang Karakter Awal	Panjang Karakter Ringkasan	Presentasi Ringkasan	Waktu Ringkasan
1	607 karakter	420 karakter	30.81%	0
2	520 karakter	361 karakter	30.58%	0.014
3	465 karakter	258 karakter	44.52%	0.016
4	549 karakter	290 karakter	47.18%	0.016
5	874 karakter	386 karakter	55.84%	0.017
6	531 karakter	308 karakter	42.00%	0.014
7	798 karakter	433 karakter	45.74%	0.020
8	633 karakter	226 karakter	64.30%	0.016
9	819 karakter	237 karakter	71.06%	0.020
10	683 karakter	238 karakter	65.15%	0.035
11	674 karakter	234 karakter	65.28%	0.018
12	676 karakter	349 karakter	48.37%	0.016
13	795 karakter	364 karakter	54.21%	0.019
14	892 karakter	282 karakter	68.39%	0.020
15	823 karakter	277 karakter	66.34%	0.018
Rata-Rata			53.32%	0.017 detik

Berdasarkan dari hasil pengujian pada Tabel 9, metode TF-IDF berhasil melakukan ringkasan teks berita surat kabar Kompas dengan presentasi rata-rata ringkasan 53.32% dan rata-rata waktu 0.017 detik

4. KESIMPULAN

Penentuan ringkasan berita pada surat kabar Kompas berhasil dilakukan menggunakan metode TF-IDF dengan bahasa pemrograman *python* sebagai *backend* dan HTML sebagai *frontend*. Berdasarkan hasil pengujian, metode TF-IDF terbukti efektif dalam melakukan ringkasan teks berita pada surat kabar

Kompas, dengan rata-rata panjang ringkasan 53.32% dari teks asli. Selain itu, metode ini memiliki efisiensi waktu yang cukup baik, dengan rata-rata waktu pemrosesan 0.017 detik. Hasil ini menunjukkan bahwa TF-IDF dapat digunakan sebagai pendekatan yang andal dalam menyederhanakan informasi berita tanpa mengurangi esensi utama dari isi teks. Metode TF-IDF dapat dijadikan solusi dalam membantu pembaca memperoleh informasi secara ringkas dan cepat.

REFERENSI

- [1] F. Halim, Liliana, and K. Gunadi, "Ringkasan Ekstraktif Otomatis pada Berita Berbahasa Indonesia Menggunakan Metode BERT," *J. Infra*, vol. 10, no. 1, pp. 162–168, 2022.
- [2] M. P. Simatupang and D. P. Utomo, "Analisa Testimonial Dengan Menggunakan Algoritma Text Mining Dan Term Frequency- Inverse Document Frequence (Tf-Idf) Pada Toko Allmeeart," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 808–814, 2019, doi: 10.30865/komik.v3i1.1697.
- [3] B. Herwijayanti, D. E. Ratnawati, and L. Muflikhah, "Klasifikasi Berita Online dengan menggunakan Pembobotan TF-IDF dan Cosine Similarity," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 1, pp. 306–312, 2019.
- [4] T. Wahyu and D. Sari, "Penerapan Text Mining Dengan Menggunakan Algoritma TF-IDF Untuk Klasifikasi Genre Novel," *Pelita Inform. Inf. dan Inform.*, vol. 10, no. 1, pp. 29–37, 2021.
- [5] D. Fatmalasari and F. R. Lumbanraja, "Peringkasan Teks Artikel Ilmiah Berbahasa Indonesia dengan Metode Pembobotan Kalimat," *J. Pepadun*, vol. 3, no. 3, pp. 314–322, 2022.
- [6] H. G. Simanullang, S. Wijono, B. Kristianto, W. Sulisty, and A. P. Silalahi, "Detecting Cyberbullying Using Machine Learning Approaches," in *2024 Ninth International Conference on Informatics and Computing (ICIC)*, 2024, pp. 1–8. doi: 10.1109/ICIC64337.2024.10956842.
- [7] A. Firdaus and W. I. Firdaus, "Text Mining Dan Pola Algoritma Dalam Penyelesaian Masalah Informasi : (Sebuah Ulasan)," *J. JUPITER*, vol. 13, no. 1, p. 66, 2021.
- [8] A. Hermawan, I. Jowensen, J. Junaedi, and Edy, "Implementasi Text-Mining untuk Analisis Sentimen pada Twitter dengan Algoritma Support Vector Machine," *JST (Jurnal Sains dan Teknol.*, vol. 12, no. 1, pp. 129–137, 2023, doi: 10.23887/jstundiksha.v12i1.52358.
- [9] L. M. Barus, H. G. Simanullang, and A. P. Silalahi, "Classification of English-Language Racial Hate Speech Using Random Forest "," *J. Intell. Comput. Inf. Syst.*, vol. 1, no. 1, pp. 1–17, 2025.
- [10] M. S. Utomo, J. S. Wibowo, and E. N. Wahyudi, "Text Summarization Pada Artikel Berita Menggunakan Vector Space Model Dan Cosine Similarity," *J. Din. Inform.*, vol. 14, no. 1, pp. 11–24, 2022, doi: 10.35315/informatika.v14i1.9163.
- [11] H. Sari, G. L. Ginting, and T. Zebua, "Penerapan Algoritma Text Mining dan TF-IDF untuk Pengelompokan Topik Skripsi pada Aplikasi Repository STMIK Budi Darma," *TIN Terap. Inform. Nusantara*, vol. 2, no. 7, pp. 414–432, 2021.
- [12] D. Septiani and I. Isabela, "Analisis Term Frequency Inverse Document Frequency (TF-IDF) Dalam Temu Kembali Informasi Pada Dokumen Teks," *SINTESIA J. Sist. dan Teknol. Inf. Indones.*, vol. 1, no. 2, pp. 81–88, 2023.
- [13] T. R. Barua, H. G. Manullang, and A. P. Silalahi, "ANALISA ALGORITMA K-MEDOIDS DALAM," vol. 4, no. 2, pp. 1–10, 2024.
- [14] H. G. Simanullang, A. P. Silalahi, and G. P. Pangaribuan, "PENGELOMPOKAN KOMENTAR PENGGUNA JKN MOBILE MENGGUNAKAN METODE K-MEDOIDS," 2025.
- [15] A. F. Lestari and M. Hafiz, "Penerapan Algoritma Apriori Pada Data Penjualan Barbar Warehouse," *INOVTEK Polbeng - Seri Inform.*, vol. 5, no. 1, p. 96, 2020, doi: 10.35314/isi.v5i1.1317.
- [16] A. Z. Z. Abidin and A. Sukmadinata, "Sistem Deteksi Kerusakan pada Sistem Operasi Menggunakan Metode TF - IDF dan Cosine Similarity," *J. Ilm. Inform.*, vol. 8, no. 2, pp. 6–11, 2020.
- [17] R. Robiyanto, N. Nugraha, and I. Apriatna, "Peringkasan Teks Otomatis Berita Menggunakan Metode Maximum Marginal Relevance," *JEJARING J. Teknol. dan Manaj. Inform.*, vol. 4, no. 1, pp. 23–32, 2019, doi: 10.25134/jejaring.v4i1.6712.